

Refining of breast cancer prognosis using digital pathology

Ayat Gamal Lashen

MB.BCH, MSc

Thesis submitted to the University of Nottingham for the Degree of Doctor of Philosophy

January 2024

School of Medicine

ACKNOWLEDGMENT

I am grateful to Allah, the Almighty, who has provided me the strength, wisdom and patience to complete my research and for granting me this opportunity and the determination to see this research through.

I would like to express my gratitude to Professor Emad Rakha for his supervision and the continuous support of my PhD study and related research. His guidance helped me from the very early stage of this research as well as giving me his experiences throughout the work. Without his guidance, this project would not have been possible. His unparalleled supervision, inspiration and huge passion for science have guided me throughout this journey.

Very special thanks to Professor Nigel Mongan. I am much indebted to him for his invaluable advice in scientific discussions, using his precious times to guide me with his fruitful ideas and for providing me unflinching encouragement and support in various ways.

I gratefully acknowledge Dr Andrew Green for his advice, supervision, and crucial contribution. His truly scientist intuition has made him as a constant supporter with limitless ideas and passions in science, I am grateful in every possible way and hope to keep up our collaboration in the future.

My sincere thanks also go to Dr Michael Toss, whose immeasurable knowledge of pathology and his continuous help and support have encouraged me during my PhD Journey.

I would also like to express my gratitude to Professor Srinivasan Madhusudan, Dr Catrin Rutland and DR Jennie Jeyapalan for their scientific support and guidance.

Many thanks are due to Mr. Christopher Nolan, Mr Alex Dean, Miss Holly Nichollas and Miss Holly Brown for their invaluable assistance and all

their spiritual support throughout my study. Without their invaluable help, I would never achieve any progress.

Thank to my parents who give me all the support and encouragement in my life.

I would like to thank everybody who was important to the successful realisation of thesis, as well as expressing my apology that I could not mention personally one by one.

This study was under the financial support of the Ministry of Higher Education Egypt; to whom I am indebted. Many thanks go to my colleagues in Pathology Department, Faculty of Medicine, Menoufia University, Egypt, for their spiritual support.

DEDICATION

Words fail me to express my appreciation to my husband, Ahmed, whose dedication, love, and persistent confidence in me, have taken the load off my shoulder. Your presence and consistent encouragement have been a source of motivation during our life abroad. Although, you have struggled too, it was a great relief to know that you are willing to provide everything you can give to support me during this journey.

To my sons, Yassin, Youssif and Yahea, I will not be this stronger without you as my main inspiration in life. You are the reason why I keep pushing; I keep facing all the struggles, pains and hardships. I love you all so much.

DECLARATION

I declare that this thesis is the result of my own work and has not in this, or any other form, been presented to this, or any other University in support of an application for any degree other than for which I am now a candidate.

Ayat Lashen

January 2024

APPROVAL

This study was approved by the Yorkshire & the Humber - Leeds East Research Ethics Committee (REC Reference: 19/YH/0293) under the IRAS Project ID: 266925. Data collected were fully anonymised.

Under the title: Development of a molecular genetic classification of breast cancer.

ABBREVIATIONS

95% CI	95% confidence interval
AI	Artificial intelligence
AJCC	The American Joint Committee on Cancer
ASCO	American Society of Clinical Oncology
BC	Breast cancer
BCSS	Breast cancer-specific survival
cAMP	Cyclic adenosine 3′,5′-monophosphate
CAP	College of American Pathologists
CISH	Chromogenic *in-situ* hybridisation
CL	Co-localisation
CM	Cumulative matrix
CNN	Convolutional neural network
DCIS	Ductal carcinoma *in situ*
DEGs	Differentially expressed genes
DFS	Disease-free survival
DL	Deep learning
DMFS	Distant metastases free survival
DP	Digital pathology
EMT	Epithelial-to-mesenchymal transition
ER	Estrogen receptor
FDA	Food and Drug Administration
FFPE	Formalin fixed paraffin embedded
FISH	Fluorescent *in-situ* hybridisation
H&E	Haematoxylin and eosin
H.H1	Histone H1
HER2	Human epidermal growth factor receptor 2
HPFs	High power fields
HR	Hormone receptors
HR	Hazard ratio
H-score	Histo-score
ICC	Intra-Class Correlation Coefficient
IHC	Immunohistochemistry
IKWG	International Ki67 in BC Working Group
ILC	Invasive lobular carcinoma
ITH	Intra-tumoural heterogeneity
LM	Light microscopy
LN	LN
LOH	Loss of heterozygosity
LVI	Lymphovascular invasion
MAI	Mitotic activity Index
MCM	Minichromosome maintenance
MCM7	minichromosome maintenance protein7
MI	Mitotic Index
NGS	Nottingham grading system

NPI	Nottingham Prognostic Index
NST	No special type
OS	Overall survival
PCNA	Proliferating cell nuclear antigen
pCR	pathologic complete response
PHH3	Phosphor-histone H3
PLK1	Polo-like kinases 1
PPAR	Peroxisome proliferator-activated receptor
PR	Progesterone receptor
RANKL	Receptor activator of nuclear factor kappa beta
RCPath	The United Kingdom Royal College of Pathologists
RNA-Seq	RNA sequencing
ROI	Region of interest
rRNA	Ribosomal RNA
RS	Recurrence score
SERDs,	Selective Estrogen Receptor Degraders
SERMs	Selective estrogen receptor modulators
SPSS	Statistical Package for Social Sciences
TCGA	Cancer Genome Atlas Program
TILs	Tumour infiltrating lymphocytes
TMA	Tissue microarray
TME	Tumour microenvironment
TN	Triple negative
WB	Western blot
WSI	Whole slide image

Abstract

Background:

Breast cancer (BC) is a heterogeneous disease, necessitating accurate risk stratification for optimal patient management. Currently, BC patients' risk stratifications depend on the clinicopathological parameters, and few clinically validated molecular markers. Using digitalised whole slide images (WSI) in clinical pathology revolutionised the way of reporting BC cases. In addition, the advent of artificial intelligence (AI) enables utilisation and integration of knowledge that is beyond human limits. However, as a newly introduced technology, further exploration of the applications and uses in the clinical setting is of utmost importance. Therefore, this study integrates digital pathology (DP) and AI to refine morphological features and identify tissue biomarkers for improved BC prognosis.

Patients and methods:

This study included multiple large cohorts of primary BCs; The Nottingham cohort and the online available cohorts including The Cancer Genome Atlas (TCGA) BC dataset. The detailed clinicopathological and outcome data was collected for the Nottingham cohort. For the DP part of the study, the haematoxylin and eosin (H&E)-stained slides were histologically reviewed and scanned into WSIs enabling the application of deep learning (DL) models. Annotation of different histological features was performed. For the immunohistochemical part of study, tissue microarrays (TMAs) were constructed. However, for Ki67 and PR assessment, full-face sections were utilised. Mining for biomarkers that have potential prognostic values in BC was performed. This includes histone H1 (H.H1), Polo kinase 1 (PLK1) and minichromosome maintenance complex component 7 (MCM7).

Results:

This study, aimed to assess the accuracy of WSI in interpreting subtle histological features. A cohort of 3089 cases was used to refine the definition of mitosis in BC. Mitotic figures were visually assessed using

light microscope (LM) and compared with their corresponding in WSI. In total,16 875 figures were examined. About half of mitotic figures showed absent hairy-like projection in WSI; however, most of the mitotic figures retained their ragged nuclear border, which distinguished them from mimickers. When the mitotic counts were considered, there was a 20% average reduction rate of mitotic counts in WSI when compared to the same area using LM. A refined mitosis identification criteria improved inter-observer concordance. Using a refined cutoff for mitosis in WSI reduced the variability in mitotic counts between WSI and LM.

To evaluate the frequency and prognostic significance of atypical mitosis, two large cohorts (n=1878) of BC were utilised to score atypical, typical mitotic figures and the ratio between both. This showed that high atypical mitosis was significantly associated with parameters characteristic of aggressive tumour behaviour. Notably, atypical-to-typical mitoses ratio had a prognostic value in BC patients independent of the overall mitotic count. Atypical-to-typical mitoses ratio was able to dichotomise triple negative-BC (TNBC) patients into two distinct prognostic groups, while the overall mitotic count was not. To evaluate the molecular mechanisms of atypical mitosis in BC, differential gene expression (DGE) analysis of TCGA cases was performed. Transcriptomic analysis identified 2494 genes associated with atypical mitosis and they were linked to pathways involved in cancer progression including chromosomal segregation and organisation and regulation of cell cycle.

Despite, the clinical value of Ki67 in BC as a surrogate of proliferation, its adoption into routine practice is limited because no consensus has been reached regarding its optimal scoring method. Therefore, a large number of cases (n=3004) was used to assess Ki67 in BC using manual and AI-assisted methods. The dynamic relationship between Ki67 and mitosis was also assessed. Comparing different methods for Ki67 assessment, quantification of 1000 cells in hotspot showed the most accurate method and showed the highest inter-observer concordance. The potential underlying molecular alterations linked to Ki67 were explored and showed 5489 genes were significantly associated with high *mKI67* gene. These genes are involved in regulation of cell cycle phase transition,

stem cell proliferation and chromosome segregation. There was a strong positive correlation between Ki67 expression using manual scores and automated scores using DL model. In addition, patients with Ki67 spatial distribution (co-localisation) score showed prognostic significance, emphasising its potential as a predictive marker. There was a significant association between Ki67 and mitosis scores. The mean proportion of BC cells in the cell cycle was 24% compared to 5% in mitosis. Tumours with a high proportion of mitoses to the overall cycling cells were associated with clinical parameters characteristics of aggressive tumour behaviour.

Considering the remaining controversy regarding the definition of PR positivity and its predictive value in BC, a large cohort of luminal BC (n=2304) was utilised to assess PR expression. Negative PR status showed a significant association with poor patient outcome. The optimal cutoff to categorise PR expression levels into two clinically distinct prognostic groups on both core needle biopsies and full-face sections was 10%. In multivariate analysis, PR was an independent prognostic factor in luminal BC patients who received endocrine therapy. Furthermore, automated PR scores using DL model showed a significant association with patient survival. When PR and Ki67 expressions were compared in multivariate analysis, PR showed independent prognostic significance in patients received endocrine therapy while Ki67 lost its significance.

The assessment of the selected biomarker panel revealed a significant association of their protein expression and patient outcome. High PLK1 protein expression showed independent association with poor outcome in TNBC patients. Similar results were observed at the mRNA levels. In the whole cohort, strong association was observed between high H.H1 expression and shorter survival. MCM7 was independently associated with shorter survival, particularly in luminal BC patients.

Understanding intra-tumour heterogeneity (ITH) is crucial for more accurate prognostication and treatment planning of early-stage BC. This study employs AI to assess ITH in early-stage BC, aiming to unveil subtle variations within tumours that may impact patient outcome. In this study, a large cohort (n=2561) of BC was utilised. Using a DL model, 162 features were identified, extracted, and quantified from WSI. These features related to tumour differentiation and tumour associated microenvironment. The highest degree of ITH was seen in larger size tumours, tumours of no special type (NST) and grade 3. When all features were considered together into one score, a high heterogeneity score was significantly associated with poor patient outcome highlighting its clinical relevance.

Conclusion:

This comprehensive study emphasises the importance of considering all morphological features in BC grading, particularly in WSI. Atypical mitosis, refined proliferation assessment, and the interplay between Ki67 and mitosis contribute valuable insights into tumour aggressiveness. PR emerges as a potential tool for identifying patients suitable for aggressive adjuvant therapy. DL models prove effective in deciphering ITH complexity, providing additional prognostic information for BC outcomes.

STUDY RELATED PUBLICATIONS

Peer-Reviewed Publications

1- Lashen A, Alqahtani S, Shoqafi A, Algethami M, Jeyapalan JN, Mongan NP, Rakha EA, Madhusudan S. **Clinicopathological Significance of Cyclin-Dependent Kinase 2 (CDK2) in Ductal Carcinoma In Situ and Early-Stage Invasive Breast Cancers.** Int J Mol Sci. 2024 May 6;25(9):5053. doi: 10.3390/ijms25095053. PMID: 38732271; PMCID: PMC11084890.

2- **Lashen AG**, Toss MS, Wootton L, Green AR, Mongan NP, Madhusudan S, Rakha E. **Characteristics and prognostic significance of polo-like kinase-1 (PLK1) expression in breast cancer.** Histopathology. 2023 Sep;83(3):414-425. doi: 10.1111/his.14960. Epub 2023 May 24. PMID: 37222669.

3- **Lashen AG**, Toss MS, Ghannam SF, Makhlouf S, Green A, Mongan NP, Rakha E. **Expression, assessment and significance of Ki67 expression in breast cancer: an update.** J Clin Pathol. 2023 Jun;76(6):357-364. doi: 10.1136/jcp-2022-208731. Epub 2023 Feb 22. PMID: 36813558.

4- **Lashen AG**, Toss MS, Mongan NP, Green AR, Rakha EA. **The clinical value of progesterone receptor expression in luminal breast cancer: A study of a large cohort with long-term follow-up**. Cancer. 2023 Apr 15;129(8):1183-1194. doi: 10.1002/cncr.34655. Epub 2023 Jan 18. PMID: 36653923.

5- **Lashen A**, Toss MS, Fadhil W, Oni G, Madhusudan S, Rakha E. **Evaluation oncotype DX® 21-gene recurrence score and clinicopathological parameters: a single institutional experience**. Histopathology. 2023 Apr;82(5):755-766. doi: 10.1111/his.14863. Epub 2023 Feb 1. PMID: 36631400.

6- **Lashen A**, Al-Kawaz A, Jeyapalan JN, Alqahtani S, Shoqafi A, Algethami M, Toss M, Green AR, Mongan NP, Sharma S, Akbari MR, Rakha EA, Madhusudan S. **Immune**

infiltration, aggressive pathology, and poor survival outcomes in RECQL helicase deficient breast cancers. Neoplasia. 2024 Jan;47:100957. doi: 10.1016/j.neo.2023.100957. Epub 2023 Dec 21. PMID: 38134458; PMCID: PMC10777014.

7- Lu W, **Lashen AG**, Wahab N, Miligy IM, Jahanifar M, Toss M, Graham S, Bilal M, Bhalerao A, Atallah NM, Makhlouf S, Ibrahim AY, Snead D, Minhas F, Raza SEA, Rakha E, Rajpoot N. **AI-based intra-tumor heterogeneity score of Ki67 expression as a prognostic marker for early-stage ER+/HER2- breast cancer**. J Pathol Clin Res. 2024 Jan;10(1):e346. doi: 10.1002/cjp2.346. Epub 2023 Oct 24. PMID: 37873865; PMCID: PMC10766021.

8- Wahab N, Toss M, Miligy IM, Jahanifar M, Atallah NM, Lu W, Graham S, Bilal M, Bhalerao A, ***Lashen AG***, Makhlouf S, Ibrahim AY, Snead D, Minhas F, Raza SEA, Rakha E, Rajpoot N. **AI-enabled routine H&E image based prognostic marker for early-stage luminal breast cancer**. NPJ Precis Oncol. 2023 Nov 15;7(1):122. doi: 10.1038/s41698-023-00472-y. PMID: 37968376; PMCID: PMC10651910.

9- Ibrahim A, Jahanifar M, Wahab N, Toss MS, Makhlouf S, Atallah N, **Lashen AG**, Katayama A, Graham S, Bilal M, Bhalerao A, Ahmed Raza SE, Snead D, Minhas F, Rajpoot N, Rakha E. **Artificial Intelligence-Based Mitosis Scoring in Breast Cancer: Clinical Application**. Mod Pathol. 2023 Dec 27;37(3):100416. doi: 10.1016/j.modpat.2023.100416. Epub ahead of print. PMID: 38154653.

10- **Lashen, A.G**., Toss, M.S. and Rakha, E.A. (2023), Reply to "Increased survival of women with luminal breast cancer and progesterone receptor immunohistochemical expression of greater than 10%". Cancer, 129: 2105-2105. https://doi.org/10.1002/cncr.34826

11- Makhlouf S, Wahab N, Toss M, Ibrahim A**, Lashen AG,** Atallah NM, Ghannam S, Jahanifar M, Lu W, Graham S, Mongan NP, Bilal M, Bhalerao A, Snead D, Minhas F, Raza SEA, Rajpoot N, Rakha E. **Evaluation of tumour infiltrating lymphocytes in luminal breast cancer using artificial intelligence.** Br J Cancer. 2023 Nov;129(11):1747-1758. doi: 10.1038/s41416-023-

12-Atallah NM, Wahab N, Toss MS, Makhlouf S, Ibrahim AY, **Lashen AG**, Ghannam S, Mongan NP, Jahanifar M, Graham S, Bilal M, Bhalerao A, Ahmed Raza SE, Snead D, Minhas F, Rajpoot N, Rakha E. **Deciphering the Morphology of Tumor-Stromal Features in Invasive Breast Cancer Using Artificial Intelligence.** Mod Pathol. 2023 Oct;36(10):100254. doi: 10.1016/j.modpat.2023.100254. Epub 2023 Jun 26. PMID: 37380057.

13- Lashen A, Toss MS, Green AR, Mongan NP, Rakha E. **Ki67 assessment in invasive luminal breast cancer: a comparative study between different scoring methods**. Histopathology. 2022 Dec;81(6):786-798. doi: 10.1111/his.14781. Epub 2022 Sep 19. PMID: 35997652; PMCID: PMC9826086.

14-Toss MS, **Lashen A**, Ibrahim AY, Rakha EA. **Assessment of proliferative activity in breast cancer: Approaches, significance and challenges.** Clinical laboratory diagnostics (2022).

15- Lashen A, Toss MS, Alsaleem M, Green AR, Mongan NP, Rakha E. **The characteristics and clinical significance of atypical mitosis in breast cancer**. Mod Pathol. 2022 Oct;35(10):1341-1348. doi: 10.1038/s41379-022-01080-0. Epub 2022 May 2. PMID: 35501336; PMCID: PMC9514994.

16-Ibrahim A, **Lashen AG**, Katayama A, Mihai R, Ball G, Toss MS, Rakha EA. **Defining the area of mitoses counting in invasive breast cancer using whole slide image**. Mod Pathol. 2022 Jun;35(6):739-748. doi: 10.1038/s41379-021-00981-w. Epub 2021 Dec 11. PMID: 34897279; PMCID: PMC9174050.

17- Wahab N, Miligy IM, Dodd K, Sahota H, Toss M, Lu W, Jahanifar M, Bilal M, Graham S, Park Y, Hadjigeorghiou G, Bhalerao A, **Lashen AG**, Ibrahim AY, Katayama A, Ebili HO, Parkin M, Sorell T, Raza SEA, Hero E, Eldaly H, Tsang YW, Gopalakrishnan K, Snead D, Rakha E, Rajpoot N, Minhas F. **Semantic annotation for computational pathology: multidisciplinary experience and best practice** recommendations. J Pathol Clin Res. 2022 Mar;8(2):116-

128. doi: 10.1002/cjp2.256. Epub 2022 Jan 10. PMID: 35014198; PMCID: PMC8822374.

18- Ibrahim A, **Lashen A,** Toss M, Mihai R, Rakha E. **Assessment of mitotic activity in breast cancer: revisited in the digital pathology era**. J Clin Pathol. 2022 Jun;75(6):365-372. doi: 10.1136/jclinpath-2021-207742. Epub 2021 Sep 23. PMID: 34556501.

19-Lashen AG, Toss MS, Katayama A, Gogna R, Mongan NP, Rakha EA. **Assessment of proliferation in breast cancer: cell cycle or mitosis**? An observational study. Histopathology. 2021 Dec;79(6):1087-1098. doi: 10.1111/his.14542. Epub 2021 Sep 20. PMID: 34455622.

20-Lashen A, Ibrahim A, Katayama A, Ball G, Mihai R, Toss M, Rakha E. **Visual assessment of mitotic figures in breast cancer: a comparative study between light microscopy and whole slide images**. Histopathology. 2021 Dec;79(6):913-925. doi: 10.1111/his.14543. Epub 2021 Sep 22. PMID: 34455620.

Conference Presentations

1- The clinical value of cyclin dependent kinases CDK4 and CDK6 in the context of p53 status in early-stage breast cancers, **(The 14th Joint Meeting of the BDIAP and The Pathological Society, 2023, Liverpool, United Kingdom).**
Poster presentation

2- Nottingham Prognostic x (NPx): a clinical decision-making tool in ER positive - HER2- negative breast cancer, (**The 14th Joint Meeting of the BDIAP and The Pathological Society, June 2023, Liverpool, United Kingdom).**
Poster presentation

3- The Clinical Value of Cyclin Dependent Kinases 2 (CDK2) in preinvasive and Invasive Breast Cancers, (**BDI Annual Research Symposium, September 2023, Nottingham, United Kingdom)**
Oral presentation

4- The clinical value of cyclin dependent kinases CDK4 and CDK6 in early-stage breast cancers, **(Nottingham Breast Cancer Research Centre Research Day May 2023, Nottingham, United Kingdom)**
Poster presentation

5- The Prognostic and predictive value of Polo-like kinase-1 (PLK1) expression in breast cancer; is it Oncogene or Tumour Suppressor Gene? (**Pathological Society Meeting January 2023, London, United Kingdom)**
Poster presentation

6- The characteristics and prognostic significance of immunohistochemical expression of Histone H1 in breast cancer (**Pathological Society Meeting January 2023, London, United Kingdom)**
Poster presentation

7- Evaluation Oncotype dx® 21-gene recurrence score and clinicopathological parameters in early-stage breast cancer (**Pathological Society Meeting January 2023, London, United Kingdom)**
Poster presentation

8- The clinical value of progesterone receptor expression in luminal breast cancer) (**Pathological Society Meeting January 2023, London, United Kingdom)**
Poster presentation

9- The clinical and biological significance of atypical mitosis in breast cancer (**Pathological Society Meeting January 2023, London, United Kingdom)**
Poster presentation

10-Ki67 assessment in ER positive / HER2 negative breast cancer: A comparative study between different scoring methods (**Pathological Society Meeting January 2023, London, United Kingdom)**
Poster presentation

11-Assessment and clinical value of progesterone receptor expression in **early-stage breast cancer (Sue Watson Postgraduate Oral Presentation event May 2022)**
Oral Presentation

12-Assessment of proliferation in breast cancer **(Nottingham Breast Cancer Research Centre Research Day May 2022, Nottingham, United Kingdom)**
Poster presentation

13- Eyeballing assessment of mitotic figures in invasive breast cancer using whole slide image: A Comparative Morphological Study with Light Microscopy (**13th Joint Meeting of BDIAP and the Pathological Society, July 2021, Manchester, United Kingdom)**
Oral presentation

14- Assessment of proliferation in breast cancer: cell Cycle or mitosis **(13th Joint Meeting of BDIAP and the Pathological Society, July 2021, Manchester, United Kingdom)**
Poster presentation

15-Mitosis evaluation in invasive breast carcinoma; a comparative morphological study between light microscopy and digitalised whole slide images **(Edinburgh Virtual Breast Cancer Special Symposium, February 2021**
Poster presentation

16- Assessment of proliferation in breast cancer: cell cycle or mitosis? An observational study **(Edinburgh Virtual Breast Cancer Special Symposium, February 2021)**
Poster presentation

17- Assessment of Ki67 in early-stage breast cancers **(Nottingham Breast Cancer Research Centre Weekly Seminar April 2020)**
Oral presentation

CONTENTS

Table of Contents

Chapter 1 General Introduction 1

1.1 BC definition and historical background 2

1.2 Incidence and Epidemiology 2

1.3 BC-related morbidity 3

1.4 Risk factors 4

1.5 Pathogenesis and cytogenetics 4

1.6 Modes of Spread: Lymphatic and Haematogenous 5

1.7 Routine Practice for BC diagnosis 6

1.8 BC Prognostic and Predictive Factors .....

1.8.1 Clinico-pathologic factors as prognostic factors in BC 7

1.8.2 Hormone Receptors ER and PR 13

1.8.3 Human Epidermal Growth Factor Receptor 2 (HER2) 13

1.9 BC Molecular Classification 14

1.9.1 Multi-gene Signatures 1

1.10 Digital pathology and artificial intelligence in BC 1

1.11 Cellular proliferation in BC 4

1.11.1 Phases of cell cycle 5

1.11.2 Mitosis in BC 6

1.11.3 Ki67 in BC 14

1.12 Progesterone receptor (PR) in BC 33

1.12.1 Function of PR in BC 34

1.12.2 PR assessment in clinical routine practice 35

1.12.3 PR in molecular subtype of BC 35

1.12.4 PR and proliferation in BC 36

1.12.5 PR and endocrine therapy 36

1.12.6 PR cutoff 36

1.13 Tumour heterogeneity in BC 41

1.13.1 Inter-tumour heterogeneity 41

1.13.2 Intra-tumour heterogeneity in BC 42

1.13.3 Assessment of intra-tumoral heterogeneity 45

1.14 Hypothesis and aims of the thesis 46

Chapter 2 Patients and methods 49
2.1 Patient cohort and data collection 50
2.2 Histological review 53
2.3 Slide digitalisation and image acquisition 53
2.4 Tissue Microarray Construction 54
2.4.1 Preparation of the donor blocks 63
2.4.2 Matching of donor blocks with their corresponding slides 63
2.4.3 Digital images marking 64
2.4.4 Recipient block construction 64
2.4.5 TMA layout design 65
2.4.6 Coring process 66
2.4.7 TMA block annealing and sectioning 68
2.5 Optimisation of antibodies used in the study. 55
2.5.1 Validation of antibody specificity by Western Blot 55
2.5.2 Preparation of samples, electrophoresis and transfer 55
2.5.3 Optimisation of antibodies by IHC 57
2.5.4 IHC technique 58
2.5.5 Assessment of immunoreactivity 62
2.5.6 Determination of the cutoff 63
2.6 AI-assisted interpretation of histopathological variables and biomarkers 64
2.6.1 Algorithm training 64
2.6.2 Tissue and tumour segmentation 69
2.6.3 Cell segmentation and classification 69
2.6.4 Prediction of morphological features of BC 70
2.7 Statistical approach 71
2.7.1 Intra-Class Correlation Coefficient (ICC) 72
2.7.2 Correlation with clinicopathological parameters 72
2.7.3 Univariate analysis with outcome 72
2.7.4 Multivariate analysis with outcome 72
2.7.5 Calculation of Sensitivity, Specificity 73
2.8 Validation of the Findings 73
Chapter 3 Eyeballing assessment of mitotic figures in BC 74
3.1 Introduction 75
3.1.1 Hypothesis and aims. 76

3.2 Patient and methods 77
3.3 Tissue preparation and digitalisation 79
3.4 Refining the definition of mitosis in WSI 79
3.4.1 Comparison of mitotic figures between LM and WSI 79
3.4.2 Identification of mitosis and mimickers with WSI 82
3.4.3 Comparison of mitotic counts between LM and WSI 83
3.4.4 Evaluation of the prevalence of mitotic phases 83
3.4.5 Refining mitotic cutoffs when using WSI 84
3.5 Statistical analysis 84
3.6 Results 85
3.6.1 Comparison of mitotic figures between LM and WSI 85
3.6.2 Refined criteria for mitosis identification in WSI 97
3.6.3 Validating the refined criteria for mitoses in WSI 99
3.6.4 Comparing between mitotic counts in LM and in WSIs 101
3.6.5 The prevalence of mitotic phases in BC 103
3.6.6 Refining existing mitosis cutoff in WSI. 108
3.7 Discussion 111
Chapter 4 The role of atypical mitosis in BC 116
4.1 Introduction 117
4.1.1 Hypothesis and aims. 118
4.2 Patients and Methods 119
4.2.1 Characteristics and significance of atypical mitosis 119
4.2.2 Identification and counting of atypical mitosis in WSI. 121
4.2.3 Prognostic and predictive value of atypical mitosis in BC 123
4.2.4 Deciphering the molecular changes linked with atypical mitoses 123
4.2.5 Statistical analysis 124
4.3 Results 124
4.3.1 Characteristics of atypical mitoses in BC 124
4.3.2 Correlation with other clinicopathological parameters 125
4.3.3 Outcome analysis 131
4.3.4 Atypical mitoses and the transcriptomic profiling 138
4.4 Discussion 142
Chapter 5 Assessment and clinical value of Ki67 in BC 146
5.1 Introduction 147

5.1.1 Hypothesis 149
5.1.2 Aims 149
5.2 Patients and methods 150
5.2.1 Study cohort 150
5.2.2 Ki67 staining: 150
5.2.3 Comparing between various methods of eyeballing Ki67 assessment 150
5.2.4 Evaluation of the quantitative differences between BC cells in the cell cycle and those in mitosis 154
5.2.5 Assessment of the best cutoff for Ki67 in luminal BC. 156
5.2.6 Assessment the genetic pathways associated with Ki67. 156
5.2.7 Assessment of automated Ki67 scores in BC 157
5.2.8 Statistical analysis 160
5.3 Results 161
5.3.1 Comparing methods of eyeballing Ki67 assessment 161
5.3.2 Clinical impact of Ki67 intra-tumoural heterogeneity 176
5.3.3 The quantitative differences between BC cells in cell cycle and in mitosis 180
5.3.4 Assessment of the best cutoff of Ki67 expression in luminal BC. 187
5.3.5 Genetic pathways associated with *MKI67*. 189
5.3.6 Assessment of the Ki67 automated scores 191
5.4 Discussion 197
Chapter 6 The clinical value of progesterone receptor (PR) in BC 211
6.1 Introduction 212
6.1.1 Hypothesis and aims: 214
6.2 Patients and Methods 214
6.2.1 Study cohort 214
6.2.2 Immunohistochemical staining of PR. 216
6.2.3 Eyeballing evaluation of PR expression 216
6.2.4 Exploring the optimal cutoff to define PR positivity. 217
6.2.5 Automated scoring of PR 218
6.3 Statistical analysis 218
6.3.1 RESULTS 219
6.3.2 Eyeballing PR expression 219

6.3.3 Automated assessment of PR expression and clinicopathological parameters 239
6.3.4 Automated PR scoring and patient survival. 242
6.4 DISCUSSION 245
Chapter 7 Identification of potential prognostic markers in BC 252
7.1 Introduction 253
7.1.1 Hypothesis and aims. 255
7.2 Patients and Methods: 255
7.2.2 H.H1, PLK1 and MCM7 immunostaining 257
7.2.3 H.H1, PLK1 and MCM7 protein expression evaluation 258
7.2.4 Statistical analysis 259
7.3 Results 260
7.3.1 H.H1 expression in BC 260
7.3.2 PLK1 in BC 270
7.3.3 MCM7 in BC 280
7.3.4 Correlation between H.H1, PLK1 and MCM7 293
7.4 DISCUSSION 294
Chapter 8 Assessment of intra-tumour heterogeneity of BC using AI 302
8.1 Introduction 303
8.1.1 Hypothesis and aims. 304
8.2 Patients and Methods 305
8.2.1 Study cohort 305
8.2.2 Training of algorithm 307
8.2.3 Computing of ITH features using DL 307
8.2.4 Segmentation and classification of nuclei 307
8.2.5 Region segmentation 308
8.2.6 Evaluation of the morphological features of ITH. 309
8.2.7 Statistical Analysis 312
8.3 Results 313
8.3.1 Assessed intra-tumoural heterogeneity features. 313
8.3.2 Association of ITH features and clinicopathological parameters 313
8.3.3 Overall heterogeneity score 318
8.3.4 Outcome analysis 321
8.3.5 Intra-tumoural heterogeneity score in sub-cohorts 329

8.4 DISCUSSION 331
Chapter 9 General Discussion 336
9.1 Background 337
9.2 Hypothesis, aims, and methods of the study. 338
9.3 Cardinal findings of the study 339
9.3.1 Eyeballing assessment of mitotic figures in BC: 339
9.3.2 The role of atypical mitosis in BC 340
9.3.3 Assessment of proliferation in BC using Ki67 341
9.3.4 The Role of PR in BC 342
9.3.5 Identification of potential prognostic markers in BC 343
9.3.6 Assessment of ITH in early-stage BC using AI 344
9.4 Final conclusions 345
9.5 Strengths of the study 346
9.6 Potential limitations 347
9.7 Future Plans 348
References 62

LIST OF TABLES

Table 1-1: Comparison between Ki67 and other proliferative markers in BC 15

Table 2-1: Clinicopathological characteristics of the study cohort..... 52

Table 2-2: Specificity of the selected antibodies for the project 56

Table 2-3: Optimum conditions of the selected antibodies..... 61

Table 3-1: Clinicopathological characteristics of the study cohort..... 78

Table 3-2: Matched mitotic figures between LM and WSI..... 85

Table 3-3: Morphological features of mitotic figures versus mimickers in WSI..... 92

Table 3-4: Comparison between existing and refined criteria for mitosis identification 98

Table 3-5: Concordance of recognising the recorded morphological features of mitoses 100

Table 3-6: Interobserver concordance in mitotic phases identification 101

Table 3-7: Comparison between mitotic scores on LM versus WSI using the existing mitosis cutoff..... 103

Table 3-8: The prevalence of mitotic phases in LM..... 105

Table 3-9*:* The prevalence of mitotic phases in WSI..... 105

Table 3-10: Morphological features of mitosis recorded by each mitotic phase..... 106

Table 3-11: Comparison between Mitotic score with LM versus with WSI with the refined cutoff 109

Table 4-1: Clinicopathological characteristics of the Nottingham cohort 120
Table 4-2: The relationship between overall mitoses and clinicopathological parameters 127
Table 4-3: The relationship between atypical mitosis and clinicopathological parameters 128
Table 4-4: Relationship atypical-to-typical mitoses ratio and clinicopathological parameters 129
Table 4-5: Differential gene expression linked with the different mitotic scores. 139
Table 4-6: The Biological process that are significantly associated with high atypical mitoses. 141
Table 5-1: Low versus high Ki67 expression groups with different methods of scoring using multiple cutoffs. 163
Table 5-2: The relationship between Ki67 expression in luminal BC assessed by various methods and clinicopathological parameters ... 167
Table 5-3: Univariate Cox regression analysis for predictors of breast cancer-specific survival according to the different methods of Ki67 assessment with different Ki67 cutoffs 171
Table 5-4: Multivariate Cox regression analysis for predictors of breast cancer-specific survival according to the different methods of Ki67 assessment with different Ki67 cutoffs 172
Table 5-5: Relationship between Ki67 heterogeneity index and clinicopathological parameters 176
Table 5-6: Distribution of Ki67 and mitotic scores 181

Table 5-7: Relationship between mitotic scores and scores of cycling (Ki67-positive) cell scores (low, intermediate, high)..... 181
Table 5-8: Relationship between proportions of cycling BC cells entering mitotic phase and clinicopathological diameters..... 185
Table 5-9 : Relationship between the discrepant mitosis/ Ki67 groups and clinicopathologic parameters. 187
Table 5-10: Gene Ontology shows the common pathways that are significantly associated with high *MKI67.* 190
Table 5-11: Relationship between Ki67 expressions in luminal BC assessed by manual versus automated scoring and clinicopathological parameters. 192
Table 5-13: Multivariate analysis of manual Ki67 scores in luminal BC 196
Table 5-14: Multivariate analysis of automated Ki67 scores in luminal BC 196
Table 5-15: Multivariate analysis of automated versus manual Ki67 scores in luminal BC..... 197
Table 6-1: Clinicopathological characteristics of the study cohort..... 215
Table 6-2: Comparison between different cutoffs of PR expression in full face sections 221
Table 6-3: Comparison between different cutoffs of PR expression in needle core biopsies..... 222
Table 6-4: Sensitivity and specificity of PR cutoff in needle core biopsies compared to full face sections. 223

Table 6-5: Relationship between PR status and clinicopathological parameters 225
Table 6-6: Relationship between H. score of PR (low /high) and clinicopathological parameters 226
Table 6-7: Comparison between the prognostic and predictive values of PR and Ki67 in univariate analysis in luminal BC. 235
Table 6-8: Multivariate analysis of Ki67 and PR in luminal BC using full face sections..... 236
Table 6-9 : Multivariate analysis of Ki67 and PR in BR using core needle biopsies. 237
Table 6-10: Relationship between automated PR status and clinicopathological parameters in luminal BC..... 239
Table 6-11: Relationship between PR density and clinicopathological parameters in luminal BC 241
Table 6-12: Multivariate analysis of automated PR expression in endocrine therapy treated patients. 244
Table 6-13: Multivariate analysis of PR density in endocrine therapy treated patients..... 244
Table 6-14: Multivariate analysis of eyeballing PR scores, automated PR scores and PR density in endocrine therapy treated patients..... 244
Table 7-1: Clinicopathological characteristics of the study cohort..... 256
Table 7-2: Relationship between H.H1 expression and clinicopathological parameters in BC..... 263
Table 7-3: Multivariate analysis of H.H1 expression in BC 269

Table 7-4: Relationship between PLK1 expression and clinicopathological parameters 274
Table 7-5: Multivariate analysis PLK1 expression in luminal BC 279
Table 7-6: Multivariate analysis PLK1 expression in TNBC 279
Table 7-7: Relationship between nuclear MCM7 expression and clinicopathological parameters 284
Table 7-8: Multivariate analysis of MCM7 expression in BC 293
Table 7-9 Multivariate analysis of MCM7 expression in luminal BC patients who were treated with endocrine therapy. 293
Table 8-1: Clinicopathological characteristics of the study cohort. 306
Table 8-2: Relationship between tumour differentiation heterogeneity clinicopathological parameters 315
Table 8-3: Relationship between stromal heterogeneity and clinicopathological parameters 316
Table 8-4: Relationship between TILs heterogeneity and clinicopathological parameters. 317
Table 8-5: Mean and standard deviation (Mean ± SD) of heterogeneity score in the different pathological parameters 319
Table 8-6: Relationship between overall heterogeneity score and clinicopathological parameters in the study cohort 320
Table 8-7 Multivariate analysis of overall intra-tumour heterogeneity score in luminal BC 328
Table 8-8: Relationship between Heterogeneity score and clinicopathological parameters in LN positive BC 330

LIST OF FIGURES

Figure 1-1: Nottingham histologic grading of BC 11

Figure 1-2: Molecular classification of BC [89]. 1

Figure 1-3: The value of AI in the pathological diagnostic process [121] 4

Figure 1-4: Phases of the cell cycle [138]. 6

Figure 1-5: Mitotic figures in different phases in WSI (40x) 11

Figure 1-6: Process of digitalisation of glass slides into WSI for mitosis identification 14

Figure 1-7: Overview of Ki67 localisation and functions. 17

Figure 1-8: Representative expression of Ki67 in BC. 19

Figure 1-9: PR/RANKL signal in human breast. 34

Figure 1-10: Overview of BC intra-tumour heterogeneity [312] 45

Figure 2-1: Regional annotation of H&E slides 66

Figure 2-2: Cellular annotation of H&E slides 66

Figure 2-3: Regional and cellular annotation for PR 68

Figure 2-4: Regional and cellular annotation for Ki67 68

Figure 2-5: Segmentation of invasive tumour and DCIS regions 69

Figure 2-6: HoVer-Net for cell segmentation. 70

Figure 2-7: Tissue region classification including tumour, normal TDLU and stroma. 71

Figure 3-1 : Recorded criteria to characterise mitoses in WSI. 81

Figure 3-2: Differentiating between prophase and apoptosis in WSI at (40x) 94
Figure 3-3: Differentiating between metaphase and elongated hyperchromatic nuclei in WSI at 40x 95
Figure 3-4: Differentiating between Prophase and lymphocytes in WSI (40x) 96
Figure 3-5: Scatter plot shows the correlation between mitotic counts with LM versus WSI. 102
Figure 3-6: Kaplan Meir curves showing the association between digital mitotic scores and both BCSS (a) and DMFS (b) with applying the refined cutoff for mitosis. 110
Figure 4-1: Morphological features of typical and atypical mitosis (40×) 122
Figure 4-2: Representative image showing different types of atypical mitotic figures in one field (20x). 123
Figure 4-3: Unfamiliar shapes of atypical mitosis 125
Figure 4-4: Kaplan Meier curves showing the associations of overall mitosis with (a) BCSS and with (b) DMFS. 132
Figure 4-5: Kaplan Meier curves showing the associations of atypical mitosis with (a) BCSS and with (b) DMFS. 133
Figure 4-6: Kaplan Meier curves showing the associations of atypical-to-typical mitoses ratio with (a) BCSS and (b) DMFS) 134
Figure 4-7: Kaplan Meier curves showing the associations of atypical mitoses with (a) BCSS and with (b) DMFS in luminal BC. 135

Figure 4-8: Kaplan Meier curves showing the associations of atypical-to-typical mitoses ratio with BCSS in luminal BC (a, b) while (c, d) show association of high ratio with both poor BCSS and DMFS in TNBC. . 136
Figure 4-9: Kaplan Meier curves showing a significant association between BCSS and atypical-to-typical mitoses ratio in TNBC patients who did receive chemotherapy (a), but no association was observed in terms of overall mitoses (b). 137
Figure 4-10: Venn diagram of obtained DEG..... 140
Figure 4-11: (a) Volcano plot obtained from DEG analysis 140
Figure 5-1: Patterns of Ki67 staining 153
Figure 5-2: Illustration of Ki67CL score..... 159
Figure 5-3: a) BC case showing very low Ki67 expression while b) another case showing high Ki67 expression. 162
Figure 5-4: Assessment of Ki67 in clinical practice..... 165
Figure 5-5: Kaplan Meier plots showing the associations of BCSS with Ki67 using all scoring methods (a-c) and the same findings were observed in terms of DMFS (d-f) 170
Figure 5-6: Kaplan Meier curves show the association of Ki67 with different scoring methods 173
Figure 5-7: Kaplan Meier curves show the association of Ki67 granularity and outcome..... 174
Figure 5-8: Kaplan Meier curve shows the association of high Ki67 expression with granular pattern and shorter DMFS in luminal BC. .. 174
Figure 5-9: Kaplan Meier curve shows a significant association between H-score of Ki67 expression and short DMFS. 175

Figure 5-10: Line graph shows the correlation between Ki67 heterogeneity index and the Ki67 expression levels 178
Figure 5-11: Kaplan Meier plots show significant associations between Ki67 heterogeneity index and shorter BCSS (a), DMFS (b) in the whole cohort and BCSS in patients who received endocrine therapy (c). 179
Figure 5-12: Box plot shows the relationship between mitotic scores in H&E and Ki67 positive cells. 182
Figure 5-13: Relationship between hotspot of Ki67 and mitosis (40x) 182
Figure 5-14: Different phases of Ki67-positive mitotic cells (40x) 183
Figure 5-15: Matched mitotic figures between H&E and Ki67(40x) ... 183
Figure 5-16: Box plot shows the relationship between proportion of BC cells entering mitosis phase and molecular subtypes. 185
Figure 5-17: ROC curves show area under the curve of 0.7 for Ki67 expression levels in hotspot and it is 0.64 for average Ki67expression levels in the whole slide. 188
Figure 5-18: Kaplan Meier curve shows of the association of high localisation score with poor survival. 194
Figure 5-19: Kaplan Meier curves shows manual versus automated Ki67 score with patient outcome. 195
Figure 6-1: Expression of PR in luminal BC 217
Figure 6-2: Normal distribution curve shows the bimodality of PR expression in BC. 223
Figure 6-3: Kaplan Meier curves show the association of PR status with a) BCSS and b) DMFS. 228

Figure 6-4: Kaplan Meier curves show a significant association between low PR H. score and a) BCSS and b) DMFS. 229
Figure 6-5: Kaplan Meier curves show a significant association of PR negativity with short BCSS in grade 2 tumours (b), but no association was found in grade 1or grade 3 tumours (a&c). 230
Figure 6-6: Kaplan Meier curves show significant associations between PR negativity and poor BCSS and DMFS in post-menopausal patients (c&d) while no association was observed in terms of pre-menopausal patients (a&b). 231
Figure 6-7: Kaplan Meier curves show significant association between PR negativity and poor BCSS (a) and DMFS (b) in patients who received endocrine therapy. 232
Figure 6-8: Kaplan Meier curves show significant association between low PR H. score and poor BCSS (a) and DMFS (b) in patients who received endocrine therapy..... 233
Figure 6-9: Kaplan Meier curves show a significant association between negative PR status with both BCSS and DMFS in low Ki67 expression tumours (a&b) as well as high Ki67 tumours (c&d). 238
Figure 6-10: Kaplan Meier curves show a significant association between negative PR status using automated scores with short survival (a) while high PR density was associated with worse survival (b). 243
Figure 7-1: Box and whisker plot*s* show a significant association *between* high *H.H1* mRNA expression and NST tumour type (a) while (b) shows that *H.H1* mRNA levels were significantly higher in Basal-like BC..... 260

Figure 7-2: Kaplan Meier curves show significant associations of high *H.H1* mRNA expression and poor outcome in the whole bc-GenExMiner cohort (a) and basal-like BC (b). 261
Figure 7-3: Western blot of H.H1 antibody shows specific bands at 30 KDa and B actin at 42 KDa. 262
Figure 7-4*:* H.H1 protein expression in BC 262
Figure 7-5 : Kaplan Meier curves show the association of high H.H1 expression and poor BCSS (a) 265
Figure 7-6: Kaplan Meier curves show the significant association of H.H1 and BCSS in TNBC (b) 266
Figure 7-7: Kaplan Meier curves show the association of H.H1 and short BCSS in endocrine therapy naïve luminal BC patients (a) but, no association was observed in endocrine therapy treated patients (b). 267
Figure 7-8: Kaplan Meier curves show the association of H.H1 and shorter BCSS in TNBC patients who didn't receive chemotherapy (a) but, no association was observed in chemotherapy treated patients (b). . 268
Figure 7-9: Box and whisker plots showing significantly higher *PLK1* mRNA levels in NST tumour (a) and in basal-like BC (b). 270
Figure 7-10: Kaplan Meier curves show association between *PLK1* mRNA levels and survival using Kaplan Meier Plotter cohort. 271
Figure 7-11: Western blot of PLK1 antibody shows a specific band at 60 KDa while B-actin shows specific bands at 42 KDa. 272
Figure 7-12: PLK1 expression in BC 273
Figure 7-13: a- Kaplan Meier curves show the association of PLK1 expression with BCSS (a) and DMFS (b). 276

Figure 7-14: Kaplan Meier curves show association of high PLK1 with longer survival in luminal and HER2+ BC (a&b) while high PLK1 expression shows association with poor outcome in TNBC (c). 277
Figure 7-15: Kaplan Meier curves shows the association between PLK1 and adjuvant therapy 278
Figure 7-16: Box and whisker plots showing significantly higher *MCM7* mRNA expression in NST tumour type (a) and basal-like BC in gene ex miner data. 280
Figure 7-17: Kaplan Meier curves showing a significant association between high *MCM7* mRNA levels and poor outcome in the whole cohort (a) and luminal BC in gene ex miner data. (b). 281
Figure 7-18: Kaplan Meier curves show association of *PLK1* and molecular subtypes..... 281
Figure 7-19: Western blot of MCM7 antibody shows specific bands at a 80 kDa while B-actin shows specific bands at 42 kDa 282
Figure 7-20: BC cases exhibiting nuclear (a), cytoplasmic (b), and both nuclear and cytoplasmic (c) MCM7 expression (x20). 283
Figure 7-21: Kaplan Meier curves showing a significant association between high MCM7 expression and shorter BCSS (a) in the whole cohort while no association was observed with DMFS (b)..... 287
Figure 7-22: Kaplan Meir curves showing the association of high MCM7 with shorter BCSS in luminal BC (a) while no association was observed in both HER2 enriched and TNBC (b&c). 288
Figure 7-23: Kaplan Meir curves showing the association between high MCM7 with short BCSS in endocrine therapy naïve patients (a) where no

association was observed in those who received endocrine therapy (b). 289
Figure 7-24: Kaplan Meier curves show a significant association between MCM7 and tumour behaviour 291
Figure 7-25: Kaplan Meier curves showing a significant association between MCM7 expression and adjuvant therapy 292
Figure 8-1: Classification for tumour nuclei for calculation of the ITH features. 308
Figure 8-2: Region segmentation 309
Figure 8-3: Evaluation of morphological features of ITH.....311
Figure 8-4: Kaplan Meier curves show the association of tumour differentiation heterogeneity with BCSS (a) and with DMFS (b). 322
Figure 8-5: Kaplan Meier curves show the association of stromal heterogeneity with BCSS (a) and DMFS (b). 323
Figure 8-6: Kaplan Meier curves show the association of TILs heterogeneity with BCSS (a) and DMFS (b). 324
Figure 8-7: Kaplan Meier curves show the association of heterogeneity score with BCSS (a) and DMFS (b) 325
Figure 8-8: Kaplan Meier curve shows the association of heterogeneity score with BCSS in intermediate risk patients (moderate Nottingham prognostic index). 326
Figure 8-9: Kaplan Meier curves show the association of heterogeneity score and adjuvant therapy 327

Chapter 1 General Introduction

1.1 BC definition and historical background

Breast cancer (BC) is an extraordinarily heterogeneous disease encompassing a diverse spectrum of tumours with varying morphological, biological, and clinical behaviour, outcome, and response to therapy [1, 2]. These distinct tumour subtypes arising within the same anatomical organ, the breast. Such heterogeneity and diversity pose significant impacts on patients' prognosis and treatments [3, 4]. The aetiology of BC is attributed to a complex interaction between various modifiable and non-modifiable factors, and it is determined by genetics, hormonal and environmental factors, which contribute to the development of this disease [5]. The history of BC is a complex maze of attempts to understand the nature of this hormone-responsive cancer and the will of clinicians to conquer this disease [6].

1.2 Incidence and Epidemiology

BC is by far the most common type of cancer in women and the second cause of cancer-related deaths in the Western World. BC is the second most common cancer and the fifth cause of cancer-related deaths overall [7] with an estimated 2.1 million new cases being diagnosed each year and representing 24.2% of all cancer diagnoses among women. BC accounts for 1 in 4 cases in women globally and contributes to 15% of mortality [8]. BC is not a disease of modern society; it was recognised as long as 1600 BC by Imhotep (the Egyptian physician-architect) who provided authentic accounts of BC. At that time, the BC case was deemed incurable if the disease was “cool to touch, bulging and spread all over the breast [5, 9].

Although BC is reported in every country globally, there is a wide geographical variation in incidence, mortality, and survival rates. Incidence rates vary widely in developing regions. Multiple studies have found that the incidence of BC is highest in high-income countries ranging from 85.8 to 91.6 cases per 100,000 [10]. Importantly, a high rate of BC appears to be in the UK where it is responsible for 30% of all cancers in women [11]. In addition, in the UK, over a third of BC occurs in females over the age of 70, and less than one in five females are under the age of 50 at diagnosis [12]. This difference in the incidence of BC between countries can be attributed to the changes in exposure to environmental risk factors, behaviour, and lifestyle factors in different population groups. In addition, genomic modification is an important factor that significantly alters the risk profile of BC tumorigenesis [13, 14].

1.3 BC- related morbidity

Morbidity in BC patients represents a wide array of symptoms and health problems ranging from mild symptoms to devastated health status that results from the spread of the tumour beyond the breast regionally and systemically [2]. More dramatic complications can occur as a result of BC including pathologic bone fractures, hypercalcaemia, neurologic manifestations, and impaired organ functions [15]. These complications arise from the spread of the tumour cells to bone, brain, and parenchymal organs such as the liver, and lungs. In addition, the different modalities of BC treatments cause several morbidities. These include lymphoedema, cutaneous numbness, and impaired sensations; all could result from axillary lymphadenectomy carried out for staging and therapeutic

purposes [16, 17] and side effects caused by systemic chemotherapy [18, 19], hormonal therapy [19, 20], or targeted therapy [21, 22].

Therefore, it could be safely concluded that morbidity in BC patients is mainly attributable to the development of loco-regional or systemic spread of cancer or recurrences [23, 24].

1.4 Risk factors

Risk factors for developing BC include a family history of BC, early menarche, late menopause, nulliparity, exposure to radiation, lack of breastfeeding, long-term usage of hormonal replacement therapy and obesity [25]. One of the highest risk factors for BC development is high breast tissue density as detected by mammography [26]. BRCA1 and BRCA2 mutations are risk factors for BC and are detected in 5% to 10% of diagnosed BC [27]. Early detection of mutation carriers in these genes, in turn, can play an important role in disease prevention [27, 28].

1.5 Pathogenesis and cytogenetics

Several factors can raise the risk of BC. Damage to the DNA and genetic mutations have been experimentally linked to the development of BC [29]. Chromosomal abnormalities and somatic loss of heterozygosity (LOH) have been widely reported in BC as a means of identifying putative tumour-suppressor genes [30]. LOH commonly occurs in chromosomes 17p, 11p and 16q [31]. DNA aneuploidy is a significant predictor for BC progression and survival, whereas the ploidy status characterises genotypic stability and potential metastatic capacity [32, 33]. BC also shows loss of tumour suppressor genes, such as *TP53* and *p16*, and overexpression of proliferation markers such as Ki67 [34-36].

In addition, the potential role of other factors such as myoepithelial cells and surrounding tumour microenvironment (TME) in the process of progression of BC has been extensively studied [37].

1.6 Modes of Spread: Lymphatic and Haematogenous

The spread of BC to regional lymph nodes (LNs) and remote organs occurs through invasion of lympho-vascular channels within the tumour vicinity by tumour cells. Both lymphatic and vascular channels are infiltrated by cancerous cells [38, 39]. Generally, lymphatics are the main route through which carcinoma cells spread to distant sites. This could be attributed to the molecular identity of carcinoma cells themselves and to the permeable nature of lymphatic channels which naturally lack basement membranes possessed by vascular channels [40, 41]. BC cells colonising regional LNs forming satellite metastases are able to make their way into systemic circulation indirectly through the lymphovascular anastomoses within LNs, and hence colonise parenchymal organs [42]. Therefore, metastatic cells within LNs form new stations for showering systemic circulation by carcinoma cells [41]. It has been proposed that primary BCs often invade lymphatics to gain access to systemic circulation, supporting the greater contribution of regional LN metastasis to systemic metastasis [43]. These phenomena justify, from a biological point of view, why the axillary LN stage is the single most important predictor of disease-free survival and overall survival in BC patients [44, 45]. One central route of cancer cell spread is epithelial phenotype transitions whereas the tumour cells are mesenchymal-like (epithelial-to-mesenchymal transition, EMT).

Through the EMT mechanism, epithelial cells in one part of the body break away from the epithelium and then travel to different organs [46].

1.7 Routine Practice for BC diagnosis

Formalin-fixed paraffin-embedded (FFPE) samples represent a cost-effective approach for pathological diagnosis as well as the measurement of valuable tissue markers for disease prognosis and therapeutic prediction [47]. FFPE is routinely collected and archived, along with patients' clinicopathological parameters and outcome data, for further clinical and research purposes [48]. Holding tissue samples and clinical data represents a fundamental requirement for retrospective and prospective research studies essential for the discovery and validation of novel biomarkers of cancer development and progression [49]. UK Royal College of Pathologists minimum dataset report for BC practice includes tumour grade, nodal status, tumour size, histological tumour type, presence of associated *in-situ* component and excision margin status. Reliable reporting on these parameters necessitates optimally fixed and processed tissue sections, is achieved through strict adherence to guidelines for specimen handling, fixation and processing [50, 51]. In addition, the use of immunohistochemistry (IHC), utilising FFPE sections, to detect estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2) status and Ki67 proliferative status is widely used in BC management and prognostication for their invaluable prognostic and predictive information [52, 53]. Moreover, an array of detection systems for molecules and tissue markers have been developed and validated using FFPE tissues including fluorescent *in-situ*

hybridisation (FISH) and chromogenic *in-situ* hybridisation (CISH) [54, 55]. Importantly, frozen section is currently being used in routine BC practice mostly to obtain a quick pathologist decision for surgical margin adequacy and/or axillary nodal status during breast surgical procedures [56].

1.8 Prognostic and Predictive Factors

1.8.1 Clinico-pathologic factors as prognostic factors in BC

1.8.1.1 Invasive carcinoma or in-situ disease

In-situ carcinoma is confined to the ductal/lobular systems. However, cancer deaths associated with *in-situ* carcinoma are due to subsequent development of invasive carcinoma or areas of invasion, which may not be detected at the time of initial diagnosis. Most patients with adequately treated *in-situ* lesions will be cured. In contrast, at least half of invasive carcinoma will have metastasised locally or distantly at the time of diagnosis [57].

1.8.2 Histological type

Histological typing, which refers to the growth pattern of the tumours, is routinely performed for every invasive BC case. Ductal carcinoma, no special type (NST) has a poor prognosis, while tubular carcinoma has an excellent prognosis, other special types of invasive carcinoma fall between these two prognostic extremes [58]. Histological typing is performed by pathologists based on published guidelines and strictly revised criteria [58]. Histological types of invasive breast carcinoma comprise the following:

1.8.2.1.1 Invasive ductal carcinoma, no special type (NST)

NST is commonest category of histologic tumour types and represents about 70% of BC types. It is described as malignant ductal proliferation along with stromal invasion. Tumour cells are pleomorphic, vary in shape and size, and usually have prominent nucleoli and numerous mitoses [59].

1.8.2.1.2 Pure special type

The assignment of a tumour into this category requires 90% purity in terms of the morphologic pattern displayed. These include:

1.8.2.1.2.1 Invasive lobular carcinoma (ILC)

ILC is the second major biologically distinct invasive mammary carcinoma other than NST. It constitutes 5%–15% of invasive breast carcinoma. ILC tumour cells are typically round, small, relatively uniform, non-cohesive and have characteristic growth pattern with single-file infiltration of the stroma [59].

1.8.2.1.2.2 Tubular carcinoma

Tubular carcinoma is a rare, well-differentiated, distinct, special type of BC with an excellent prognosis, constituting about 2% of BC. Microscopically, it is characterised by the proliferation of angulated, oval, or elongated tubules with haphazard arrangement [59].

1.8.2.1.2.3 Invasive cribriform carcinoma

Invasive cribriform carcinoma is a special, rare type of BC associated with favourable prognosis. It is composed of islands of uniform tumour cells with low-grade atypia that have a cribriform appearance [59].

1.8.2.1.2.4 Mucinous carcinoma

Mucinous carcinoma is a rare subtype of BC associated with good prognosis. It accounts for only 2% of total BC. The tumours comprising small clusters of uniform epithelial tumour cells with mild nuclear atypia float in abundant mucus [59].

1.8.2.1.2.5 Invasive papillary and

Invasive papillary carcinoma is also a rare type of BC with invasive elements, showing predominantly papillary architecture whereas the papillae formed by malignant cells are having mild-to-moderate nuclear atypia with delicate fibrovascular core devoid of myoepithelial cells. [56].

1.8.2.1.2.6 Micropapillary carcinoma

Invasive micropapillary carcinoma is characterised by the growth of cohesive tumour cell clusters within prominent clear stromal spaces, resembling dilated angiolymphatic vessels [56].

1.8.2.1.3 Mixed tumour type

This is a relatively common pattern of invasive BC. This category is defined as tumours in which at least 50% of the tumour has a specialised pattern and 10%-90% of the tumour has a nonspecialised pattern [59].

1.8.2.1.4 Other types.

Several rare types of NST BC with an incidence of less than 1% are recognised including lipid-rich carcinoma, secretory carcinoma, Oncocytic carcinoma and adenoid cystic carcinoma [60].

1.8.2.2 Tumour size

Tumour size is one of the most powerful prognostic factors and predictors of tumour behaviour in BC. It is an indication of how long the tumour has been there[61].

1.8.2.3 Histological Grade

Histologic grading is a simple and inexpensive method to assess tumour behaviour and prognosis of invasive BC [62]. Nottingham modification of the Scarff-Bloom-Richardson histological grading of invasive BC is one of the strongest prognostic factors providing significant prognostic information [63]. Grading plays an important role in the selection of patients for adjuvant therapy, being a part of the Nottingham Prognostic Index (NPI) [63-65]. Nottingham grading system is a three-tier system based on reporting on three parameters reflecting the degree of tumour differentiation: degree of tubule formation, mitotic frequency scores and nuclear pleomorphism (**Figure 1-1**). Histological grade is a measure of the biological aggressiveness of the tumour and usually does not change with time [65, 66].

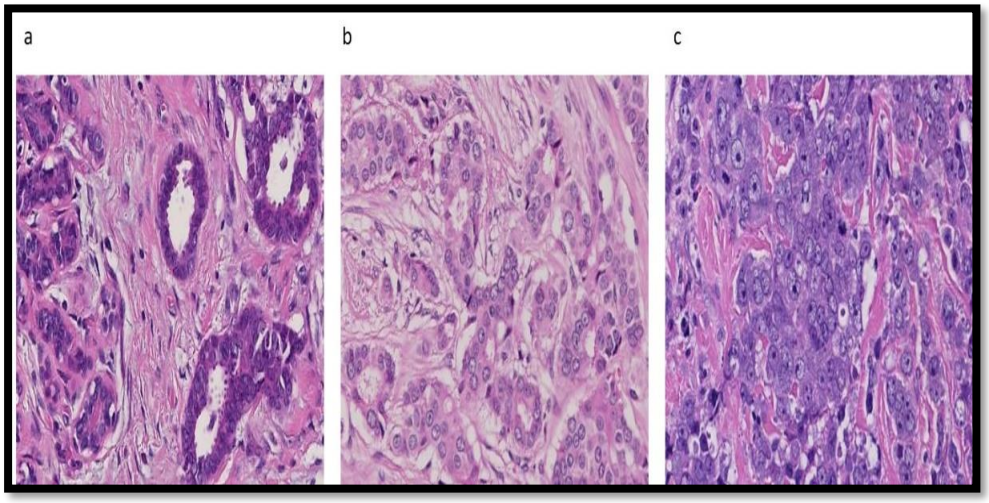


Figure 1-1: Nottingham histologic grading of BC

a- grade 1 tumour shows prominent tubule formation with low degree of polymorphism, b- grade 2 tumour shows increased degree of pleomorphism, c- grade 3 tumour shows high degree of pleomorphism, prominent nucleoli and absence of tubule formation.

1.8.2.4 LN Status

Axillary LN status has been repeatedly shown to be the single most important predictor of the overall survival in BC patients. Twenty percent to 30% of node-negative patients will potentially develop distant recurrence within 10 years of diagnosis, compared to up to 70% of patients with positive axillary nodal involvement [67]. In addition, the number of involved axillary nodes showed prognostic importance. LNs of BC specimens are routinely assessed, and the number involved is reported: tumours with no involved LNs = stage 1, tumours with 1-3 involved LNs = stage 2, and tumours with 4 or more involved nodes = stage 3 [68].

1.8.2.5 Nottingham Prognostic Index (NPI)

NPI is a prognostic stratification tool for the assessment of BC patients comprised of tumour grade and size and LN stage [69]. Each patient's NPI score is based on the clinicopathological tumour and LN scores; the formula used to calculate NPI is as follows: NPI = 0.2 × tumour size (cm) + LN stage (1 to 3) + tumour grade (1 to 3). Using the results of the calculation, NPI is categorised as being 1) good, 2) moderate, or 3) poor. The good prognosis NPI group is defined as patients with NPI score of < 3.4, while the moderate prognosis NPI category covering NPI scores of 3.4–5.4, and all scores exceeding 5.4 fall within the poor prognosis NPI group [70]. Not only does NPI provide prognosis for BC patients, but it also serves in selecting patients who may derive benefit from systemic therapies [71].

1.8.2.6 Lymphovascular Invasion

Lymphovascular invasion (LVI) reflects the presence of tumour cells in lymphatic spaces around the tumour. In BC, LVI correlates strongly with patient survival [72]. LVI is identified by microscopic examination of the morphology of the primary tumour that may or may not have endothelial-specific markers. LVI is a marker of metastatic potential if identified in the primary tumour. Several studies have established the prognostic value of LVI in BC [73-75].

1.8.3 Hormone Receptors (ER and PR)

Current assays use IHC to detect the expression of hormone receptors (HR) in tumour nuclei. Women with HR positive cancers have a slightly better prognosis than with HR negative carcinomas [76]. Eighty percent of tumours with ER and PR expression respond to hormonal manipulation while only 50% of ER+/PR- patients respond to hormone therapy [77]. Patients' responses to adjuvant hormonal therapy can be predicted based upon the co-expression status of ER and PR [78].

1.8.4 Human Epidermal Growth Factor Receptor 2 (HER2)

This is a cell membrane tyrosine kinase receptor, normally involved in the signal transduction pathways leading to cell growth and differentiation [79]. The tyrosine-protein kinase receptor erbB-2, which is encoded by the ERBB2 (HER2) gene, is overexpressed in approximately 15% of BC patients [80]. There is a strong correlation between HER2 overexpression and aggressive BC phenotypes and poor survival [81]. HER2+ BCs can be treated with trastuzumab (Herceptin), which is a humanised

monoclonal antibody that targets HER2+. Thus, trastuzumab is an excellent example of the emerging discipline of precision medicine, where the molecular features of specific tumours can be specifically targeted by therapeutics [82]. However, the effectiveness of trastuzumab declines as the disease progresses and is limited to women with early-stage BC [83].

1.9 BC Molecular Classification

Distinct patterns of gene expression, related to the different BC subtypes, were first reported in 2001 [84]. HR+ and HR^- BC were identified from DNA microarray experiments using hierarchical clustering analysis using the Prediction Analysis of Microarrays (PAM50) gene panel [85]. Clustering analysis revealed that HR+ BC could be sub-categorised according to tumour gene expression profiles. Five classes of tumours were distinguished by clustering analysis: 1) luminal A, 2) luminal B, 3) HER2-amplified, 4) basal-like, and 5) normal breast-like. These subclasses represent the expression profiles of ER, PR, and HER2 as luminal-A (ER+, PR+/-, HER2-), luminal-B (ER+, PR+/-, HER2+), HER2+ (ER-, PR-, HER2+), normal like (ER+ and/or PR+,HER2-), and basal-like (ER-, PR-, HER2-)

(

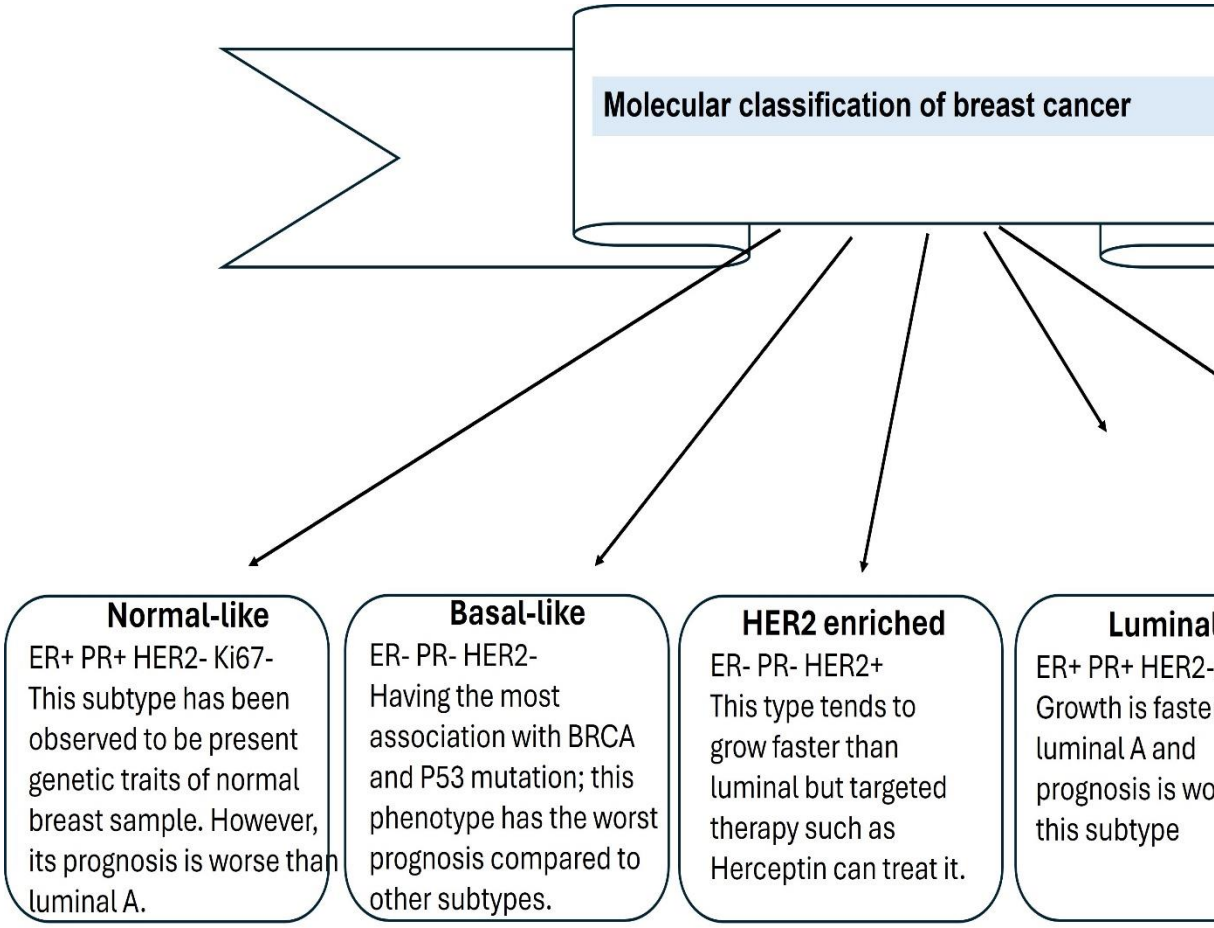


Figure 1-2) [86, 87]. ER, PR, or HER2 receptors are expressed in 15%–54% of Basal-like tumours. This suggests that triple negative (TN) BC and basal-like cancers are not necessarily molecularly similar even though they share a similar histological appearance [88].

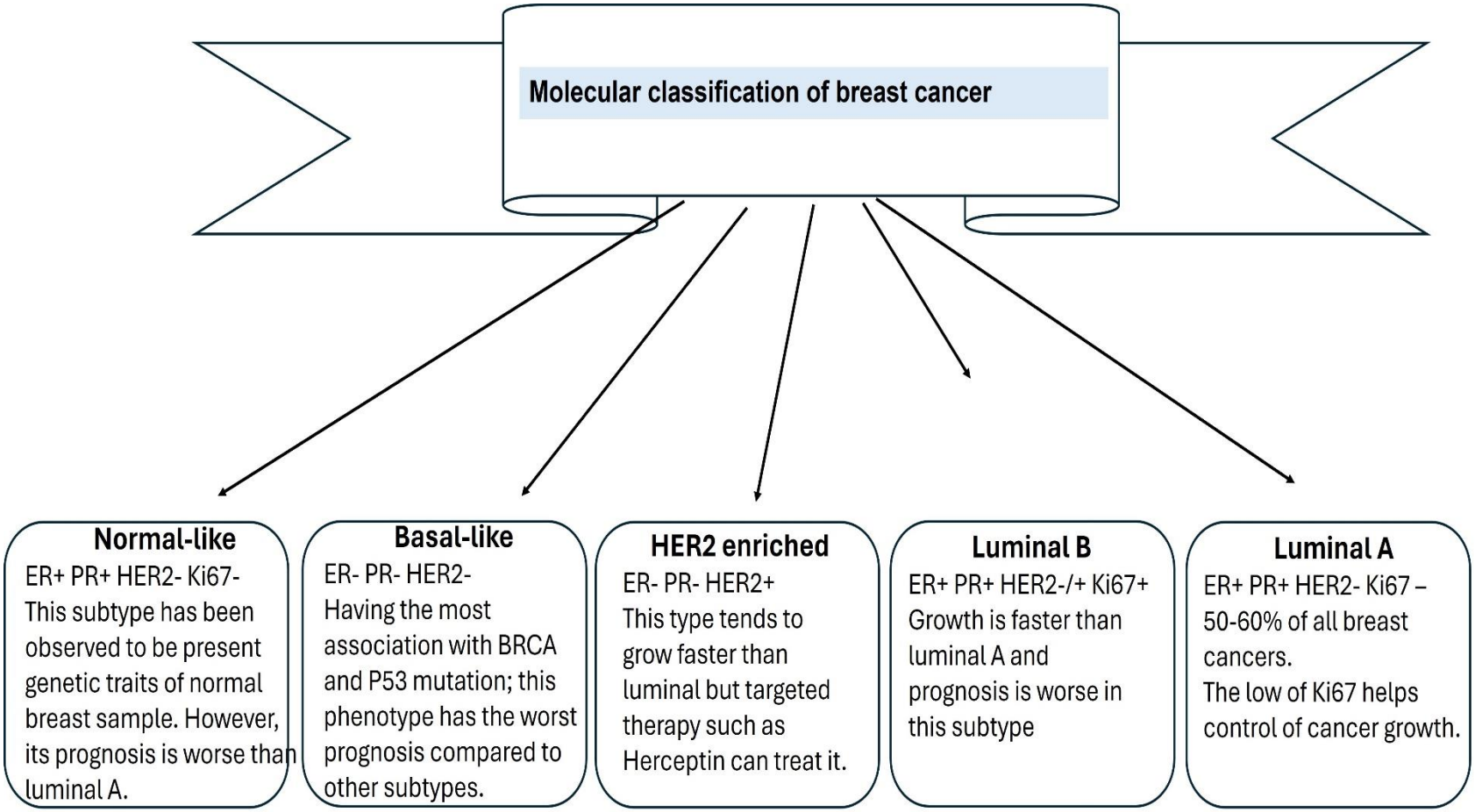


Figure 1-2: Molecular classification of BC

1.9.1 Multi-gene Signatures

Determining which patients with early-stage BC should receive chemotherapy is an important clinical issue. There are several multi-gene prognostic signatures for BC recurrence but there is less evidence that these prognostic signatures are predictive of therapy benefit [90]. Multigene signatures provide information on survival, risk of recurrence, and treatment benefit from chemotherapy or hormone therapy [91]. These gene signatures have been identified using expression profiling technology and different statistical approaches, including the 70 gene signature or MammaPrint, the 76 gene signature the genomic grade index (GGI), Oncotype DX recurrence score (RS), OncoMasTR (OM), MammaTyper assay and EndoPredict assay (EP) [92, 93]. A common character shared by all these signatures is the utilisation of gene panel, rather than using single genes, to predict a certain outcome that reflect the overall genetic derangements underlying the complex tumour biology [94]. Of critical importance is the very minimal to negligible degree of overlap between these signatures regarding their gene sets [95]. Moreover, the prognostic power of some of these signatures is time dependent in the way that their prognostic power is substantially reduced from five to 10 years of follow-up [96].

1.10 Digital pathology and artificial intelligence in BC

In modern clinical practice, digital pathology (DP) has a crucial role in the scientific laboratory environment. DP is defined as utilising a computer to visualise histopathological images in whole slide image (WSI) format, obtained by the high-resolution scanning of tissue samples [97]. The

automated WSI scanner is a robotic microscope capable of digitising an entire histological specimen on a glass slide with the use of specific software to merge or stitch individually captured images into a digital image [98]. It can store and analyse the images and their associated meta-data for diagnostic purposes using general techniques of “pattern recognition”, particularly through “deep learning” (DL) techniques [97]. With availability of faster networks, and cheaper storage solutions, it becomes easier for pathologists to manage digital slide images and share them for clinical use [99]. Many studies on histopathological diagnoses using WSIs have shown non-inferior to microscopic diagnosis [100-102]. It has been shown that a percentage of concordance of between 89 and 99% was detected between using WSIs and the traditional methods, which use slides and optical or electronic microscopy [103]. Moreover, there are many advantages of WSI over glass slides; including easy image accessibility, annotating and manageability to automated image analysis which has been shown to improve objectivity and productivity within pathology practice [104]. A previous study showed that algorithms can outperform human observers in this task and pose a very promising method to overcome this challenge [105]. In addition, WSI can be read by multiple examiners in different locations, allowing remote consultations and reducing the time and financial costs of transferring glass slides [106]. Therefore, WSI integration in different pathology applications facilitates education, consultation, frozen section diagnosis, and clinicopathological conferences in addition to the great improvement of research [107, 108]. Several important guidelines on the validation and

implementation of WSIs for primary diagnosis have been developed by professional bodies such as the College of American Pathologists (CAP), the Canadian Association of Pathologists, the Royal College of Pathologists in the United Kingdom, the Royal College of Pathologists of Australasia, and the American Telemedicine Association [98]. In Europe, several WSI scanners are approved, and in the US, the US Food and Drug Administration (FDA) has approved the Philips IntelliSite Pathology Solution to be marketed [109]. Notably, WSI serves as an enabling platform for the application of artificial intelligence (AI) in DP. AI has become a significant part of the pipeline in research and clinical laboratories [110] **(Figure 1-3)**. A related concept, that has become more popular, is that of “computational pathology”, which regards the computerised analysis of large imagistic data volumes and meta-data, acquired through multiple sources, to extract relevant biological and clinical features and to identify specific patterns [111]. Through the last decade, researchers have developed DL algorithms to detect and classify BCs. A DL model was proposed to automate the multiclassification of BC histopathologic types, such as ductal carcinoma, lobular carcinoma, mucinous carcinoma and papillary carcinoma [112]. Cruz-Roa et al built a convolutional neural network (CNN) model to estimate the extent of invasive foci and the degree of infiltration on entire WSIs [113]. Similarly, several fully automated systems to count mitosis for accurate histologic grading were developed [114, 115]. Recently, Huang and his colleagues developed an AI-assisted LN assessment workflow using CNNs trained on a huge number of LN images [116]. Additionally, computational

assessment of ER/PR IHC using AI algorithms provides opportunities to improve precision performance [117-119]. Furthermore, AI provides new information in predicting prognosis and therapy response in BC [120].

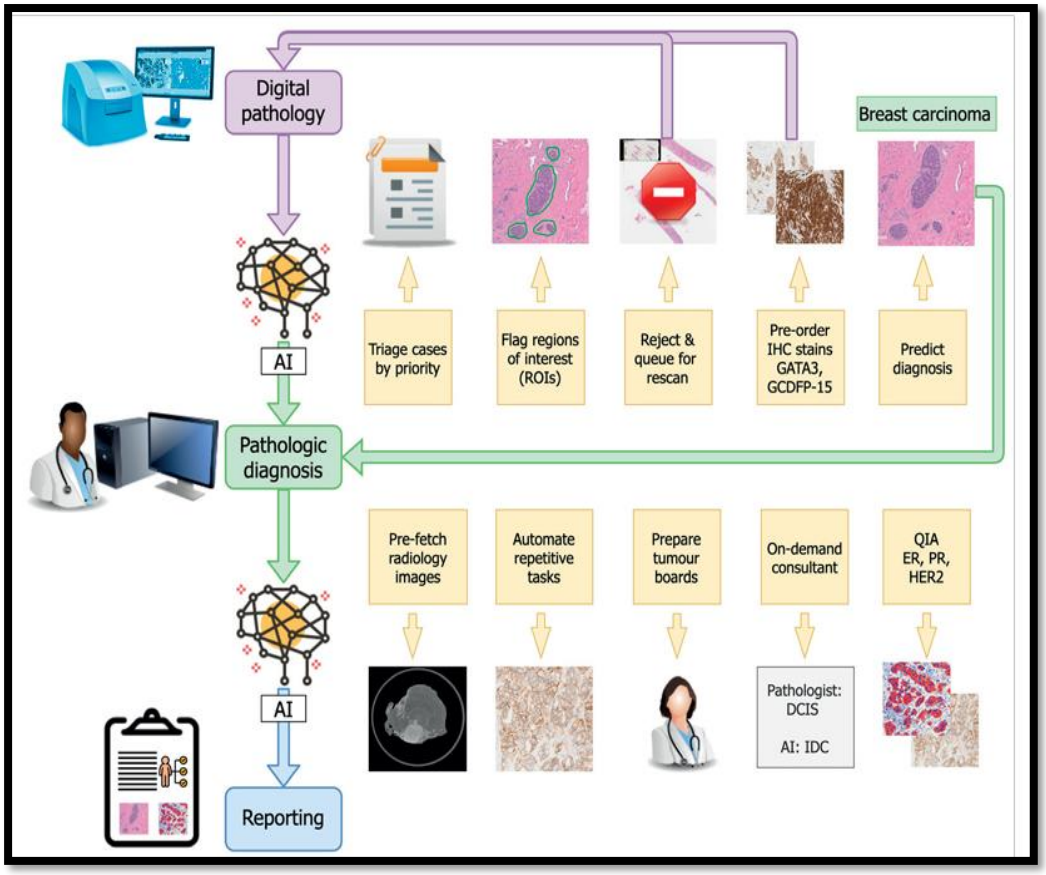


Figure 1-3: The value of AI in the pathological diagnostic process [121]

1.11 Cellular proliferation in BC

The prognosis and aggressiveness of BC are also dependent on the proliferation index, and tumour proliferative activity is considered one of the most thoroughly investigated cellular functions in BC [122, 123]. Uncontrolled cell proliferation is a hallmark in BC and counting of mitotic figures in Haematoxylin and eosin (H&E-stained) sections is the oldest and the most widely used form of assessment of proliferation of BC [124].

Mitotic count is an important component of histopathological grade and forms part of a routine evaluation of BC biology [124]. Proliferation rates can provide useful information on the prognosis and aggressiveness of individual cancers and can be used to guide treatment protocols in clinical practice [125]. Other techniques have been used to estimate cellular proliferation, including bromodeoxyuridine and tritiated thymidine, which were previously considered the gold standard measures of cellular proliferation [126]. The percentage of cells in the S-phase of the replicative cycle can also be measured by flow cytometry [127]. However, these techniques are not easily replicated in clinical practice due to the complexity of requiring fresh tissue. Evaluation of proliferative activity using IHC offers an accessible, cost-effective, and reliable method of assessing tumour cell proliferation in BC [128]. Several proteins have been used to assess proliferation in BC including Ki67, which remains the most used biomarker in the clinical setting [36, 129]. However, adoption of Ki67 in routine practice is still limited.

1.11.1 Phases of cell cycle

Most cells in the body go through a cycle of life in which their genetic information is retained, fixed, and passed down to daughter cells through a highly coordinated and regulated process known as the cell cycle **(Figure 1-4)** [130]. The cell cycle begins with resting cells in the (G0) phase, which are stimulated to enter the active cycle at the first gap (G1) phase. At that point of time, the cell prepares for DNA synthesis (the S phase), which is followed by a second phase of relative inactivity (G2)

and prepares for the separation of the chromatids in the mitotic (M) phase, which is the shortest phase [131].

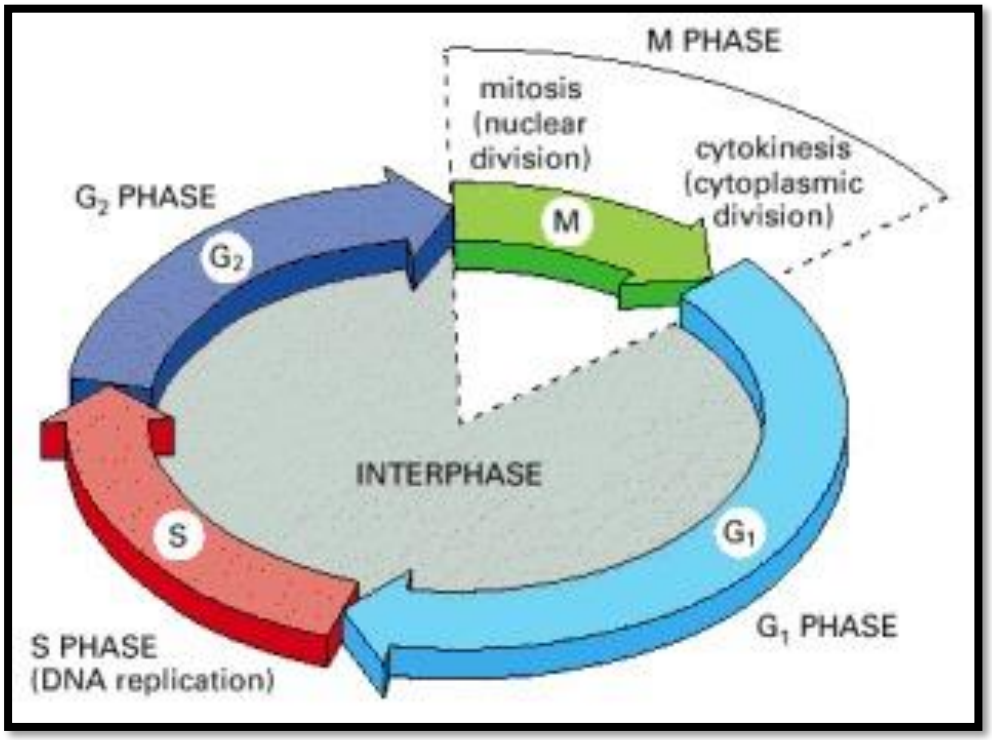


Figure 1-4: Phases of the cell cycle [132].

1.11.2 Mitosis in BC

Mitoses refers to the number of dividing cells visible in H&E-stained slides and it occurs during the M phase of the cell cycle, in which duplicated DNA is organised for equipartitioning. Mitosis is widely acknowledged as a good predictor of tumour aggressiveness through adding information on cellular tumour proliferation [133, 134]. The mitotic count can be reported as Mitotic Index (MI), which is defined as a ratio between mitotic cells and those not undergoing mitosis, in a given area of a tumour [135].

Mitotic activity Index (MAI) is assessed on glass slides using conventional microscopy where mitosis is counted in 10 high power fields (HPFs) or per unit area (2 mm^2) in the most active part of the tumour, mostly at the periphery [136]. As part of the Nottingham grading system, the absolute number of mitotic figures per 10 HPF is transformed into three scores with cutoffs adjusted based on the microscope field diameter, and area size [137]. Several phases of mitosis including typical and atypical forms are identifiable, some of which phases can significantly overlap with other structures such as apoptotic bodies, dark hyperchromatic nuclei and tissue artefacts [138].

1.11.2.1 Phases of mitosis

Mitotic phases are a multistep process where the duplicated chromosomes are aligned, separated, and moved to opposite poles of the cell, which is then divided into two new identical daughter cells including the separation of the cytoplasm. The five phases of mitosis, called prophase, prometaphase, metaphase, anaphase, and telophase [139].

1.11.2.1.1 Prophase

Prophase is the first stage of mitosis, where the chromatin fibres condense into chromosomes that are visible through a light microscope, each replicated chromosome appears as two identical sister chromatids joined at their centromeres, which is necessary in later phases for attachment to microtubules for migration. Furthermore, the centrosomes begin to move to opposite poles of the cell, and they are propelled by the lengthening microtubules between them. In addition, the network of

cytoskeletal components begins to break down and the mitotic spindle forms. During prophase, the nuclear membrane and nucleolus start to disappear [140].

1.11.2.1.2 Prometaphase

In prometaphase, the nuclear envelope disintegrates into small membrane vesicles then microtubules invade the nuclear space and bind to some of the chromosomes. Later, microtubules begin to attach to chromosomal kinetochores, which is a specialised protein structure at the centromere [140].

1.11.2.1.3 Metaphase

Metaphase occupies a large portion of mitosis, and in this phase, the two centrosomes begin pulling the chromosomes towards opposite ends of the cell. The resulting tension causes the chromosomes to align along the metaphase plate, an imaginary line that is centrally located between the two centrosomes. The complete disappearance of the nuclear membrane occurs in this phase [140].

1.11.2.1.4 Anaphase

In anaphase, which is the shortest stage of mitosis, the paired chromosomes (sister chromatids) separate and begin moving to opposite poles of the cell. At the end of this phase, each pole contains a complete compilation of chromosomes [140].

1.11.2.1.5 Telophase

Telophase is the terminal phase of mitosis and is characterised by the chromosomes being turned into distinct new nuclei in the emerging

daughter cells a new nuclear envelope forms around each set of separated daughter chromosomes [140].

Cytokinesis

Cytokinesis is not a phase of mitosis but rather a separate process, necessary for completing cell division, and characterised by the division of the cell's cytoplasm. It begins before the end of anaphase and completes shortly after telophase [140].

1.11.2.2 Abnormal mitosis (Atypical mitosis):

Chromosome segregation errors during cell division can cause genetic abnormalities in neoplastic cells. These errors result in various nuclear abnormalities and can be observed morphologically as atypical mitotic figures [141, 142]. Cytological smears and FFBE embedded tissue samples are useful materials for evaluating atypical mitotic figures [138]. Cells in typical mitosis are often seen in normal tissues exhibiting rapid turnover, such as the epithelium of the gut; however, atypical and bizarre mitotic figures are important morphologic features of malignancy [143]. These abnormal mitotic figures are commonly seen in oral epithelial dysplasia and squamous cell carcinoma and increased numbers of atypical mitosis carry an increasing rate in the grading of cellular dysplasia [144]. For example, atypical mitosis correlates with poor prognosis and outcome of several types of tumours [145-147]

1.11.2.2.1 Classification of atypical mitosis

Atypical mitosis includes two broad categories: mitotic/polar asymmetry and abnormal segregation of chromosomes. More complex aberrations

are derived from combinations of both categories [142]. The common types of atypical mitosis figures are: a- polar asymmetry that is described as unequal sizes of the metaphase axes or anaphase poles where the uneven division of chromosomes to either pole results in two asymmetrical chromosomes, b- tripolar mitosis has three spindle poles resulting in movement of the chromosomal content in three or more directions at anaphase and are easily seen in cytologic and histologic specimens, c- multipolar mitosis has more than three spindle poles at any stage of the mitotic cycle and it is easily appreciable from metaphase through telophase, d- anaphase bridging which is defined as a string of chromatin stretches from one pole of anaphase to the other, e- lagging chromosome has been associated with multiple abnormalities in cellular signalling systems whereas the chromosome or chromatid does not properly migrate during anaphase, f- complex atypical mitoses which can be described as a combination of the previous names [148], g-other atypical mitotic figures (OAMFs), these have a morphological appearance which is not consistent with one of the abovementioned types. OAMFs included ring, asymmetrical, and dispersed mitoses [148].

1.11.2.3 Morphology of mitosis

The appearance of mitotic figures in H&E-stained preparations can be summarised by a few typical characteristics. Mitotic figures appear as hyperchromatic objects without nuclear membrane and hairy extensions of nuclear material and the shape of chromosomes differ according to the phase (**Figure 1-5)** [149].

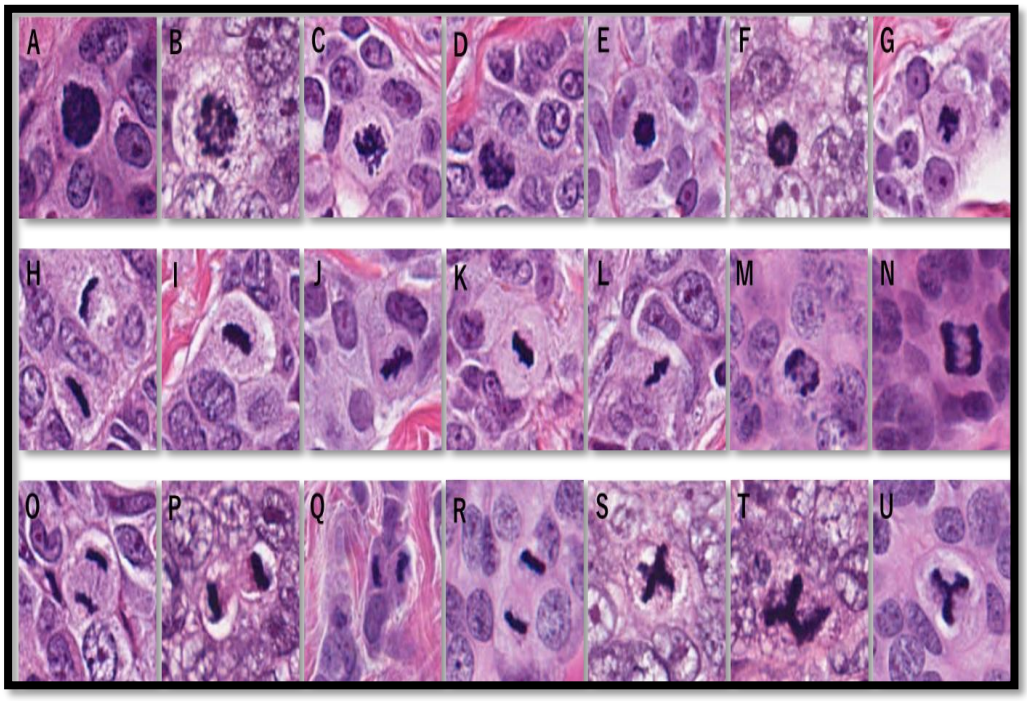


Figure 1-5: Mitotic figures in different phases in WSI (40x)
A-G: Prophase mitoses show hyperchromatic and round-shaped nuclei with hair projection. H-L: Metaphase mitoses show linear to club-shaped hyperchromatic nuclei. M-O: Anaphase mitoses show club-shaped nuclei pulled away from each other. P-R: Telophase mitoses show two linear completed and separated club-shaped nuclei. S-U: Atypical mitoses show irregular-shaped nuclei with enlarged size.

1.11.2.4 Identifying and challenges in counting mitosis

Identifying mitosis depends on certain features including absence of the nuclear membrane, a hairy extension of nuclear material (condensed chromosomes), either clotted (beginning metaphase), in the plane (metaphase/anaphase) or in separate clots (telophase)] [150]. Although it seems as simple terms, mitosis recognition has been a great challenge among pathologists in addition to great inter-observer variability [149]. Furthermore, concordance data from the interpretative National Health Service Breast Screening Programme external quality assurance scheme of breast pathology in the UK indicates a low concordance rate

of mitotic count (κ=0.34) compared with other components of histologic grade such as tubule formation (κ=0.46) [151]. Identification of individual mitotic figures has only moderate agreement among pathologists because the shape of mitotic figures is different in each of the phases (typical and atypical) and some phases have more than one shape. Moreover, the presence of mitotic mimickers like apoptoses, dense hyperchromatic nuclei and artefacts leads to misidentification. These discrepancies in mitotic cell detection may affect grade interpretations. The methodology for the assessment of mitotic figures remains subject to interpretation and individual pathologist performance [152]. Although standardisation of area for mitotic count was published in the 2019 fifth edition of the WHO classification of breast tumours, this was based on defining the area of mitotic count using 10 HPF [59], rather than defining the best size area to be used for counting and then adjusting the number of HPF based on the microscope field diameter. Therefore, standardisation of how to identify mitosis and sample area would increase the reproducibility of mitotic counts and improve BC grading by the reporting pathologist [153].

1.11.2.5 Mitosis identification using digital pathology.

The establishment of WSI as a medical diagnostic device allows pathologists to evaluate mitotic activity with this technology. Mitotic cell detection has limitations regarding accuracy and reproducibility and with the advantage of digital imaging, these limitations may be sorted out. For example, algorithms can outperform human observers in this task and pose a very promising method to overcome this limitation in mitosis

identification and counting (**Figure 1-6)** [154]. Despite the high concordance that was detected between WSI and light microscopy (LM), most pathologists do not have as detailed knowledge of the advantages and limitations of WSI analysis compared to their expertise in pathology using traditional microscopy. Additionally, few studies have reported on the agreement between the use of WSI and microscopy for analysis of histological features at the level of individual cells (e.g. mitotic figure quantification) [155]. Although the problems with different HPF sizes in microscopes may be solved in the DP era, the field diameter selection, area chosen for counting and whether a representation of either peripheral or central parts should be included or not still needs more investigation. The difficulty in identifying mitotic figures digitally has been discussed in many studies [156-158], and it has been shown that not all mitotic figures seen using microscopes could be identified on WSI. Moreover, the differentiation from different mimickers using WSI is still challenging [159]. Williams et al [160] pointed out the main two reasons for difficulty in discerning mitotic figures in digital images; a- less contrast between chromatin and the background on digital slides and b- the absence of the fine focus on suspected mitotic figures on digital slides, a function that can be performed on glass slides to confirm the identity of mitoses. Moreover, other challenges like cellular level variations in colour, intensity, and morphological shape/size, can also increase the challenges [161]. Additionally, the use of different scanners and adjusting their calibration may affect the counting of mitosis. Although the reproducibility (intra-observer agreement) is similar for Nottingham grade

using glass slides or WSI, the tendency of lower mitotic scores in WSI was pointed out [160, 162].

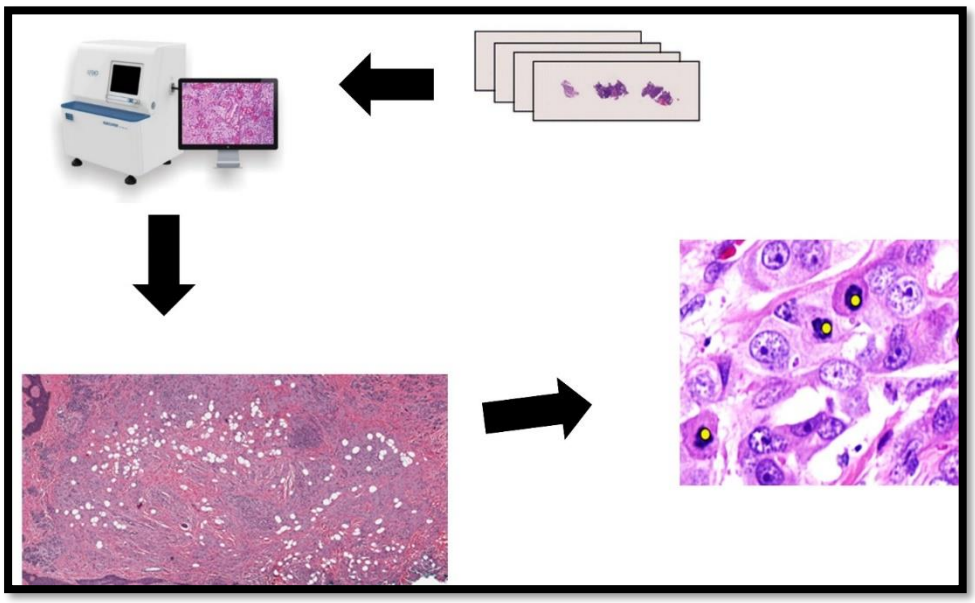


Figure 1-6: Process of digitalisation of glass slides into WSI for mitosis identification

1.11.3 Ki67 in BC

Ki67 protein has been widely used as a proliferation marker for BC for decades [163, 164]. It was first identified as an antigen in Hodgkin lymphoma cell nuclei structure [165]. Although topoisomerase IIα, phosphor-histone H3 (PHH3), proliferating cell nuclear antigen (PCNA), and minichromosome maintenance (MCM) proteins have been used to assess proliferation in BC, Ki67 remains the best IHC marker to assess proliferation (**Table 1**)

Table 1: Comparison between Ki67 and other proliferative markers in BC

Markers	Ki67	PCNA	PHH3	MCM family	Topoisomerase II
Function	Required during interphase for normal cellular distribution of heterochromatin antigens.	Involved in DNA excision repair, cell cycle control, and RNA transcription	Involved in chromatin condensation.	Essential to ensure DNA is replicated only once per cell cycle.	Required to segregate replicated chromosomes.
Cell cycle	Expressed in all cell cycle phases except in G0.	Expressed in S and G phases.	Expressed in the M phase only.	Increased during the G1/S phases	Expressed in S/G2/M phases.
Prognostic role in BC	Correlation with poor disease-free survival (DFS), overall survival (OS) and BC-specific survival (BCSS).	Association with a shorter DFS and OS.	A prognostic marker in patients with LN-negative BC. Correlation with short DFS.	Association with OS, BCSS, regional recurrence, and distant metastases free survival (DMFS).	Association with a shorter DFS and OS.
Disadvantages	No standard scoring protocol or accepted cutoff until now.	Not specific for proliferation.	Expression is affected by fixation time.	MCM is not superior to Ki67 in BC	Not exclusively expressed by proliferating cells.
Correlation with ki67		Poor correlation.	Poor correlation.	Strong correlation.	Strong correlation.

1.11.3.1 Functions of Ki67.

Ki67 is a labile non-histone protein, involved in the early steps of polymerase I-dependent ribosomal RNA synthesis [166, 167]. Ki67 is a nuclear antigen expressed in all proliferating cells, during cell cycle phases S, G1, G2, and M, but not in G0 [167]. During interphase, Ki67 is located in the nucleolar cortex and the dense fibrillar components of the nucleolus, whereas it becomes associated with the periphery of condensed chromosomes during mitosis **(Figure 1-7)**. Ki67 levels are low in the G1 and S phases with peak levels during mitosis, especially metaphase, then sharply decrease during anaphase [168]. Similar to other cell cycle markers, Ki67 is affected by external factors such as cellular nutrient deprivation, which leads to an underestimation of the number of cycling cells [169]. The half-life of the Ki67 antigen is between 1-1.5 hours, regardless of the stage of the cell cycle [170].

Ki67 is required during interphase for normal cellular distribution of heterochromatin antigens and the nucleolar association of heterochromatin [163]. Ki67 binds with DNA and its protein-bound complexes to promote chromatin compaction [171]. During mitosis phase, the chromosomes become tightly packed structures and coated by a layer of proteins, called the perichromosomal layer. Its absence led to problems with the ability of daughter cells to synthesise protein building blocks and increased rate of spontaneous cell death [172]. Notably, Ki67 is essential for the organisation of perichromosomal layer, and it is one of the first proteins to join the perichromosomal layer after the beginning of mitotic phase. When Ki67 is depleted, chromosomes in cells appear to

lack most or all their perichromosomal layer [171]. These features make Ki67 one of the best markers to assess cell proliferation for its crucial role in the cell cycle.

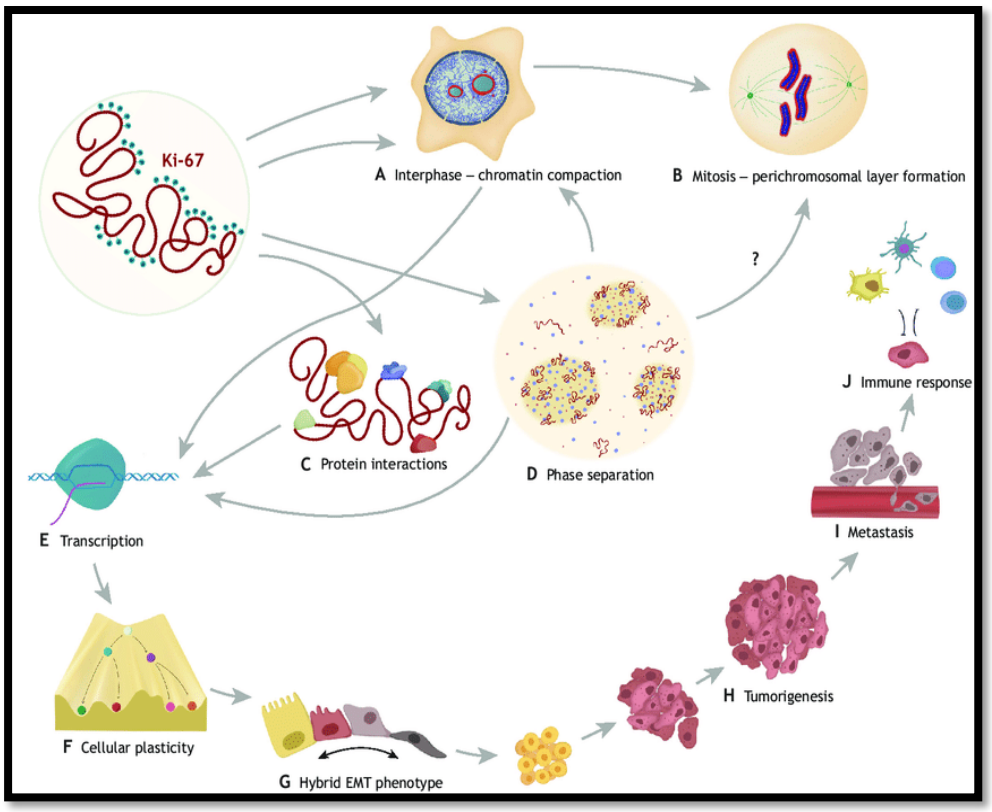


Figure 1-7: Overview of Ki67 localisation and functions.

During interphase, Ki67 is localised to peri-nucleolar heterochromatin and regulates chromatin compaction (A). In mitosis, Ki67 relocates to the surface of chromosomes (B,). Through protein-protein interactions (C) it helps in separation of subnuclear compartments (D), Ki67 regulates transcriptional programmes (E) for cellular plasticity (F) and, it is required to maintain a hybrid EMT in cancer cells (G), for tumour development (H) and metastasis (I), as well as anti-tumour immune responses (J) [173].

1.11.3.2 Ki67 antibodies

Several antibodies are used for Ki67 antigen detection, including Ki67 and MIB1 monoclonal antibodies directed against different epitopes [174]. MIB1 has a higher sensitivity for detecting the Ki67 antigen [175]. Other antibodies against Ki67 are also available, such as SP6 which recognises

the same epitope as MIB1 and has been used to assess proliferation in BC [176, 177]. However, one study has shown that SP6 is not superior to the MIB1 antibody with a head-to-head comparison of both markers [177]. Therefore, MIB1 continues to be widely used for Ki67 assessment.

1.11.3.3 Role of Ki67 in cell division

The role of Ki67 in cell division has been confirmed in *in-vitro* studies from the arrest of cell proliferation observed with Ki67 blockage either by antibodies or by inhibition of serine and threonine dephosphorylating [171]. Moreover, Ki67 binds with satellite DNA and presents in the protein complexes that bind to DNA [171]. This was confirmed by acute Ki67 depletion, which caused DNA damage, and the damage occurred even in the absence of checkpoint signalling; this damage is robustly detected in mitotic cells [178]. Other study showed that Ki67 is involved in the early steps of ribosomal RNA (rRNA) synthesis [179] and in the stabilization and maintenance of the mitotic spindle during mitosis, by recruiting Hklp2 to the mitotic chromosomes [180].

1.11.3.4 Localisation and patterns of Ki67 IHC staining.

Ki67 is a nuclear protein and other non-nuclear staining patterns are not considered in the scoring [181] and their significance is still unclear [182, 183]. Cross-reactivity with other proteins, technical artefact, or re-localisation of Ki67 within the cell are possible explanations [182]. There are different patterns of nuclear Ki67 expression observed in BC including homogeneous staining pattern of the nucleoplasm, a granular pattern, which stained nucleoli or granules of different sizes dispersed

throughout the nucleoplasm and a mixed pattern that showed strongly stained granules against a diffusely positive background of a lower staining intensity [184]. The clinical relevance and impact of these distinct Ki67 staining patterns in BC remain poorly understood.

1.11.3.5 Intensity of Ki67 staining

The intensity of nuclear staining of Ki67 ranges from weak, moderate, to strong staining [185] and this may reflect the fluctuation of Ki67 levels in the different phases of the cell cycle **(Figure 1-8).** However, the clinical value of Ki67 intensity of staining has not been well explored and the nuclear Ki67 positivity of any intensity is considered in the scoring.

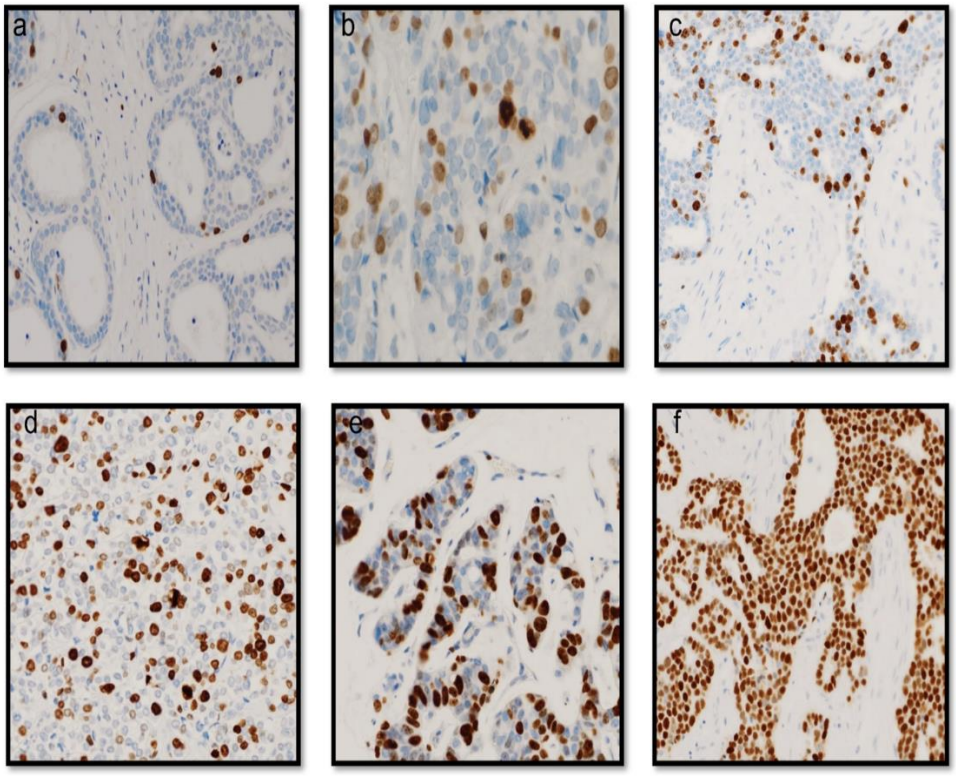


Figure 1-8: Representative expression of Ki67 in BC.

Very low Ki67 expression (a), moderate Ki67 expression (b,c), high Ki67 expression (d,e) and very high ki67 expression in almost all BC cells (f) (x20).

1.11.3.6 Heterogeneity of Ki67 staining

The intra-tumoural heterogeneity (ITH) of Ki67 staining across biological BC subgroups, especially in luminal tumours was previously studied [186]. Some BC types are characterised by low Ki67 expression such as tubular and invasive cribriform carcinoma, classical lobular carcinoma, and other grade 1 special type carcinomas while other tumour types show variable degrees of expression from low to high such as NST tumour type. These tumours also show ITH [187, 188]. This has been explored using a Ki67 heterogeneity index (the difference between Ki67 expression levels in hotspots and global Ki67 expression in the whole tumour) [189]. This may reflect the presence of highly proliferative clones in heterogeneous tumours compared to homogeneous tumours, in which the majority are low proliferative. This heterogeneity also supports the use of hotspots to assess Ki67 score which should represent the most proliferative part of the tumour rather than the average proliferative activity which will be lower than that of the hotspots.

1.11.3.7 Ki67 expression in core biopsies and excision specimens

The concordance of Ki67 expression in core biopsies and full-face sections has been a point of interest. A recent systematic review involving 22 studies revealed a wide concordance kappa coefficient range from 0.2 to 0.7 between both biopsy types [190]. However, different methodologies and cutoffs were used to explain the results. Ki67 has been reported to be more highly expressed in core than excision biopsies [191] which may be related to fixation issues, particularly in cohorts enriched with cases

that had suboptimal fixation of surgically resected specimens. This result supported the recommendations of the International Ki67 in BC Working Group (IKWG) for testing Ki67 on core biopsies in general. However, other authors demonstrated that Ki67 is higher in the surgical specimen than in the paired baseline core biopsies, which is expected to consider the intra-tumoural heterogeneity of Ki67 and the distribution of its expression in hotspots [192]. The heterogeneity in Ki67 expression denotes that those full-face sections, under good fixative protocols, are the optimal platform to assess Ki67 properly and avoid the underestimation of Ki67 expression in heterogeneous tumours which is common. However, a core biopsy can be similar to or better than excision specimens in the assessment of Ki67 expression when the fixation of the excision specimen is suboptimal or in tumours at the end of the spectrum of expression (low or diffusely high expression) [193]. Core biopsy is ideal to test the tumour proliferative activity in the setting of neoadjuvant therapy as the residual tumours' proliferative activity can be influenced by the treatment effect.

1.11.3.8 Methodological challenges in Ki67 assessment

1.11.3.8.1 Pre-analytical setting:

Several pre-analytical issues affect Ki67 expression such as type and duration of fixation [194, 195]. Optimal Ki67 results are obtained when 10% neutral buffered formalin is used while prolonged fixation may cause a reduction in Ki67 expression levels [196]. Importantly, long-term storage in paraffin affects the accuracy of Ki67 assessment as it is also more sensitive to antigen decay [197]. It is recommended that samples

for Ki67 staining should be processed according to the American Society of Clinical Oncology (ASCO) and CAP guidelines for HR and HER2 to preclude many pre-analytical challenges [198].

1.11.3.8.2 Analytical and post-analytic issues of Ki67 assessment

The degree of inter-laboratory variability in Ki67 staining results from differences in staining methodologies, including staining platform, antigen retrieval, primary antibody, detection system, and counterstain [198]. However, this degree of inter-laboratory discordance is not different to ER and PR staining and other similar diagnostic and prognostic IHC markers in routine practice and should not be used as a major criticism against its clinical use.

There are many post-analytic issues of Ki67 assessment, whereas cell counting for Ki67 is the most challenging task for pathologist. Another caveat is assessment using hotspots or using an average value across specimens, counting non-tumour cells such as tumour infiltrating lymphocytes or non-invasive tumours, such as ductal carcinoma *in situ* (DCIS) which may contribute to scoring problems.

1.11.3.9 Scoring methods

IHC scoring is commonly based on the percentage of stained tumour cells reported as the Ki67-labelling index ; however, less labour-intensive Ki67 scoring methods have also been proposed [199]. The common scoring methods include assessing Ki67 in 500 or 1000 tumour cells, either in whole tissue sections or focussing in highly stained ‘hotspot’ areas [36, 200-204]. Others have used a global subjective assessment

of Ki67 positivity by rapidly scanning and estimating the percentage of Ki67-positive nuclei at high power magnification (termed the “quick-scan” method) [205, 206]. Some authors have also suggested scoring Ki67 in 200-400 cells in 10 high-powered fields at 40x magnification [207] while others have recommended assessment of the whole tissue section at low power and recording the overall average percentage score [208]. Recently, IKWG has suggested calibrated standardised visual scoring using online scoring applications [209]. However, it requires routine digitalisation of the stained sections, which are not available in all pathology laboratories.

Depending on the degree of the homogeneity of staining, previous studies advocated that Ki67 should be assessed in at least three randomly selected HPFs at 40x objective [36, 210, 211] or in two peripheral along with one central heterogeneous HPFs [212]. However, in heterogeneous staining, scoring is suggested to be in three HPFs from the tumour edges or hotspots [36]. However, adopting these approaches does not consider the density of the different tumours, which has an important impact on the assessment in Ki67 [189].

1.11.3.10 Ki67 cutoffs for clinical use

A consistent cutoff is crucial when considering the prognosis for luminal BC patients. Cutoffs to distinguish “Ki67 high” from “Ki67 low” in luminal BC vary from 1% to 29%, which is one of the perceived limitations of Ki67 use in the clinical setting. In 2011, the Saint Gallen Consensus Meeting defined the “low proliferation” tumours to have Ki67 <14 % [213], which

was proposed based on a study by Cheang and colleagues [214]. In the 2013 Saint Gallen Conference, the majority voted that a threshold of ≥ 20% was indicative of “high” Ki67 status. However, in the St. Gallen conference of 2015, the majority favoured a cutoff value of 20–29% [215]. A previous meta-analysis of 64,196 patients showed that cutoff at least 25% had better discriminatory power compared to other cutoffs in luminal BC [216]. Another study included 8088 patients and used automated Ki67 scoring to conclude that Ki67 >12 % had a worse 10-year BCSS in ER-positive BC patients [217]. Moreover, Ki67 has been compared with EndoPredict risk scores and Ki67 above 25% had a strong overlap with the high-risk EndoPredict test [218]. Therefore, it seems that cutoff levels ranging from 10% to 30% have been the most common to dichotomise populations [219, 220]. Al Eskandarany et al demonstrated that a 10% cutoff for Ki67 could stratify BC with grades I and II into statistically significant prognostic groups [221]. Others recommend using the median value as the cutoff [215]. The IKWG recently concluded that a Ki67 index of 5% or less, or 30% or more, can be used to estimate prognosis in early BC [209]. However, there is still a large gap between 5% and 30% and a large percentage of patients would be involved in this zone.

In TNBC, a 40% cutoff was considered an optimal cutoff of Ki67 to classify patients with a greater risk of recurrence and death compared with patients with lower expression rates [222]. Other study found that 30% is the relevant cutoff value for Ki67 for the prognosis of TNBC [223]. Although different cutoffs for Ki67 were proposed in TNBC, it is cohort-based and with limited prognostic value.

Also, the Ki67 cutoff should be changeable according to tumour type. For example, a 2% cutoff can predict the risk of developing BC in atypical hyperplasia [224]. Also, patients with Ki67>10% after pre-operative endocrine therapy are considered a candidate for further adjuvant treatments in the POETIC clinical trial [225]. Furthermore, a 20% Ki67 cutoff was accepted by the FDA for the addition of CDK4/6 inhibitors with endocrine therapy to luminal high-risk BC patients [226] and also for prediction the response of CDK4/6 inhibitors [227].

1.11.3.11 Ki67 and histologic tumour grade

Histologic tumour grade is determined by assessing morphological features according to the Nottingham grading system [228]. Therefore, histologic tumour grade may be indirectly related to Ki67 expression based on the mitotic count. Furthermore, grade and Ki67 appear to be complementary rather than independent prognostic factors [229]. Evaluating the association between Ki67 and tumour grade in BC has been controversial where some claim that high-grade tumours are correlated with high expression of Ki67 within the different molecular subtypes [230-232]. However, others did not find any association despite using a large cohort of cases [233].

1.11.3.12 Ki67 and mitosis

Ki67 and MI are indicators of tumour proliferation rate and both are significantly correlated [36]. Analogous to Ki67, the advancing edge of the tumour, probably in the areas of highest mitotic counts, is likely to be representative of tumour clones that drive the biologic potential of the

tumour [200, 234]. Mitotic count, which can easily be assessed in a defined area on H&E-stained slides, represents the dividing cells while Ki67 positive cells represents those in the cell cycle. However, not all cells in the cell cycle will proceed to the mitosis phase and become committed to cell division. Although grade is one of the important prognostic parameters in BC that attracted a lot of attention, mitosis counts as part of the 3 components of grade has not attracted similar attention to Ki67 as a measure of proliferation [163]. This may be related to several factors including the low concordance rate among pathologists in assessing mitotic scores and the lack of evidence on its prognostic significance as a standalone parameter in BC. The prognostic significance of mitotic count in the clinical setting remains a component of grade. However, after improved standardisation of mitotic count in BC and defining the area of counting rather than the number of HPF may improve concordance of assessment. In addition, it will be possible to develop a mitotic cutoff of prognostic significance similar to Ki67 rather than the 3 scores of grade [235]. In contrast to MI, which is subject to fixation factors, Ki67 can represent the proliferative activity in tumours with suboptimal fixation and can provide an alternative for mitotic count during assessment of BC grade in sub-optimally fixed tissue [221].

1.11.3.13 Ki67 and tumour size

Previous studies reported a significant association between Ki67 expression and tumour size, where smaller tumours had lower Ki67 values, and the larger tumours were associated with higher Ki67 expression [230, 236].

1.11.3.14 Ki67 and histologic tumour types

High Ki67 expression was significantly associated with NST [212]. However, ILC, which represents 5–15% of the invasive breast tumours, showed a lower Ki67 index, and some investigators consider this an explanation for the good behaviour of lobular carcinoma [237]. Notably, rare types of BC like the lipid-rich and sebaceous BC typically present with high Ki67 levels [237, 238].

1.11.3.15 Ki67 and LN status

In BC, the spreading of primary tumour cells to axillary LNs is a crucial step in the process of metastasis [41]. It has been suggested that Ki67 is one of the markers that can predict LN metastasis [239]. Many studies found that the incidence of LN metastasis was higher in high Ki67 positive group than in low Ki67 positive group [230, 240-242]. In a previous study, it has been shown that Ki67 assessment in LNs is superior to its evaluation in primary BC as Ki67 expressions >10% in the metastatic LNs showed an association with poor survival while in primary tumours, no association was found [243]. However, others reported only a marginally significant relationship between LN metastasis and Ki67 expression [244].

1.11.3.16 Ki67 and hormonal receptor (ER, PR) and HER2

The relationship between Ki67 index and HR was studied in previous studies. In most of these studies, a negative correlation was found between HR and Ki67 levels [230, 245, 246]. Regarding, the correlation between Ki67 and HER2 status, it was controversial, as some of the researchers have found a negative correlation between Ki67 and HER2

[247] while others found a positive correlation [248, 249]. This may be due to the difference of Ki67 positive cutoffs used. In some studies, the cutoff of Ki67 positivity were greater than 30% positive cells, while in others, it was less than 12 or 25 % positive cells [250].

1.11.3.17 Ki67 and molecular subtypes

Luminal tumours show a wide spectrum of proliferative activity, this is reflected in the ability of Ki67 to stratify HR-positive BC into prognostically distinct groups based on the proliferative activity. The Saint Gallen Guidelines recommend Ki67 to distinguish the molecular subtypes luminal A and luminal B BC subtypes [251]. The majority of HR-negative BC are highly proliferative. A correlation is seen with high Ki67 index and TNBC [201] and high Ki67 expression is significantly associated with aggressive clinical behaviour in TNBC [252].

1.11.3.18 Ki67 and Oncotype DX recurrence score

Due to its prognostic significance in HR-positive BC, *MKi67* gene is one of the main 16 genes that are measured in the Oncotype DX RS [253-255]. Oncotype DX RS which generates a recurrence score between 0 and 100 that correlates to the likelihood of disease recurrence within 10 years of diagnosis [256, 257]. There is a high concordance rate between Oncotype DX RS and Ki67 index [192, 258, 259]. In addition, a previous study showed that high Ki67 tumours are at higher risk of relapsing among patients with low-risk Oncotype DX, and Ki67 status may help to identify a subset of low-risk Oncotype DX patients who could benefit from

adjuvant chemotherapy [260]. In addition, Ki67 proliferation value is a major determinant and predictor of Oncotype Dx score [261].

1.11.3.19 Ki67 and Nottingham Px prognostic index

Nottingham Px is a prognostic index incorporating tumour grade, tumour size, PR and Ki67 with a formula calculated for BCSS and DMFS. Nottingham Px was able to stratify the challenging group of early BCs into two clinically relevant classes more effectively than the conventional NPI [262]. In addition, this helps decisions regarding therapy plans, where the high-risk class can be considered as candidates for chemotherapy or further testing using multigene assays [262].

1.11.3.20 Prognostic role of Ki67 in BC

Elevated levels of Ki67 expression in BC correlate strongly with the more tenacious proliferation of BC cells and several studies have shown the relationship between Ki67 expression and poor outcome in BC patients [36, 230, 263]. A meta-analysis, based on 46 studies including 12 155 patients, confirmed that high Ki67 levels confer an adverse effect on OS and DFS [264].

1.11.3.21 Predictive role of Ki67 in BC

It has been suggested that Ki67 can serve as a tool to identify BC patients who especially benefit from chemotherapy or endocrine therapy [265]. Patients with luminal A tumours characterised by low proliferation can avoid the toxic side effects of chemotherapy and receive hormonal therapy only. While those with luminal B subtype are associated with a high risk of recurrence and poor outcome, therefore, the addition of

chemotherapy could be associated with a recurrence risk reduction [266]. Ki67 could justify the use of adjuvant chemotherapy in the setting of patients with highly proliferative tumours.

1.11.3.22 Ki67 and adjuvant therapy

Adjuvant systemic therapy improves outcomes for patients with early-stage invasive BC [267, 268]. The selection of adjuvant endocrine therapy or adjuvant anti-HER2 therapy is based on ER and HER2 status. Selection of appropriate patients for adjuvant chemotherapy following surgery remains one of the greatest challenges during the management of early BC patients, and the choice of adjuvant chemotherapy is principally based on the estimation of prognosis and residual risk [198]. Ki67 is an important part of a prognostic algorithm for residual risk in early BC patients treated with endocrine therapy [269]. Previous studies have found that Ki67 is a marker of chemosensitivity for BC as most cytotoxic agents require cells to be in the cell cycle [198]. Furthermore, previous studies showed that positive responses to certain forms of therapy can be predicted with high Ki67 scores, such as some chemotherapy combinations (e.g., docetaxel, fluorouracil and epirubicin) in ER-positive tumours [270]. However, some studies found that Ki67 alone cannot predict the benefit of adjuvant chemotherapy in LN-negative BC patients [271, 272].

1.11.3.23 Ki67 and neoadjuvant therapy

1.11.3.23.1.1 Neoadjuvant endocrine therapy

In patients with luminal tumours, endocrine therapy accounts for most of the gains obtained with adjuvant systemic treatment. Neoadjuvant

endocrine therapy is a logical alternative to neoadjuvant chemotherapy for ER-positive tumours for many reasons including toleration, simplicity to deliver, and proven benefits [225, 273]. Previously, many neoadjuvant trials had shown that after a 2-week on-treatment, Ki67 predicted outcomes better than its baseline and, unlike baseline Ki67, it was significant in multivariable analysis [225, 273]. Therefore, Ki67 can provide a simple and inexpensive test to personalise adjuvant treatment in patients with hormone receptor-positive, and HER2-negative BC.

1.11.3.23.1.2 Neoadjuvant chemotherapy

Neoadjuvant chemotherapy, is an approach whereby patients receive chemotherapy before surgery and it is a standard treatment strategy for patients with local advanced and also operable BC [274]. Therefore, it is important to predict the response to neoadjuvant chemotherapy and identify patients who might benefit from this type of therapy. Two studies showed that high Ki67 has been associated with a good response to neoadjuvant chemotherapy [275, 276]. It has been observed that patients with high post-treatment Ki67 levels were at higher risk for disease relapse and death compared with patients with low or intermediate Ki67 levels. Also, post-treatment Ki67 levels provided more prognostic information than pre-treatment Ki67 levels or changes in its levels from pre- to post-treatment. [275]. Importantly, a meta-analysis of 36 studies showed that Ki67 might be an independent predictor for pathologic complete response (pCR) [277]. However, not all studies agree [278, 279].

1.11.3.24 Ki67 and artificial intelligence

The traditional scoring method of Ki67 staining by IHC may be time-consuming, poorly reproducible for many pathologists, and liable to inter/intra-observer variability [280, 281]. Fortunately, in the era of WSI technology, it is now possible to combine histopathological image analysis with AI [282]. This may enable highly accurate and rapid workflows [283]. However, the question arises whether AI could be used to solve the problem of accurate Ki67 assessment on IHC stained sections. To date, the development of such automated approaches has been limited, including a lack of automated selection for regions of interest or the hotspot areas [280]. Stalhammar et al reported that automated Ki67 assessment methods are superior to manual assessment in terms of sensitivity and specificity and can improve inter-observer reproducibility of Ki67 assessment [284]. However, automated scoring methods have some defects in the identification of invasive tumour cells, especially in lymphocytic-rich tumours where some Ki67-positive lymphocytes may be identified as tumour cells leading to its overestimation [210]. To overcome this, a semi-automated evaluation method of Ki67 index which allows for the determination of the exact proliferation index value by marking tumour and non-tumour cells manually, followed by counting the cells automatically has been proposed [210]. Also, new software in development could overcome this problem through virtual double staining for tissue classification. For example, the same tissue section is stained for both cytokeratin and Ki67; tumour cells are recognised by positive cytokeratin expression, and only cells that co-

express both markers are automatically counted as positive tumour cells [284, 285]. Efforts should be exerted to develop a standard methodology and recognise constant cutoffs with the development of computer-assisted AI guidelines, which can improve the reproducibility of Ki67 assessment [286].

1.12 Progesterone receptor (PR) in BC

Estrogen and progesterone, and their nuclear receptors (ER and PR), are involved in BC development [287]. Molecular crosstalk between ER and PR occurs where the PR gene is regulated by ER as an ER-dependent gene product. PR expression is dependent on estrogen and at the same time, PR can modulate ER action [288]. PR expression is generally regarded as a surrogate for a functional ER pathway and hence, the presence of PR means that the ER pathway is intact. However, some ER positive/PR negative tumours respond to ER antagonists suggesting that the ER pathway remains functional and that other regulatory mechanisms are involved [289]. In a study on 666,852 HR-positive BC, 82.9% were ER-positive/PR-positive, 15.0% were ER-positive/PR-negative, and 2.0% were ER-negative/PR-positive [290]. PR expression in ER+ tumours is associated with a good prognosis [291, 292]. Several multigene signatures contain PR as one of the main component genes in their testing gene set [1]. PR levels by IHC have been used by various prognostic tools, such as Magee equations, the American Joint Committee on Cancer (AJCC) eighth edition prognostic stage groupings, and nomograms that predict the Oncotype DX RS results [293], in

addition to various predictors of response to neoadjuvant therapy [294, 295].

1.12.1 Function of PR in BC

PR is expressed in only a fraction of the luminal epithelial cells, while in the basal epithelium PR is not expressed. Progesterone upregulates receptor activator of nuclear factor kappa beta (RANKL) expression by binding to its receptor in PR-positive breast luminal cells. Then, RANKL binds to RANK expressed on the surface of neighbouring PR-negative luminal cells or basal cells, activating downstream pathways of cell proliferation, expansion, and survival [287]. The paracrine mediators of progesterone-mediated proliferation include Wnt family member 4 (Wnt4), RANKL, amphiregulin, calcitonin and the chemokine C-X-C ligand 12 (CXCL12**) (Figure 1-9)** [287]. PR can drive mammary gland development by expansion of the mammary stem cell population [296].

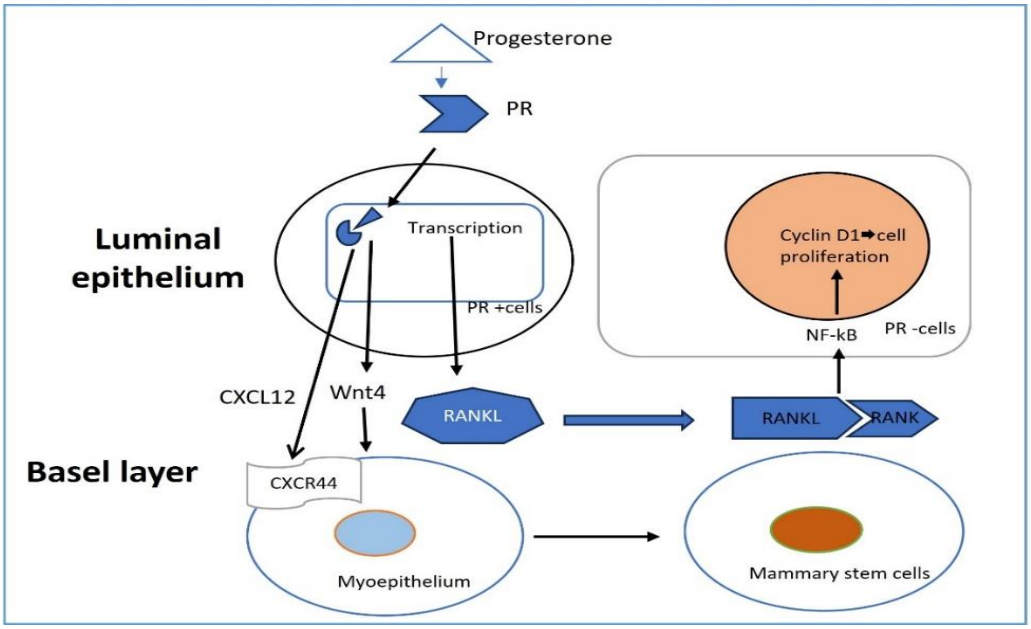


Figure 1-9: PR/RANKL signal in human breast.

Progesterone increases RANKL in PR-positive luminal epithelial cells mainly by inducing the transcription of RANKL mRNA. Then RANKL induces proliferation of neighbouring PR-negative luminal cells by activating NF-κB and cyclin D1 pathways. RANKL is also involved in transition of myoepithelium to mammary stem cell. PR can mediate CXCL12-CXCR4 and Wnt Pathway [287].

1.12.2 PR assessment in clinical routine practice

PR is also routinely assessed by IHC; however, testing for PR in BC remains optional in some countries and centres. In 2010, ASCO/CAP guidelines [195] recommended that ER and PR status should be evaluated in all BC while the 2020 guidelines reported that PR is used mainly for prognostic purposes in ER-positive BC. They also recommended that the principles of ER testing apply to PR [293]. However, the UK Royal College of Pathologists guidelines and some others regard PR testing in BC as optional [297]. A recent paper showed PR should be assessed routinely for several reasons including that PR staining can provide a quality assurance measure for ER staining [227]. In addition, PR status can be used for the molecular classification of BC into luminal A and B subtypes and it provides prognostic significance similar to Ki67 expression in BC [227].

1.12.3 PR in molecular subtype of BC

BC subtypes that expressed genes characteristic of luminal epithelial cells, including ER and PR, and those that were negative for these genes were identified. Luminal-B BC has been defined as ER-positive BC with increased proliferation [298]. High expression of PR is more frequently observed in tumours with a better baseline prognosis (ie, luminal A) than in tumours with a poor baseline prognosis (ie, luminal B) [299]. It was

found that ER+/PR- tumours, as defined by gene expression profiles, share gene expression patterns with both ER+/PR+ and ER-/PR-; and bear resemblance to the luminal B subtypes [300].

1.12.4 PR and proliferation in BC

It was suggested that PR helps regulation of cellular proliferation in BC and a significant association between PR and Ki67 was reported [301]. In addition, previous studies showed that PR was an independent prognostic factor in patients with rapidly proliferating BC [302, 303].

1.12.5 PR and endocrine therapy

Endocrine therapy is generally well tolerated and an easy-to-administer option for HR-positive BC patients. PR-negative, ER-positive tumours have a worse prognosis and are associated with endocrine resistance [287].PR-negative, ER-positive tumours consistently respond less well to endocrine therapy than PR-positive BC, although some of them do benefit [304]. Therefore, hormonal therapy alone is avoided in ER-positive/PR-negative BC. Low PR expression may be a predictive benefit of adding chemotherapy in pre- and perimenopausal ER-positive patients [305].

1.12.6 PR cutoff

PR positivity is usually observed in 65-75% of BC, and it is seen in about 80% of ER positive cases [306]. This proportion varies depending on the cutoff used for PR positivity. A few studies focussed on validating IHC assays for PR and the cutoff of positivity [288, 292]. 1% cutoff has been recommended for PR similar to ER positivity [293]. However, this

seems very low in view of the different clinical indications and the use of both ER and PR status in BC. Notably, no evidence was published to suggest that tumours with low PR (1-10%) show response to endocrine therapy in ER negative BC or that endocrine therapy response of ER positive tumours is significantly different if PR is expressed at a low level [307, 308]. Previous studies showed that the prognostic value of PR was validated using a cutoff of 10%, 20% or greater [292, 308-310]. Therefore, the need to explore an optimal cutoff for PR positivity in luminal BC is warranted.

1.13 Prognostic markers in BC

Because of the well-established clinical and pathologic heterogeneity of BC, the main limitations of current markers remain sensitivity and specificity [187, 311]. There is an urgent need to identify new biomarkers that demonstrate improved prognostic value to better and more precisely inform individualised response to therapy [312].

In this study, 3 potential prognostic markers were identified based on the genes that were associated with atypical mitosis (see chapter 6). These markers include:

1.13.1 Histone H1

Chromatin-related proteins have emerged as important players in the initiation and maintenance of cancer [313]. In addition to the established role of histone-modifying enzymes and chromatin remodelling in promoting and maintaining cancer cells, it was shown that the basic components of chromatin such as histone proteins, in particular the linker

histone H1 (H.H1), display both transcriptional and genetic alterations in a variety of cancers [314].

H.H1 family genes cluster on chromosome 6 and translate to 11 isoforms [315]. H.H1 interacts with DNA and is involved in the formation and maintenance of higher-order chromatin structures. It plays role in gene and cell cycle regulation [316, 317]. H.H1 is widely expressed in many different cell types in a replication-dependent manner with a peak of expression in S phase and therefore expressed at high levels only in dividing cells [318]. H.H1 seems to regulate DNA methylation and hence gene expression [319]. In addition, it displays both transcriptional and genetic alterations in a variety of cancers. Previously, using clinical samples from different cancers, it was shown that the decrease in H.H1 expression is a common feature of neoplastic cells and the low H.H1 levels can predict poor prognosis [320]. It has been shown that decreased H.H1 expression is a common feature of neoplastic cells in lymphoma, and it was recognised that H.H1 acts as a tumour suppressor [321, 322]. However, high H.H1 expression was associated with malignant potential of prostate cancer [323].

1.13.2 Polo-like kinase1 (PLK1)

Polo-like kinases (PLKs) are important members of the serine/threonine protein kinase family, which control many crucial biological processes, including regulation of the cell cycle [324]. *PLK1* is a key component of *PLK* family, it promotes neoplastic transformation, and it is critical for cell division. Moreover, it has a key role in all stages of mitosis including

centrosome maturation, bipolar spindle formation, chromosome segregation, and cytokinesis. In addition, *PLK1* plays a role in DNA replication during S phase, and it acts as a modulator of the DNA damage response and maintenance of genome stability during replication [325, 326]. In tumours, overexpression of *PLK1* enables tumour cells to override mitosis checkpoints leading to genetic instability [327, 328]. Therefore, *PLK1* was considered a potential oncogene, and it is overexpressed in a variety of tumours and tumour cell lines. *PLK1* inhibitors have recently been developed and are being tested as potential anticancer agents [329, 330]. However, recent data showed that down regulation of PLK1 can induce aneuploidy and tumourigenesis, where lower levels of PLK1 cause improper segregation and delay in centrosome maturation that led to aneuploidy and tumourigenesis and hence it acts as a tumour suppressor gene [331]. The fact that PLK1 may function as a tumour suppressor instead of an oncogene does not necessarily argue against the use of PLK1 inhibitors in cancer therapy. Many essential components of cell proliferation may be used as cancer targets despite having no oncogenic activity, owed to non-oncogene addiction of cancer cells for specific cellular processes such as cell division [332].

High PLK1 expression is observed within proliferating normal tissues, including placenta and colonic epithelium and in many types of cancer, including gastric [333], colorectal [334], hepatocellular [335], prostate [336], ovarian [337], and non-small cell lung carcinomas [338-340]. In BC, elevated PLK1 expression was associated with aggressive tumour

characteristics, such as vascular invasion, markers of proliferative activity and lack of detectable ER [341].

1.13.3 Minichromosome maintenance complex component 7 (MCM7)

DNA replication is a central procedure of cell proliferation, whereas aberrant DNA replication is indicated to be a driving force of oncogenesis [342]. Minichromosome maintenance (MCM) complex is considered to be a critical first step in the formation of a pre-replication complex during the initiation stage of DNA replication, and also has a key role in cellular proliferation [343]. The MCM complex is composed of six highly conserved MCM proteins, namely MCM2–7 and several MCM proteins are highly associated with tumourigenesis [344, 345]. MCM7 plays an important role in initiating DNA replication it is also considered as a marker of proliferation in different cancers [346-348] including prostate cancer [349], lung cancer [344], and hepatocellular carcinoma (HCC) [344]. In addition, high MCM7 expression was associated with a negative outcome in patients with HCC. Therefore, inactivation of this gene could be a promising therapeutic target in a variety of cancer types [344]. Indicators of potential outcomes include the suppression of cell proliferation, as shown in HCC *in vitro* following downregulation of MCM7 [347]. Meanwhile, MCM7 amplification and overexpression has been shown to promote cell proliferation, colony formation and migration in oesophageal carcinoma [350].

1.14 Tumour heterogeneity in BC

Tumour heterogeneity is one of the hallmarks of malignancy [351]. BC displays striking clinical, morphological, and behavioural diversity within a single tumour and between tumours. Morphological heterogeneity of BC reflects an underlying spectrum of genetic and epigenetic portraits that control BC behaviour [351]. BC shows molecular heterogeneity, not only between tumours (inter-tumour heterogeneity) but also among the same tumours (intra-tumour heterogeneity) [352]. Both inter-tumour and intra-tumour heterogeneity imply that each BC could have a different prognosis and would benefit from specific therapy [352]. For each individual patient's tumours, heterogeneity can also be classified into spatial and temporal. The spatial heterogeneity involves distinct areas within a tumour that have differences at the phenotypic, transcriptomic, epigenetic and genomic levels [352]. Temporal heterogeneity refers to variations occurring over time during tumour progression between primary and metastasis or among different metastatic lesions and it is also important when comparing between pre-invasive and invasive disease within the same tumours [353]. Heterogeneity in BC is the result of a co-ordinated interplay between the tumour cells themselves and between the tumour cells and the different stromal elements [3]. Understanding features of heterogeneity is a fundamental requisite to understand how the tumour progresses.

1.14.1 Inter-tumour heterogeneity

Inter-tumour heterogeneity has been associated with different BC histological subtypes, treatment sensitivity response, and clinical

outcomes [354]. Inter-tumour heterogeneity of BC is best illustrated by clinical staging of the disease based on physical examination and imaging findings. Standard BC treatment is based on the tumour characteristics, including clinical stage, histopathologic features, and biomarker profile. The morphologic heterogeneity of BC constitutes the basis for the histopathologic classification of BC into several histologic types, differ in their morphology and patient outcome as well [60]. Grading of BC also highlights its tumour heterogeneity, whereas BCs of different grades also show different profiles by proteomic, genomic and transcriptomic analysis [77, 355]

According to ER and PR status, Ki67 proliferation index and HER2 expressions, BC are grouped in Luminal-A, luminal B+, HER2+ and triple-negative. These molecular subtypes have different prognosis and different responses to therapy. For example, the presence of ER positivity or HER2 positivity leads to endocrine therapy benefit or target-therapy sensitivity, respectively [352]. In addition, it has been shown that each BC subtype has distinctive molecular hallmarks; whereas *TP53* mutations were more common in TNBC, *PIK3CA* mutations were more frequent in luminal BC [356].

1.14.2 Intra-tumour heterogeneity in BC

Intra-tumour heterogeneity (ITH) can be classified according to the molecular and morphologic levels.

1.14.2.1 Genetic heterogeneity

BC shows ITH in terms of chromosomal and genomic alterations. Genetic heterogeneity affects many processes and functions, such as signalling pathways, antitumor immunity, cell senescence, migration and metastasis, angiogenesis, treatment response, and metabolic pathways [357]. Different cell clones can either segregate in different areas of the tumour or scatter within the same area. Genetic heterogeneity is characterised by bulk sequencing and single-cell or single-molecule sequencing [358].

1.14.2.2 Epigenetic heterogeneity

Epigenetic heterogeneity refers to modifications in gene expression without DNA sequence changes. Epigenetic heterogeneity is caused by variations in regulatory mechanisms, including epigenetic, posttranscriptional, and posttranslational modifications. Morphologic heterogeneity

In BC, it is common to find well-differentiated areas adjacent to poorly or moderately differentiated areas. Attempts are currently being made to quantify these areas and grade tumours according to the least-differentiated area or the area with the highest degree of cytologic malignancy [360]. Breast tumour grade varies within the same tumour, especially mitoses and atypia [60].This intra-tumoral differentiation is often patchy, and not well defined at the molecular terms and can lead to differences in the expression of some of the therapeutic targets and, therefore, the response to a specific treatment [361]. Morphological

heterogeneity of BC was previously described in the most common histological type (NST).Tumour microenvironment heterogeneity

TME is the cellular environment within which the tumour exists including the surrounding blood vessels, either pre-existing or newly formed, immune cells including lymphocytes, macrophages, and other inflammatory cells, fibroblasts, and extracellular matrix [363]. It has been shown TME has an impact on patient outcome and the cross-talk and interaction between tumour and TME contributes to early tumourigenesis as well as disease progression beside the response to therapies [364]. Different types of stroma are associated with BC. Some tumours are associated with a florid desmoplastic reaction such as Basal-like carcinoma, while other tumours, such as lobular carcinoma, do not initiate a stromal response. This likely represents different mechanisms of tumour stromal interaction [351]. **Figure 1-10** shows illustration of ITH in BC.

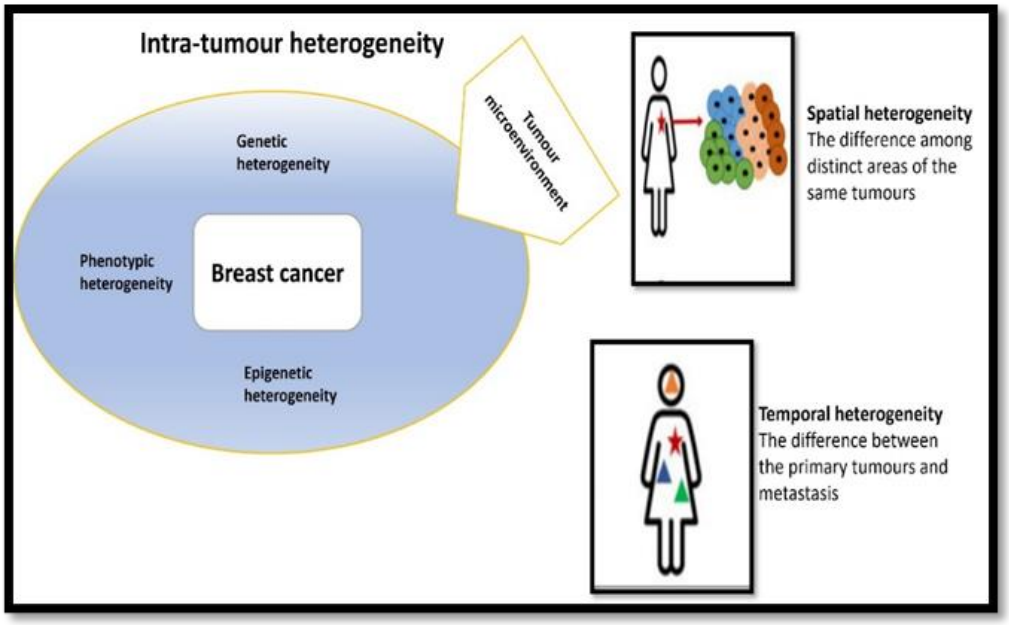


Figure 1-10: Overview of BC ITH [352]

1.14.3 Assessment of ITH

How to integrate the knowledge of ITH into clinical practice remains to be a major challenge. It is well known that a more heterogeneous population of cancer cells can behave more aggressively, because of probable presence of clones that can proliferate more rapidly. However, how to ideally quantify heterogeneity and whether such quantitative measures correlate well with clinical outcome have yet to be defined. Assessment of ITH and tumour-host interaction within the TME is becoming increasingly important for innovative cancer therapy decisions because of the unique information it can generate about the state of the disease [365]. Although extensive sequencing approaches, for example multiregional sequencing [366], single cell sequencing [367], ultra-deep sequencing [368] provide detailed information for understanding intra-tumoral heterogeneity, these techniques are not readily applicable in routine clinical practice [369]. Similarly, next-generation sequencing coupled with skilled quantification was able to measure intra-tumoral heterogeneity in BC patients [370], but again its clinical application can be difficult. The progress in DP, especially the computational power of machine-learning methods, brings opportunities to quantify morphologically features that can be subvisual by routine microscopy assessment, which is an imperative task for intra-tumoral heterogeneity [365]. There are different computer-aided methods, ranging from hand-crafted features to deep-learning neural networks, have been proposed

for diverse tasks in tumour pathology [371]. In addition, these methods provide heat maps of feature expressions within the tissue that augmented visualisation for decision support. It was shown that deep-learning enables quantification of the immune abundance (lymphocyte fractions to other cells) based on spatial distributions and proximities between lymphocytes and/or cancer cells in BC [365].

1.15 Hypothesis and aims of the thesis.

The field of BC pathology is challenging with discordance noted even amongst expert pathologists. Using digitalised WSI in clinical pathology extends the frontiers of the pathologist's view beyond a microscopic slide, The advent of DP and AI can increase the concordance among pathologists and refine the prognosis of BC patients. The hypothesis of this thesis is:

Expression patterns of current and novel tissue biomarkers at the translational level in BC will act as useful prognostic or predictive tools to be integrated into simplest robust systems for BC prognosis suitable for use in routine clinical laboratories. In this thesis, I focussed on assessment of proliferation in BC using digital pathology.

Therefore, this thesis aims to assess the accuracy of DP and AI in interpreting subtle histological features including mitotic figures in order to improve grading of BC in routine clinical practice. In addition, it aims at refining tissue biomarkers using digital DP and AI to increase the reproducibility of these markers.

The specific aims of this thesis are:

1- Collect the clinical information and outcome data of a large series of BC patients diagnosed at Nottingham University Hospitals from 2007 to 2018.

2- Construct Tissue Microarrays (TMAs) from representative sections of BC according to the molecular classification of BC.

3- Define the morphological features of mitoses to enable reliable and reproducible identification of mitotic figures in WSI compared to LM.

4- Evaluate the clinical and biological significance of atypical mitosis in BC using WSIs.

5- Compare between various methods of visual Ki67 assessment while considering practicality, reproducibility, and the prognostic and predictive value.

6- Characterise the quantitative differences between BC cells in the cell cycle (represented by Ki67 expression levels) and those in mitosis (represented by mitotic count).

7- Assess Ki67 using AI-based tools and compare the manual with the automated scores.

8- Define the positivity of PR expression and evaluate its prognostic and predictive significance in ER+ luminal BC.

9- Compare the expression of PR using visual versus automated assessment.

10- Explore a range of potential biomarkers closely relevant to BC biology and progression.

11- Assess the morphologic ITH in BC using DL models.

Chapter 2 Patients and methods

2.1 Patient cohort and data collection

A large cohort (n = 5847) of patients with primary operable BCs presenting to the Nottingham Breast Institute, Nottingham City Hospital, Nottingham, UK between 1999 and 2018, was included in this study. This cohort is well-characterised with a long-term clinical follow-up with a median of 138 months and ranged from (0-216 months). Full detailed clinicopathological data were available, including patient age at diagnosis, menopausal status, tumour size, histological tumour grade, histological tumour type, molecular subtypes and NPI risk groups **(**

Table 2**)** [354]. In addition, outcome data were also available in terms of BCSS, defined as the time (in months) from the date of the primary surgery to the time of death from BC, and DMFS, defined as the time (in months) from the primary surgery until the first event of DM. For the purpose of different aims, the cohort was classified into subcategories. The study cohort for each project is described in detail in each chapter. Adjuvant treatment was scheduled based on patients' tumour prognostic and predictive factor status including the NPI, ER status, and menopausal status. Before 2000, patients within the good prognostic group (NPI ≤3.4) did not receive adjuvant therapy after surgery. Adjuvant endocrine therapy was prescribed to patients with ER positive tumours and NPI scores of >3.4 (moderate and poor prognostic groups). Therefore, a significant number of patients in this project did not receive adjuvant therapy. However, from 2000 onward, all patients with ER+ tumours were offered adjuvant endocrine therapy. Premenopausal patients within the moderate and poor NPI prognostic groups were candidates for chemotherapy. ER negative patients received chemotherapy. HER2 enriched BC patients were treated with Herceptin None of the patients in this cohort received neoadjuvant therapy. The data for patients, who were diagnosed from 2007 to 2018 were collected from patients records by the author and Dr Wakkas Fadel.

Table 2: **Clinicopathological characteristics of the study cohort.**

Parameter	Number of cases (%)
Patient Age	
≤ 50 years	1548 (27)
> 50 years	4265 (73)

Menopausal status	
Premenopausal	3598 (63)
Post-menopausal	2100 (37)
Tumour size	
≤ 2cm	3849 (66)
> 2cm	1989 (34)
Tumour grade	
Grade 1	1157 (20)
Grade 2	2825 (48)
Grade 3	1865 (32)
Histological tumour type	
No special type (NST)	3564 (61)
Lobular carcinoma	707 (12)
Other special types	343 (6)
NST mixed	1233 (21)
Lymph node status	
Negative	4117 (70)
Positive	1726 (30)
Lymphovascular invasion (LVI)	
Absent	4626 (79)
Present	1221 (21)
ER receptor status.	
Negative	664 (12)
Positive	5125 (88)
PR receptor status	
Negative	1594 (29)
Positive	4040 (71)
HER2 status	
Negative	5240 (91)
Positive	380 (8)
Molecular classes	
Luminal	3549 (76)
HER2 enriched.	433 (9)
Triple negative	717 (15)
Nottingham prognostic index	
Good prognostic group	2632 (45)
Moderate prognostic group	2705 (46)
Poor prognostic group	500 (9)

2.2 Histological review

Histological review was performed by the author and two experienced pathologists (Dr Mohammed Aleskandarany and Dr Wakkas Fadhil). Histologic review was started by retrieving FFPE blocks from storage archives. The blocks that were marked as having tumour tissue were

selected where between 4 to 6 blocks were retrieved for each case. 1 full-face 4-µm thick section was cut from each block and stained with H&E. A quality check was performed where slides with air bubbles, no invasive tumour, extensive tissue folding or slides with poor staining were failed. Re-sectioning of failed slides was performed in case of air bubbles or poor staining. Presence of DCIS was scored into 3 grades, whereas 0= no DCIS, 1= DCIS was within the invasive tumour and 2 = DCIS was 1mm beyond the invasive tumour. Tumour infiltrating lymphocytes (TILs) were assessed by estimation of the percentage of TILs within the invasive tumours. Tumour types and histologic grade were assessed and recorded if there was a difference compared to those originally assigned in patient record. Specific findings were also assessed and recorded including, biopsy site changes, areas of haemorrhage, calcification, previous TMA empty cores and fibrocystic changes.

2.3 Slide digitalisation and image acquisition.

After completing the histologic review, the H&E-stained slides were scanned into digital images. High resolution digital slide scanning was achieved with the Pannoramic digital slide scanner (250 Flash III, 3DHISTECH®, Budapest, Hungary), at 20x magnification. This was followed by viewing the images using Pannoramic Viewer software v 1.15.4 (3D HISTECH®, Budapest, Hungary) to check the quality of the scanning. Failed slides on histologic review were excluded and slides with improper scanning were rescanned.

2.4 Tissue Microarray Construction

Expression of tissue markers potentially involved in tumour progression in BC was assessed using IHC. To assess the expression of these markers in a large number of samples, tissue microarray (TMA) was chosen as the appropriate tissue format. TMA technique enables high-throughput analysis of biomarkers expression due to the capacity to manage several hundred samples in a single slide and being able to apply techniques such as IHC [372]. Using TMA saves a considerable amount of tissue, time and reagents. In addition, it is an economical technique, making the assessment quicker and easier to perform [373]. Furthermore, it improves laboratory standardisation because all tissue samples are exposed to the same experimental conditions [374]. Preparation of TMA involves careful steps: collection of the study cohort to be arrayed, preparation of the recipient array blocks, designing the array grids, construction of the individual array blocks, and preparation of the array block for sectioning [375]. The initial tasks of preparing donor blocks and TMA design are the most critical parts of TMA construction and are therefore a quite time-consuming task. Details of TMA construction process is described in detail in the appendix.

2.5 Optimisation of antibodies used in the study.

2.5.1 Validation of antibody specificity by Western Blot

Western blotting (WB) was performed to determine the specificity of the antibodies before IHC staining. A mixture of cell lysates from 6 cell lines: MCF-7, MCF10A- MD-MBA-231, MD-MBA-293, SKBR3, and

HeLa that were obtained from the American Type Culture Collection; Rockville, MD, USA was used to reflect the expression of the selected proteins. Dr Abrar Aljohani did the cell culture part.

2.5.2 Preparation of samples, electrophoresis and transfer

Cell line lysate was prepared in a maximum volume of 20 µl (13µl of lysate, 5µl of NuPAGE LDS sample buffer (4x), and 2µl of reducing agent, (10x), all reagents were bought from Thermo Fisher Scientific. This mixture was heated for 5 minutes at 100 °C and left to cool on ice. The prepared sample was loaded onto 10% SDS–NuPAGE (4-12%) (Thermo Fisher Scientific) in 20x MES SDS running buffer (Thermo Fisher Scientific). The pre-stained Thermo Scientific Protein Ladder LC5602, (Magic ladder) was used as a molecular weight standard and gels were run at 170 volts for 60 minutes. Proteins were transferred to a nitrocellulose membrane by wet transfer and electro-blotting at 10 volts for 120 minutes in transfer buffer. To ensure that the transfer was successfully performed, the membrane was put in 1ml balance boat containing Ponceau S solution SLS (Sigma) P7170-1L.

For blocking, the membranes were incubated in 5% non-fat dry milk (Marvel dried skimmed milk, UK) for 60 minutes. The blocking solution was prepared by adding 2.5g of milk to 50ml washing solution (Phosphate Buffered Saline (PBS, Sigma) containing 0.1% Tween-20 (Sigma, P 2287, Polyethylene glycol sorbitan monolaurate, Polyoxyethylene sorbitan monolaurate, Polysorbate 20). The

membrane was washed three times with PBS Tween-20 and incubated with the primary antibody overnight at 4°C using the optimal concentration as instructed by the manufacturer **(Table 3).** Membranes were washed with PBS Tween-20 3 times, each for 10 minutes. The membranes were incubated with the secondary antibody in a dark room. Secondary antibody was used (anti mouse or anti rabbit, Licor Biosciences). The membrane was washed three times with PBS Tween-20. The membrane was visualised using Odyssey Infrared Imaging System (LI-COR Biosciences, Nebraska, USA) using Enhanced Chemiluminescence (ECL) solution, 1705061, (Bio-Rad). Mouse β-Actin (A5441, Sigma-Aldrich; Clone AC-15; Sigma, UK) or Rabbit β-Actin (SAB5600204, Sigma-Aldrich; Clone RM112; Sigma, UK) were used as a house-keeping protein at concentration of 1:5000 and 1.15000, respectively.

Table 3: Specificity of the selected antibodies for the project

Antibody	Molecular Weight	Concentration	Supplier
Histone H1	30	1:500	Thermo Fisher Scientific
PLK1	60	1:1000	Abcam
MCM7	80	1:1000	Thermo Fisher Scientific
Mouse β-Actin	42	1: 5000	Sigma-Aldrich
Rabbit β-Actin	42	1: 15000	Sigma-Aldrich

2.5.3 Optimisation of antibodies by IHC

For each antibody, optimisation of the staining conditions was performed by changing several variables in the staining protocol until the optimal result was obtained, which was considered as the specific subcellular localisation for the antibody with no or faint background staining. These variables include a- dilution of the antibody, starting with the dilution recommended by the manufacturer plus dilutions above and below, b- antigen retrieval solution: microwave pre-treatment in citrate buffer at pH 6.0, or water bath pre-treatment in EDTA buffer (AR9640, Bond^{TM} Epitope Retrieval Solution 2), pH 9.0, c- Incubation period for antibody started with the duration recommended by the manufacturer, then change the period according to the obtained results of staining ranged from 30 minutes to overnight incubation. For optimisation of antibodies, the slides were stained using Sequenza trays (Thermo Scientific).

To test the specificity of an antibody, cells should show immunoreactivity within the specific cellular location (membranous, cytoplasmic, or nuclear) with no or minimal background staining, which allows the most accurate evaluation of protein expression status. All the staining procedures including optimisation were performed by the author except ER and PR staining in core biopsies, which were performed by the Histopathology Department, Queen's Medical Centre, Nottingham. In addition, Ki67 and PR full-face slides were automatically stained in

Source Bioscience lab, Nottingham. For comparison, the author has stained a batch of Ki67 and PR manually (n=50).

2.5.4 IHC technique

IHC was performed on 4µm full face tissue or TMA sections mounted on permanent positive charged glass slides (Surgipath, X-tra™ Adhesive, Peterborough, UK). For the studied antibodies, sections were finally stained using Novolink™ Polymer Detection System (RE7150-K). To affirm adhesion of tissue sections with the slide surface, sections were placed on a hot plate (60 °C) for 20 minutes and left to cool. Slides were dewaxed in xylene (Genta Medica, York, UK) twice for 5 minutes, rehydrated twice in graded Industrial Methylated Spirit (IMS) or methanol (2 minutes/each) followed by 5 minutes in running tap water. Heat induced retrieval of antigen epitopes was performed, when required, by Whirlpool JT359 microwave (1000 W) treatment of the slides in 1 litre of sodium citrate buffer (pH 6.0), using full power for 20 minutes. Slides were cooled for 10 minutes under continuous stream of tap water and were rinsed with Tris Buffered Saline (TBS) (0.05 mol/L Tris HCL, 0.15 mol/L NaCl; pH 7.6) 2 times, each for 5 minutes. Endogenous peroxidase activity was blocked by peroxidase block (RE7101, 3-4% hydrogen peroxide) for 5 minutes. Slides were in incubated in protein block (RE7102, 0.4% casein in phosphate-buffered saline, with stabilisers, surfactant, and 0.2% Bronidox L as a preservative) for 5 minutes and primary antibodies, optimally diluted using Leica antibody diluent (AR9352) were added. 300 µl of the primary

antibody in optimal working dilution was added for each slide. Each of these steps was followed by 3×5 minutes washing with 0.1% TBS-Tween 20. Post primary block (RE7111, Rabbit anti mouse IgG (<10µg/mL) in 10% animal serum in TBS/0.09% ProClin™ 950) was incubated for 30 minutes followed by 3×5 minutes washing with TBS-Tween 20 0.1% and then Novolink Polymer (RE7112, Anti-rabbit Poly-HRP-IgG (<25µg/mL) containing 10% animal serum in TBS/0.09% ProClin™ 950) was added for 30 minutes. After washing sections well with TBS, peroxidise reaction was visualised using freshly prepared filtered solution of 3,3'-Diaminobenzidine (DAB) working solution [(DAB chromogen, RE7105, 1.74% w/v 3,3' - diaminobenzidine, in a stabiliser solution) and DAB substrate buffer (RE7143, buffered solution containing ≤0.1% hydrogen peroxide and preservative) of a ratio 1:20 was added for 5 minutes and then followed by 3×5 minutes washing. Counterstaining with Meyer's Haematoxylin (RE7107, <0.1% haematoxylin) for 6 minutes was performed. The slides were dehydrated with alcohol and cleared with xylene (Genta Medica, York, UK). Finally, slides were mounted with distyrene, a plasticiser, and xylene. Tissue sections used as negative as well as positive controls were included during staining of each antibody. A negative control was obtained by omitting the primary antibody and incubating slides in antibody diluent only. Additionally, tissues with known reactivity to each antibody according to the supplier's data sheet were used as a positive control **(Table 4)**

Table 4: Optimum conditions of the selected antibodies

Antibody	Supplier	Species	Incubation time	Dilution	Positive control
PR	Agilent	Mouse Monoclonal	30 minutes	1:150	Normal breast ducts
Ki67	DAKO	Mouse Monoclonal	30 minutes	1:100	Tonsil tissue
Histone H1	Thermo Fisher Scientific, UK	Mouse Monoclonal	30 minutes	1:100	Tonsil tissue
PLK1	Abcam	Rabbit polyclonal	30 minutes	1:250	Testicular tissue
MCM7	Thermo Fisher Scientific, UK	Rabbit Polyclonal	30 minutes	1:500	Colon tissue

2.5.5 Assessment of immunoreactivity

Prior to application of each primary antibody to the TMA using IHC, staining of randomly chosen 14 full-face tumour sections was performed from FFPE BC tissues with each primary antibody to assess homogeneity of immunoreactivity in terms of staining intensity and distribution. When even staining distribution across the tissue was ensured, IHC staining was be applied to TMA sections. In this project, PLK1, histone H1, and MCM7 markers were stained using TMA. However, Ki67 and PR are known to have heterogeneity in distribution of staining positivity [376, 377]. In addition, Ki67 showed low concordance between TMA and full-face sections has been reported [378]. Therefore, Ki67 and PR were assessed using full face sections.

High resolution digital images (0.45µm/pixel) of stained slides were obtained using a the Pannoramic digital slide scanner (250 Flash III, 3DHISTECH®, Budapest, Hungary), at 20x magnification and accessed using Xplore viewing software viewer (Philips) or by direct visualisation of the digital slides using Pannoramic Viewer software. Immunoreactivity was visually assessed at 20x magnification using a minimum of a 24” high resolution screen (1920x1080).

Each TMA core was assessed whether invasive tumour occupied >20% of the total area of the core. If this was the case, the core was eligible for scoring. Assessment of staining was started by assessing the positivity with cutoff >1% then the percentage of protein expression was

evaluated. In addition, Semiquantitative approach; Histo-score (H-score) taking into consideration the intensity (0=negative, 1=weak, 2=moderate and 3=strong staining) and the percentage of stained tumour cells (0-100%) was also calculated. The intensity was multiplied by the percentage and added together to obtain the final H-score with a range from 0-300 [379]. All cases were scored blindly without prior knowledge of patients' pathological or outcome data.

2.5.6 Determination of the cutoff

For dichotomisation of studied biomarkers into low and high expression, different cutoffs were used. Cutoffs were individually assessed and determined by several methods 1) using cutoffs from previous studies that assessed large number of data 2) using the mean or median according to the distribution of data whether normally or not normally distributed respectively or 3) using a non-commercial research tool; X-tile software (version 3.6.1, 2003-2005, Yale University, USA http://x-tile.software.informer.com). This software randomly divides the patient cohort into two separate equal sets; training and validation sets by producing separate lists of 'censored' and 'uncensored' observations, ranked by patients' follow up time [380]. The patients who developed distant metastasis were considered as an event (uncensored), while patients did not develop metastasis or lost follow-up during this study were considered censored during the analysis at the time of that event. The optimal cutoffs were determined by locating the brightest pixel on

the X-tile plot diagram of the training set [381]. Statistical significance was tested by applying the obtained cutoff to the validation set.

2.6 AI- assisted interpretation of histopathological variables and biomarkers

In recent years, DL- based analyses of H&E-stained histology images have produced remarkable results for objective assessment of morphological features [354]. Moreover, AI has the potential to assist pathologists in increasing the reproducibility of biomarker assessment, including Ki67 by developing algorithms to objectively evaluate these markers. In this project, algorithms were developed to assess the histological features and to evaluate Ki67 and PR protein expression.

2.6.1 Algorithm training

The H&E and the IHC stained slides were scanned into high-resolution digital images at 20× magnification using 3D Histech Panoramic 250 Flash II scanner (3DHISTECH Ltd., Budapest, Hungary). The staining and scanning quality of images were important for good annotations and hence better machine learning (ML) models. Therefore, a quality check was performed using AI tool and only passed images were used. The WSIs were uploaded to annotation server (WASABI), a customised version of Histomics TK [382]. The workflow of the annotation was proposed by a group of computer scientists led by Professor Nasir Rajpoot, University of Warwick, UK.

To develop DL models for cell detection and invasive tumour/stroma/DCIS region segmentation, different annotations were marked both at cell-level as well as region-level. The main cell type annotations included tumour, normal epithelial, stromal, and immune cells.

2.6.1.1.1 Regional and cellular annotation of H&E features

The regional annotation comprised of different shapes for each region chosen from the tools of the online channel. Regions of invasive as well as preinvasive regions (DCIS) were annotated. Non-tumour areas including normal breast ducts and stroma were also annotated. Exhaustive annotation of tumour and non-tumour regions was also performed. Fibrofatty and fibroblastic stroma was separately marked based on their morphology **(Figure 2-).** Other regions were annotated such as areas with ink, any artifact and areas of lymphoid collection. Cellular annotation of H&E-stained slides were carried out using smaller boxes. Nuclear features of pleomorphism were annotated, whereas cells with score 1= mild degree pleomorphism, score 2= moderate degree pleomorphism and score 3= high degree pleomorphism. Defining the degree of pleomorphism was performed by comparing the tumour cells with normal epithelial cells. Mitotic cells with clear features were annotated. TILs were annotated within the context of invasive tumours. Normal epithelial cells and stromal cells were also annotated **(Figure 2-).**

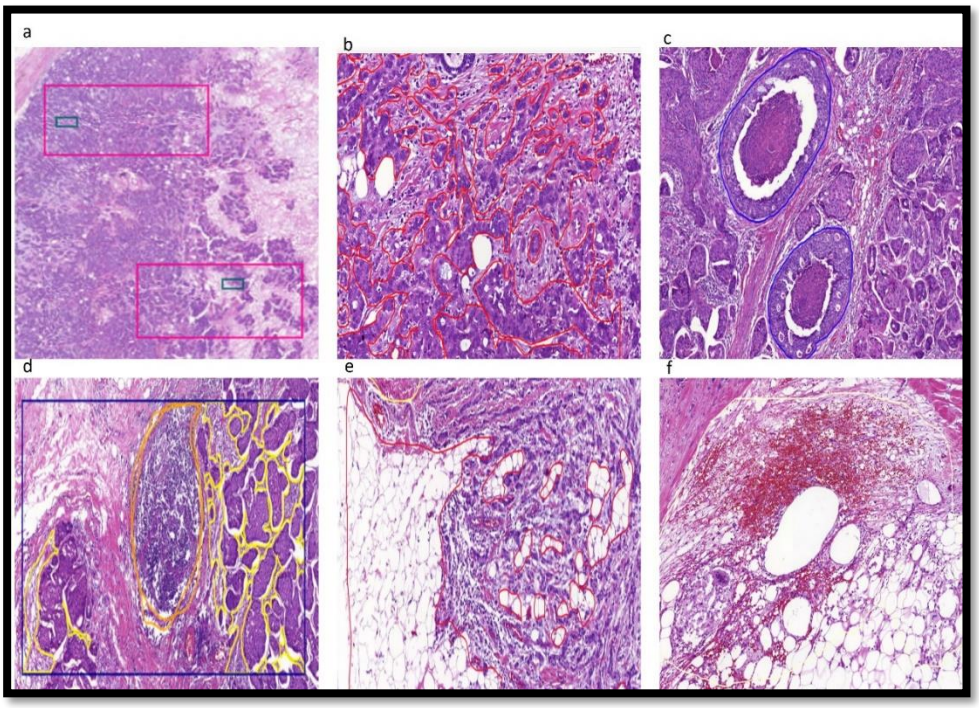


Figure 2-1: Regional annotation of H&E slides

a- annotation boxes, b- annotation of tumour areas, c- DCIS annotation, d- lymphoid collection and fibroblastic stroma surrounding the tumour, e- areas of fibrofatty stroma, f- biopsy site changes**.**

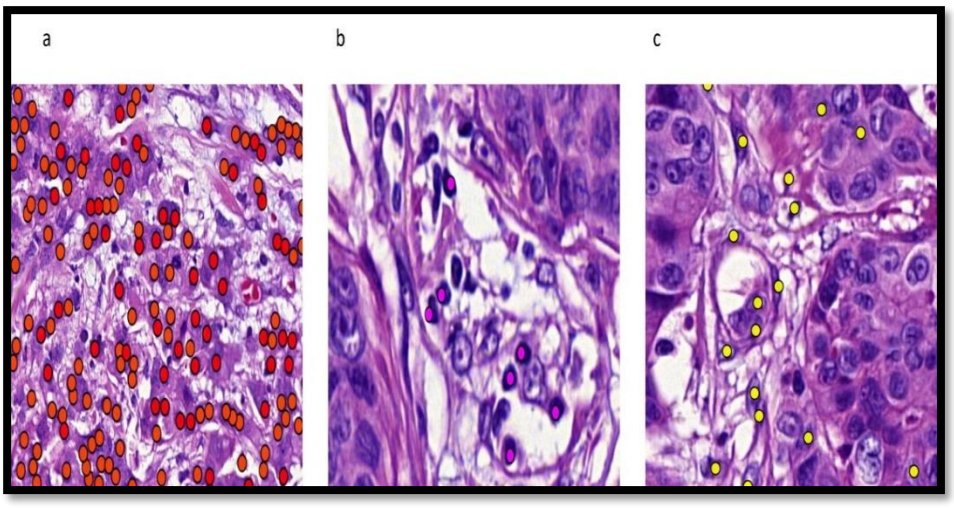


Figure 2-2: Cellular annotation of H&E slides
a- pleomorphism, b- TILs, c-stromal cells.

2.6.1.1.2 Regional and cellular annotation of Ki67 and PR

Regional annotation of Ki67 and PR was performed by marking the areas of tumour positivity and areas of negativity. For Ki67, areas of hotspot, defined as areas with high expression of Ki67 compared to other areas of tumours, were annotated.

For cellular annotation of PR, different intensities of staining expression (weak, moderate, and strong) were included. In addition, annotation of the negative tumour cells for PR expression was also performed. Annotation of Ki67 was performed using WASABI server where 3000 patches of Ki67 were annotated according to 4 types of cells (positive tumour cells, positive non-tumour cells, negative tumour cells and negative non-tumour cells) (**Figure 2- and Figure 2-**).

The annotation was performed by the author and group of pathologists (Dr Michael Toss, Dr Islam Milligy, Dr Ayaka Katayma and Dr Asmaa Ibrahim).

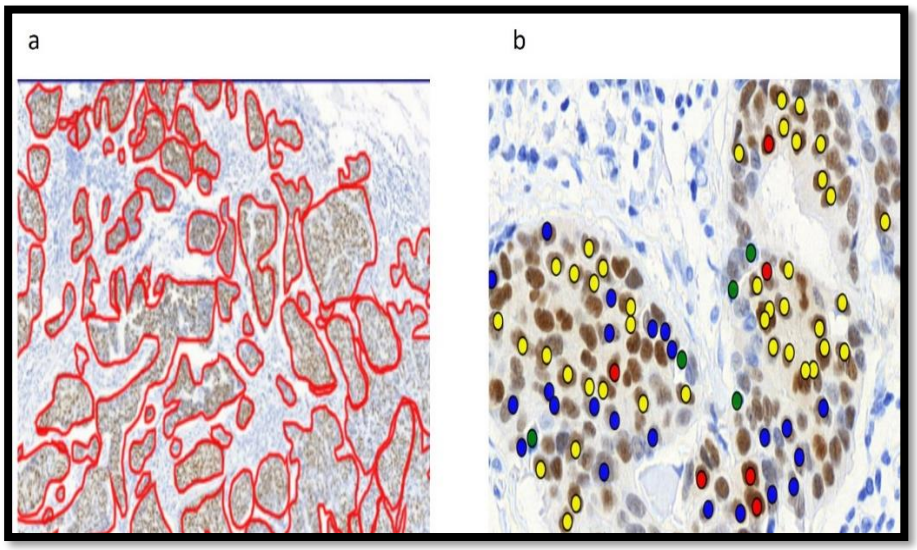


Figure 2-3: Regional and cellular annotation for PR

a- areas of tumour showing PR positivity, b- different intensity of PR expression.

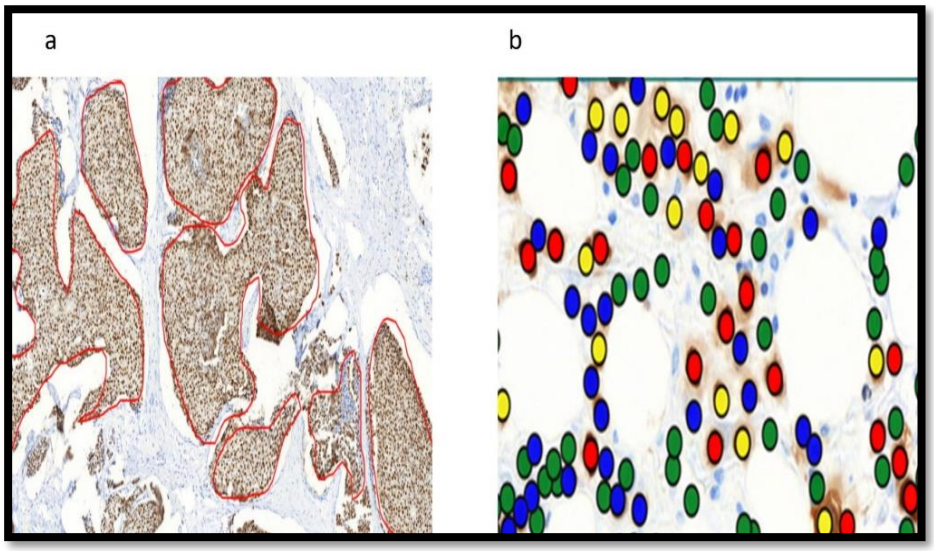


Figure 2-4: Regional and cellular annotation for Ki67

a- hotspot areas of Ki67, b- tumour and non-tumour cells with positive or negative Ki67 expression.

2.6.2 Tissue and tumour segmentation

The DL part of this project was performed by a computer scientist (Wahab Noorul, University of Warwick). Segmentation was required to remove the backgrounds areas which are usually present in the WSI. Using image thresholding and morphological operations, a tissue mask was generated for a WSI to exclude all irrelevant image background or artefacts from further processing [383]. Different morphological operations were performed resulting in the desired tissue area. Araes of DCIS were excluded using Efficient UNet-based model (**Figure 2-**) [384].

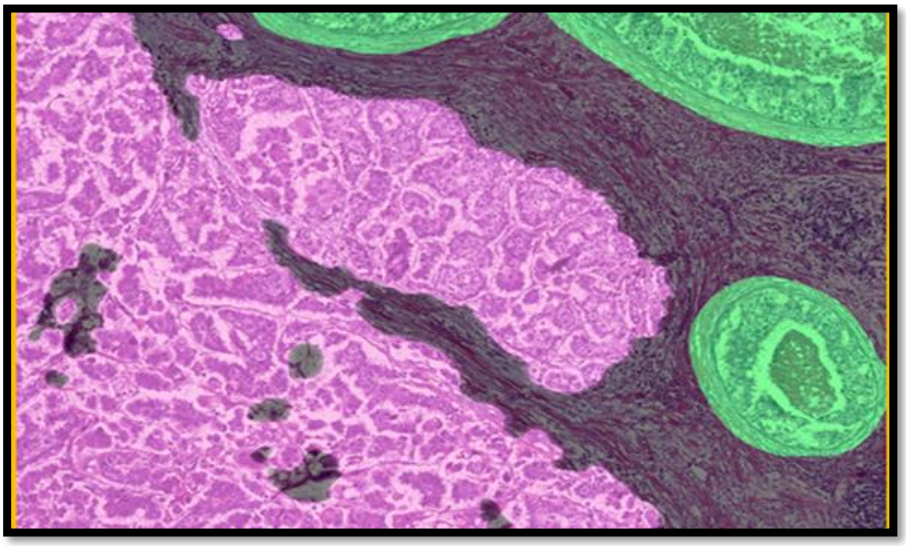


Figure 2-5: Segmentation of invasive tumour and DCIS regions

2.6.3 Cell segmentation and classification

To restrict the training of DL model to tumour areas, which represents the pathologist's assigned WSI-level label, a region of interest (ROI) was selected from each WSI. For this purpose, different types of cells such as tumour and normal epithelial cells were segmented and classified

using a state-of-the-art HoVer-Net, which was pretrained on the BC subset of the PanNuke dataset (an open Pan-Cancer histology dataset for nuclei instance segmentation and classification) **(Figure 2-)** [385, 386]. The trained model was used to generate WSI-level cell-contours to measure tumour-cell density.

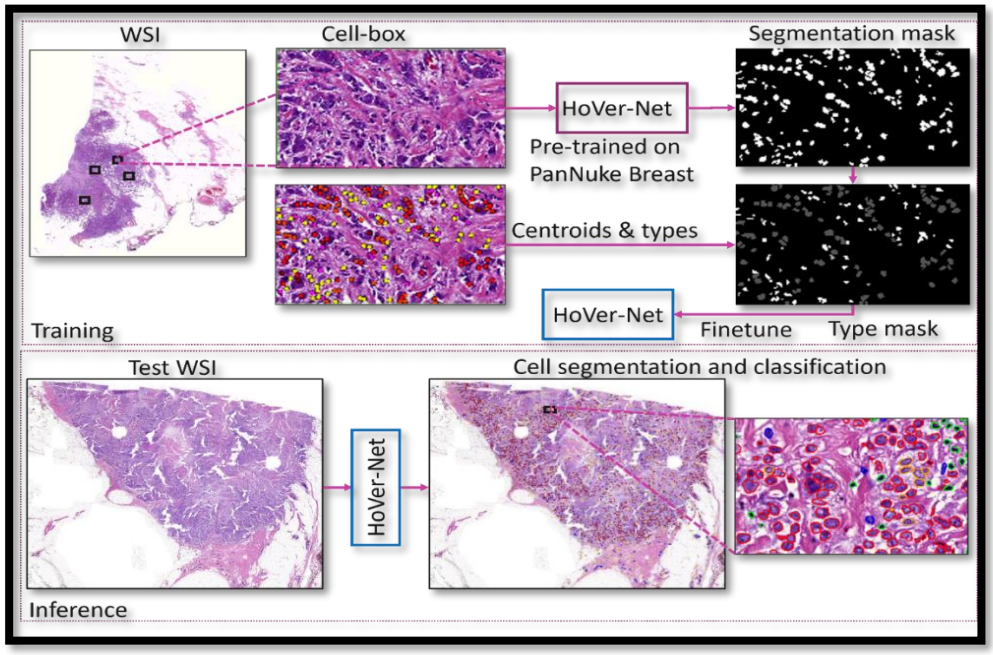


Figure 2-6: HoVer-Net for cell segmentation.
Cell-box is an area where annotations of different types of cells was performed. Segmentation mask of different cells by the pre-trained model.

2.6.4 Prediction of morphological features of BC

Tumour detector was utilised to select an ROI based on tumour nuclei density, size and shape and instead of using patches from all over the WSI, patches from the selected ROIs were used to train the model. Each ROI was of size 5600 × 5600 pixels (about 1344 µm) at 40× magnification. To allow the model to pay more attention to the tumour morphology, patches below a threshold of fifteen tumour cells were

discarded. One ROI per WSI was used to train a convolutional neural network (CNN). The trained model was used to predict features from the entire WSI. To predict tissue region types, a separate CNN was trained on patches from pathologist annotated areas of tumour, stroma, and normal terminal duct lobular units (**Figure 2-**).

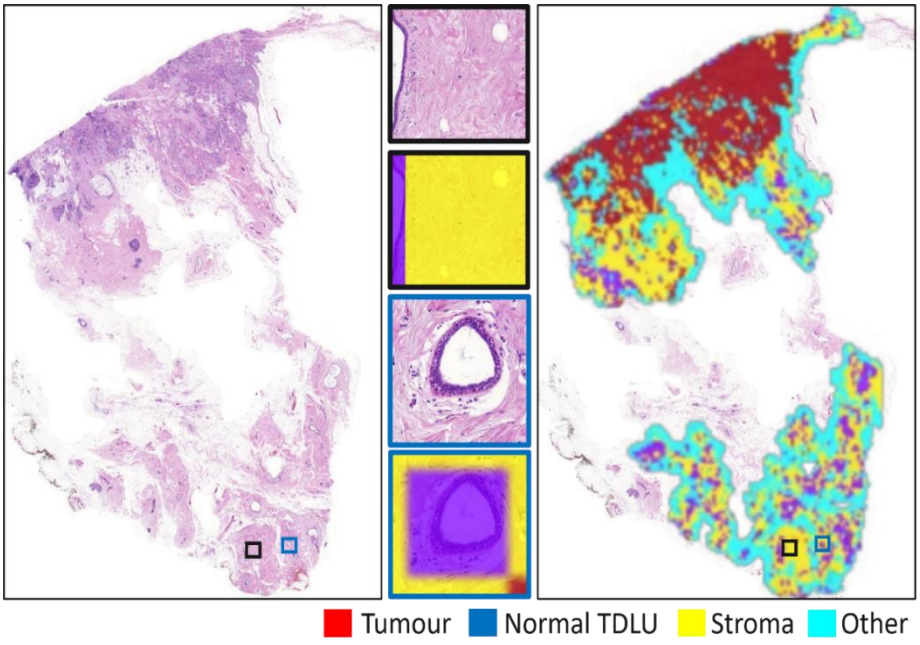


Figure 2-7: Tissue region classification including tumour, normal TDLU and stroma.

2.7 Statistical approach

To test the significance of the results, several statistical approaches were used. Statistical analyses were performed using the Statistical Package for Social Sciences (SPSS) software version 27 (SPSS Inc., Chicago, IL, USA). Chi square test was used to assess the categorical variables. Continuous scale data were assessed using correlation coefficient test. For all statistical tests, the level of significance was set as $p < 0.05$ (two-tailed).

2.7.1 Intra-Class Correlation Coefficient (ICC)

To increase the reliability of the scoring methods, an ICC test was conducted for least 20% of the whole cohort, which was re-evaluated by an additional scorer. ICC correlation was considered satisfactory if the threshold of agreement was higher than 75%. If there were discrepancies between slide scores, the slides were re-scored, and the discrepancies were discussed.

2.7.2 Correlation with clinicopathological parameters

Chi-squared analyses were used for inter-relationships between categorical variables. Regression analysis model was used when comparison with several parameters was indicated. Correlation between continuous variables was performed using Spearman or Pearson tests.

2.7.3 Univariate analysis with outcome

Survival estimates for patient BCSS and DMFS were analysed by Kaplan-Meier curves with significance determined by the log rank test.

2.7.4 Multivariate analysis with outcome

Multivariate analysis was performed using Cox regression proportional hazard analysis to test for confounders and prognostic or predictive independency of the studied biomarker from standard prognostic/predictive factors.

2.7.5 Calculation of Sensitivity, Specificity

The sensitivity and specificity of biomarker was calculated between the expression in core biopsy and in full face sections. Sensitivity was calculated based on the percentage of recorded proven BC tumours positivity in full face sections while specificity was calculated based on the percentage of recorded proven BC tumours interpreted as negative in full face sections.

2.8 Validation of the Findings

Data was sought from online databases for the purposes of comparing and validating the study findings. The prognostic value of the prognostic genes identified was evaluated using the BC Gene Expression Miner v4. 0 (Bc-GenExMiner) database, which includes RNA sequence expression data from 4,713 BC patients and Kaplan Meir plotter cohort, which includes 4929 patients. In this cohort, patient samples are split into two groups according to various quantile expressions of the proposed biomarker. The two patient cohorts are compared by a Kaplan-Meier survival plot, and the hazard ratio (HR) with 95% confidence intervals and log rank P value are calculated. All statistical analyses were performed by the author.

Chapter 3 Eyeballing assessment of mitotic figures in BC

3.1 Introduction

Uncontrolled cellular proliferation is a hallmark of cancer, and it is one of the key predictors of tumour behaviour [387-390]. In BC, the assessment of cell proliferation is a key feature for diagnosing various pathological lesions and predicting patient outcome [152]. Counting mitosis is the most widely used method to estimate cellular proliferation in BC [234, 391, 392]. It is considered a rapid, inexpensive test that can be performed by experienced pathologists with a microscope and glass slide or with a computer and WSI technology [391, 393]. Not only is mitosis used as a part of the grading scheme for assessment of prognosis but also, it can be used alone for clinical decisions in BC patients [138, 394, 395]. Although, several methods have been used to assess the proliferative activity of BC, visual quantification of mitotic figures on H&E-stained sections remains the gold standard method to assess proliferation activity in BC [157, 396]. Identification of mitotic figures in BC relies upon a diversity of morphological features that overlap, and mimic other structures. This causes difficulties in identifying mitotic figures and distinguishing them from other mimickers [1]. These difficulties are demonstrated by the low concordance rates reported among even experienced pathologists and lead to under- or over-estimation of tumour grade and hence affect patients' management [397]. The era of WSI in clinical pathology has undoubtedly many advantages [397]. However, as a newly introduced technology, adaptation of current practice guidelines to WSI is an essential step prior to adopting a fully digitalised system, especially in areas which already show high levels of discordance among

pathologists, such as mitotic counting. Therefore, visual assessment of mitoses in WSI should not follow methodologies derived from LM until comparative studies demonstrate their effectiveness or modifications are recommended to ensure non-inferiority in WSI [393]. In addition, the cutoff for dichotomisation mitotic count into different scores in WSI must be established and validated [398]. The low concordance in mitosis assessment attributed to the variability in pathologists' performance and reflected the subjectivity and variation in methodologies used for the assessment of mitosis [399]. Mitotic count often represents the highest proliferative activity of the tumour, and they are usually obtained by counting mitotic figures within hotspots (areas showing the highest number of mitoses within the tumour) [208, 391].

3.1.1 Hypothesis and aims.

This chapter hypothesised that the low concordance in mitosis assessment is mainly attributable to challenges in consistent identification of true mitotic figures. To test this hypothesis, this chapter aims to (1) define the morphological features of mitoses to enable reliable and reproducible identification of mitotic figures in WSI compared to LM, (2) assess the impact on whether mitotic figures are quantified using digital WSI or using LM using the same area and (3) determine whether new cutoffs are needed when BC is graded using WSI.

3.2 Patient and methods

This study was conducted on a large cohort of well-characterised primary invasive BC specimens. A total of 3089 cases from the Nottingham cohort were included as shown in **Table 5**. This cohort was categorised into sub cohorts for the different aims as follows:

1- For refining the definition of mitosis, 506 cases were used.

2- For refining the mitotic cutoff in WSI, 2583 cases were utilised.

Table 5: Clinicopathological characteristics of the study cohort

Parameter	Number (%)
Patient Age	
≤ 50 years	1056 (34)
> 50 years	2033 (66)
Tumour size	
≤ 2cm	1816 (57)
> 2cm	1273 (43)
Tumour grade	
Grade 1	511 (17)
Grade 2	1174 (38)
Grade 3	1402 (45)
Histologic tumour types	
No special type (NST)	1979 (65)
Lobular	329 (11)
Other special types	169 (5)
NST mixed	547 (19)
Lymph node status	
Negative	1679 (65)
Positive	904 (35)
Lymphovascular invasion	
Negative	2152 (70)
Positive	931 (30)
Estrogen receptor status.	
Positive	2431 (78)
Negative	650 (22)
Progesterone receptor status	
Positive	1803 (70)
Negative	1221 (30)
HER2 receptor status	
Positive	389 (14)
Negative	2657(86)
Molecular subtypes	
Luminal	2302 (80)
Triple negative	433 (15)
HER2 enriched	153 (5)

3.3 Tissue preparation and digitalisation

Four tumour sections per case were histologically reviewed as described in chapter 2.3, and the representative section, defined as the section with highest tumour grade was selected to be included. Freshly cut, 4-µm full-face H&E sections were prepared from the selected FFPE blocks and digitally scanned at 40× magnification using a high-throughput automated scanner (Panoramic 250 Flash III: 3D-Histech, Budapest, Hungary). Digital images were viewed using Case Viewer software (version 2.2.0.85; 3D-Histech) on a full-screen panel (21-inch screen with 1366 × 768 pixels resolution) at 40x magnification. For comparative analysis, the same section for each case was examined simultaneously using LM (Eclipse Ni-U; Nikon Instruments Inc, Tokyo, Japan) with a × 400 and field diameter of 0.63 mm.

3.4 Refining the definition of mitosis in WSI

3.4.1 Comparison of mitotic figures between LM and WSI

Features of mitosis and mimickers, as prescribed in the literature, were applied to identify mitotic figures on the glass slides in a set of cases (n=20) [149, 400-402]. Mitotic figures were identified under LM at 40x magnification. Each figure identified using LM was compared with its corresponding figure in WSI. The similarities and differences between mitotic figures in WSIs and LM were assessed and recorded. Based on the difference in mitotic figures between LM and WSIs, eight morphological features were identified to define mitotic figures in WSI, and to distinguish true mitoses from their mimickers: These features are:

1. Presence of nuclear membrane 2. Nuclear outline (regular / irregular), 3. Presence of hairy-like projection, 4. Shape of the chromatin mass (symmetrical / asymmetrical), 5. Chromatin homogeneity (homogeneous / heterogeneous), 6. Cytoplasmic features (slightly eosinophilic / deeply eosinophilic / pale / basophilic / rim or not seen/), 7. Relative cell size compared to nearby neoplastic cells (similar / enlarged / increased), 8. Relationship to surrounding cells (no specific relation / compressed between two cells / stretched in the periphery of sheets or nets of cells). In this study, irregular outline was defined as a ragged border, and this is different from the hairy-like projections which are fine nuclear extensions visible on LM. Relationship to surrounding cells was assessed separately (**Figure 3-1)**.

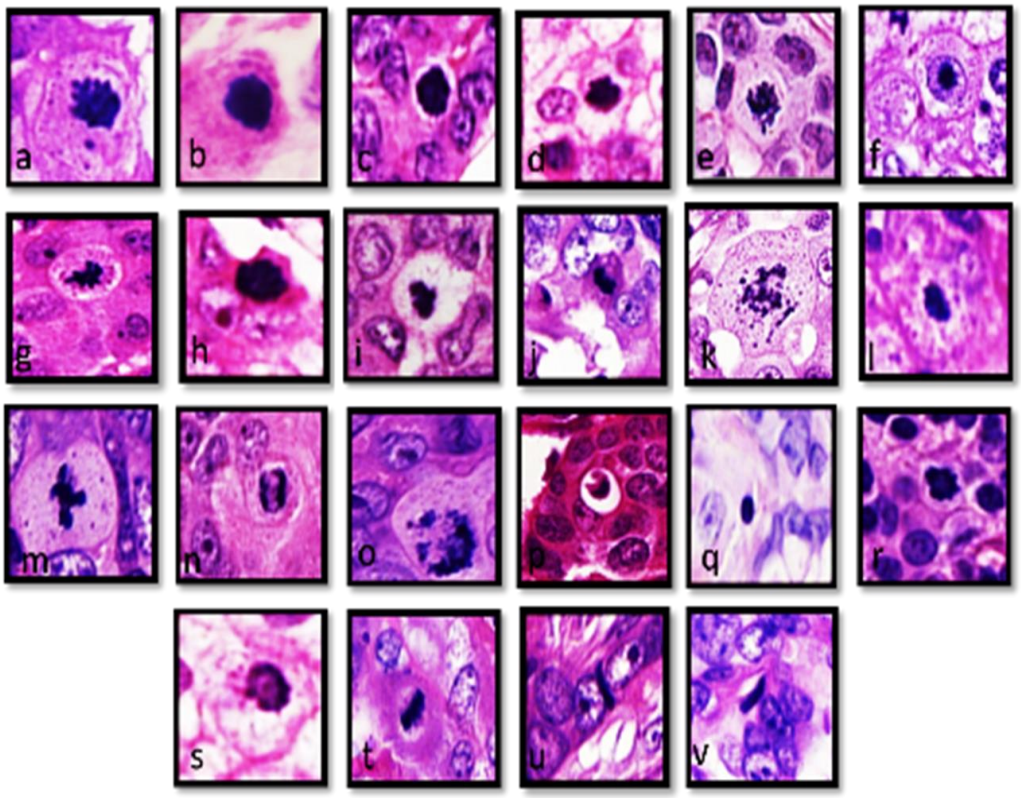


Figure 3-1 : Recorded criteria to characterise mitoses in WSI.
(a) Mitosis with obvious hairy-like projection while b) mitosis without evident hairy-like projection; (c) prophase shows mildly regular nuclear outline compared to (d) another prophase with an irregular outline; (e) prophase with absent nuclear membrane; (f) malignant cells with dark hyperchromatic nucleolus mimicking mitosis but nuclear membrane is smooth; (g) mitosis with slightly eosinophilic cytoplasm while (h) shows deep eosinophilic cytoplasm, (i) mitosis with clear cytoplasm, (j) slightly basophilic cytoplasm in mitotic cells, (k) cytoplasmic granularity ; (l) chromatin symmetry in typical mitosis compared to (m) which shows atypical mitosis with asymmetrical chromatin assembly; (n) anaphase mitosis with similar size to the surrounding cells, (o) atypical mitosis with relatively large cell compared to other surrounding cells; (p) apoptosis with small cell size compared to surrounding cells and surrounded by clear space; (q) lymphocyte with dark small nucleus, without cytoplasm and regular border, and small size ; (r) prophase with homogenous chromatin distribution, in contrast to (s) which shows heterogeneity in chromatin distribution and staining quality, (t) metaphase with obvious cytoplasm and no relation to surrounding cells while (u) and (v) show dark hyperchromatic structures mimicking mitosis in metaphase stage.

3.4.2 Identification of mitosis and mimickers with WSI

To identify the morphological features that are most likely to distinguish mitoses from mitotic mimickers on WSI, a second set of n = 20 cases was assessed. Mitotic figures and mimickers (apoptotic bodies / elongated hyperchromatic structure / lymphocytes / indeterminate figures) were identified, and the same morphological features as defined above were recorded for each figure. The necessity to distinguish elongated hyperchromatic structures from mitosis stemmed from close similarity to metaphase, which can be false identified as mitosis, especially with identification of mitotic figures on WSI.

To validate the morphological features recorded and to investigate the rate of subjective interpretation when assessing them, 300 patches of mitotic figures and mimickers were imported at 40x from WSIs and circulated to three experienced breast pathologists (the author, Dr Ayaka Katayama and Dr Michael Toss). They were asked to independently identify and record the presence or absence of the specified morphological features for each of the 300 mitoses. The concordance for identifying each criterion was calculated. Also, one pathologist (Dr Karim El dib) was asked to quantify all the mitoses in 20 WSI in the whole slide based on their experience, then was given these features to identify mitosis and has been asked to re-evaluate the same cases. The degree of improvement in recording mitosis was assessed.

3.4.3 Comparison of mitotic counts between LM and WSI

To evaluate the reduction rate between the mitotic counts in LM and WSIs if any, the mitotic count was evaluated in hotspots in a set of cases (n=341) in an area of 3mm^2, which is approximately equivalent to 10 HPFs of the wide field microscope used in this study. Hotspot areas were selected based on being the most cellular area and was usually located at the peripheral invasive front of the tumour. Initial examination of WSI was carried out at 10× magnification to identify the cellular regions harbouring the highest mitotic counts which were quantified at 40× magnification. The same area was assessed for mitoses using both LM and WSI. Next the College of American Pathologists criteria were used to assign absolute mitotic counts determined into categories as follows: Score 1:0-11 mitoses per 3 mm^2, score 2:12-22 mitoses in 3mm^2, and score 3:> 22 mitoses per 3 mm^2 [403, 404]. Area of 3mm^2 in hot spot was used to count mitosis as it showed best area size to represent mitotic count as described in detail in a previous study [235]. The difference in the number of identified mitoses quantified using LM and WSI in entire slides was evaluated to assess the prevalence of mitotic phases in BC.

3.4.4 Evaluation of the prevalence of mitotic phases

To further refine the features of mitosis and to identify the magnitude of the challenges related to the different shapes of mitotic figures with corresponding mimickers, the phases and pattern of mitotic figures in a randomly set of 125 BC cases were reviewed. Mitotic phases (prophase, metaphase, anaphase, telophase and atypical) as well as the pattern of

mitotic figures (typical or atypical) were assessed in the entire digitalised images. In addition, these cases were also assessed by LM.

3.4.5 Refining mitotic cutoffs when using WSI

The reduction rate of mitosis in WSI compared to LM was used beside patient outcome to refine the existing cutoff of mitosis in BC to be appropriate for counting mitosis using WSI. Firstly, the mitotic count scored with LM was plotted against BCSS using X tile software, then the mitotic count with WSI, using the same area size, was analysed against BCCS to evaluate the best mitotic cutoff that can be used when counting mitoses using WSI. The correlation between both LM and digital mitotic scores using the refined cutoff was next assessed. To validate the refined cutoffs, a large cohort of WSI (n=2583) was used. Mitosis was counted in this cohort using the refined criteria over an area of $3mm^2$ as described before in detail [235] and the refined cutoff was applied. For comparative analysis, the existing cutoff for the same area was applied to both LM mitotic count and digital mitotic count.

3.5 Statistical analysis

Statistical software 27.0 (SPSS, Chicago, IL, USA) was used for statistical analysis. Chi-square was used for analysis of categorical data. To refine cutoff for mitotic count in WSI, X-tile bioinformatics version 3.6.1 (Yale University, USA) was utilised with BCSS as an endpoint. The degree of inter-observer agreement in detection of mitosis within the selected region of interest was assessed using ICC and the kappa test for continuous and categorical data, respectively; 0 is interpreted as poor,

0–0.2 as slight, 0.2–0.4 fair, 0.4–0.6 moderate, 0.6–0.8 good and > 0.8 excellent. The difference rate between mitotic counts in WSIs and LM was assessed using the correlation coefficient test. For all tests, $P < 0.05$ (two-tailed) was considered statistically significant. The Kaplan–Meier survival method and the log-rank test were used for survival curves. For all tests, $P < 0.05$ (two-tailed) were considered statistically significant.

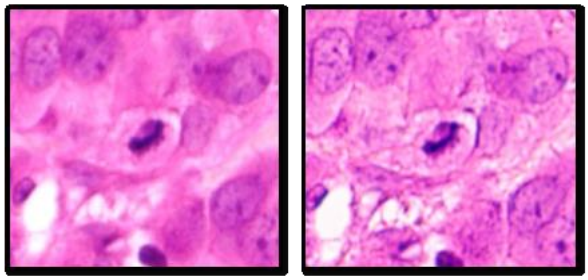
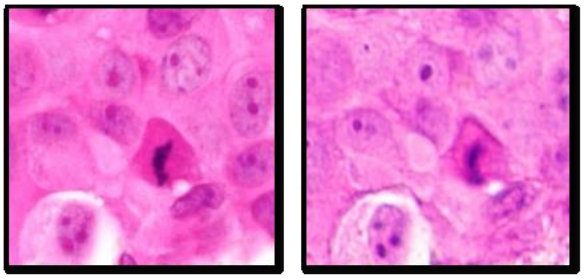
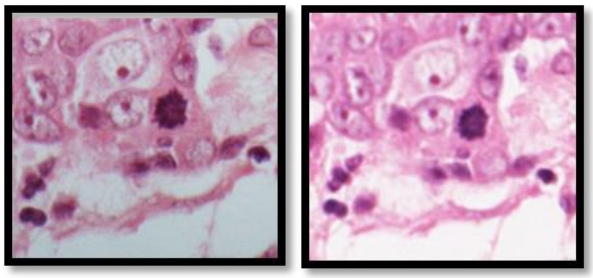
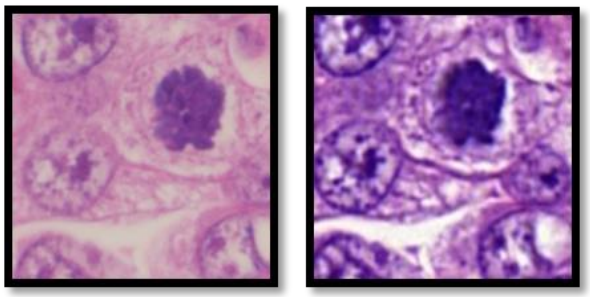
3.6 Results

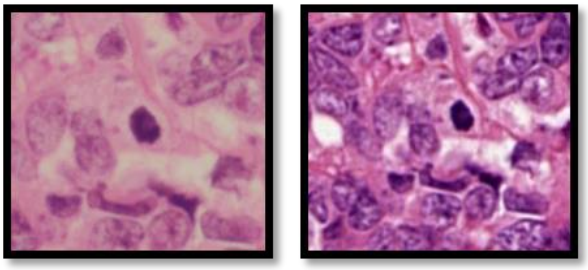
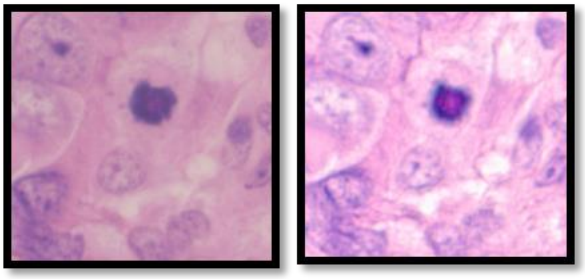
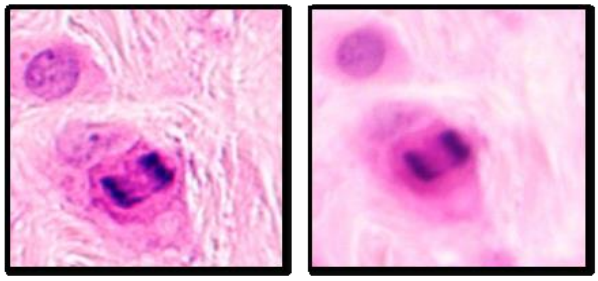
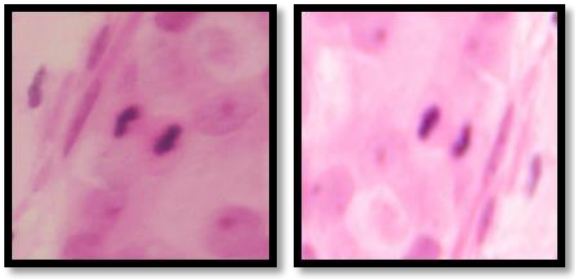
3.6.1 Comparison of mitotic figures between LM and WSI

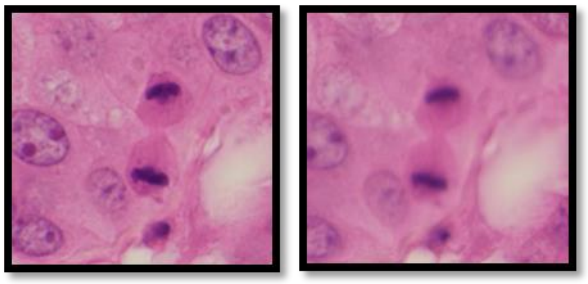
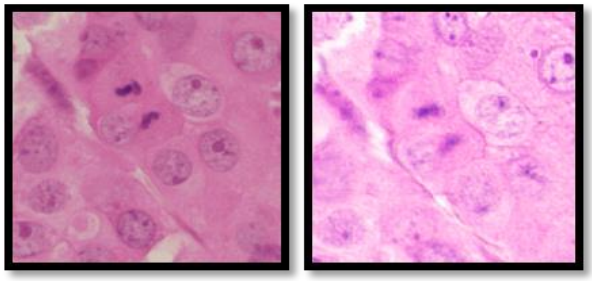
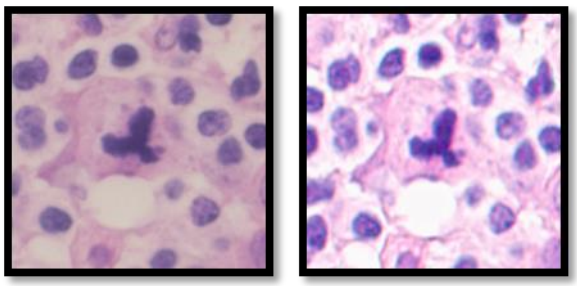
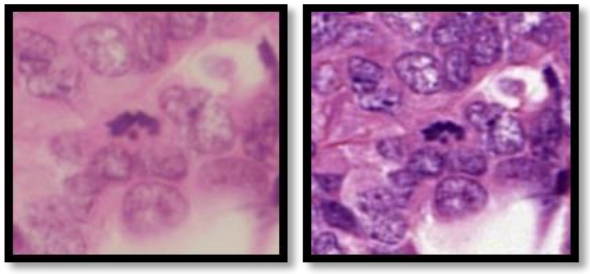
A total of 2000 mitoses were identified by LM using existing criteria for mitosis. To observe the fine details of mitotic figures, the fine-tuning facility of microscope was used. 336 mitotic mimickers were also identified at ×400 magnification. These included apoptotic bodies, elongated hyperchromatic structures, lymphocytes and indeterminate figures. The features of these mimickers were also used to refine the morphological definition of true mitoses in WSIs. Mitotic figures in WSIs appeared darker and more basophilic compared to their corresponding in LM and importantly, the fine details of mitoses were less evident in WSIs when compared to LM (**Table 6-Table 7**).

Table 6: Matched mitotic figures between LM and WSI

Microscopic	Digital	Description and comment

	**Description: Metaphase** Shape: curved rod-like perpendicular to the long axis. Nuclear membrane: absent Hairy projections: present Cytoplasm: eosinophilic **Comments:** both figures showed the same criteria
	**Description: Metaphase** Shape: linear and perpendicular to the long axis Nuclear membrane: absent Hairy projections: present Cytoplasm: deeply eosinophilic **Comments:** not recognised in WSI, due to lack of focus.
	**Description: Prophase** Shape: rounded Nuclear membrane: absent Hairy projections: present Cytoplasm: eosinophilic **Comments:** both figures showed the same criteria
	**Description: Prophase** Shape: rounded Nuclear membrane: absent Hairy projections: present Cytoplasm: slightly basophilic **Comments:** both figures showed the same criteria

	**Description: Prophase** Shape: rounded Nuclear membrane: absent Hairy projections: present Cytoplasm: eosinophilic **Comments:** in WSI, irregular outline replaces hairy outline.
	**Description: Prophase** Shape: rounded Nuclear membrane: absent Hairy projections: present Cytoplasm: eosinophilic **Comments:** not recognised in WSI, due to lack of focus.
	**Description: Anaphase** Shape: 2 uniform groups of dark chromatins are being pulled away from each other. Nuclear membrane: absent Hairy projections: present Cytoplasm: deeply eosinophilic **Comments:** although, it is slightly out of focus in WSI, it is still recognised.
	**Description: Anaphase** Shape: 2 uniform groups of dark chromatins are being pulled away from each other. Nuclear membrane: absent Hairy projections: absent Cytoplasm: eosinophilic **Comments:** although, it is slightly out of focus in WSI, it is still recognised.

	**Description: Telophase** Shape: two uniformly sized separate clusters of chromosomes. Nuclear membrane: absent Hairy projections: absent Cytoplasm: eosinophilic **Comments:** Both figures showed the same criteria
	**Description: Telophase** Shape: Two uniformly sized separate clusters of chromosomes. Nuclear membrane: absent Hairy projections: absent Cytoplasm: eosinophilic **Comments:** although, it is slightly out of focus in WSI, it is still recognised
	**Description**: Tripolar atypical mitosis Shape: 3 spindle poles with irregular outline Hairy projections: present Cytoplasm: eosinophilic **Comments:** Both figures showed the same criteria
	**Description:** Lag atypical mitosis Shape: one chromosome is on one side of metaphyseal plate. Nuclear membrane: Absent Hairy projections: present Cytoplasm: eosinophilic

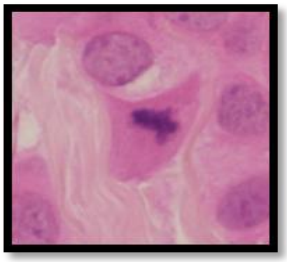
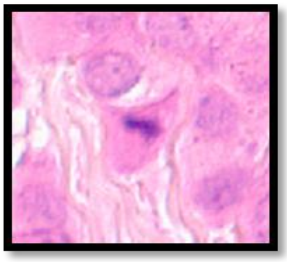
	Comments: Both figures showed the same criteria
 	**Description**: Tripolar atypical mitosis Shape: 3 spindle poles with irregular outline Hairy projections: absent Cytoplasm: eosinophilic **Comments:** not recognised in WSI, due to lack of focus

Table 7: Features of Mitoses in LM versus WSI

Mitoses features	LM	WSI
Hairy-like projection	Present in almost all mitotic figures	Absent in about 60% of mitotic figures
Irregular (ragged border)	Not recognised well	Appreciated in almost all mitotic figures
Fine details of mitosis	Easily appreciated due to fine tuning ability of LM	Difficult to be recognised due to the one focal plane of WSI
Cytoplasmic features	Mostly, slightly eosinophilic	Appeared darker
Nuclear membrane	Absent	Absent
Homogeneity of chromatin	Most of the mitotic figures have homogenous chromatin	Similar

3.6.2 True Mitotic figures versus mimickers

58% of mitoses showed absent hairy-like projection in WSI while none of the mimickers showed hairy-like projections. An irregular, ragged border nuclear outline was a valuable distinguishing feature to differentiate mitoses from their mimickers as 89% of mitoses showed this feature, while most mimickers showed regular nuclear outlines. Although nuclear membrane was absent in all mitoses, this was also difficult to observe in all mitotic mimickers including hyperchromatic dense nuclei and apoptotic bodies. However, the nuclear membrane was easily identified in all lymphocytes. Although mitotic cells showed granular or clumped chromatin, the staining pattern was homogeneous in most cases (91%). Almost all mitoses (99%) showed abundant cytoplasm while most of the elongated hyperchromatic structures (80%) and all lymphocytes showed absent or only a thin rim of cytoplasm. Deep eosinophilic cytoplasm was seen in 84% of apoptotic bodies, while only 8% of mitoses showed deep eosinophilic cytoplasm; 81% of mitoses showed slightly eosinophilic cytoplasm, whereas 7% showed clear cytoplasm. In addition, cytoplasmic granularity was observed in 31% of mitoses and this was completely absent in mimickers. Cell size was a distinguishing feature between mitotic and apoptotic cells with 67% of mitoses being relatively similar in size to nearby neoplastic cells, whereas 70% of apoptotic cells were of smaller size. Lymphocytes were clearly smaller than neoplastic cells. However, 92% of elongated hyperchromatic structures were relatively similar in size to nearby neoplastic cells and mitotic cells. Fragmentation and cytoplasmic retraction (represented by clear space

around apoptotic cells) were unique features of apoptosis, with 25% and 30% of apoptotic bodies showed fragmentation and cytoplasmic retraction, respectively, while no mitoses presented these features. 87% of elongated hyperchromatic structures showed a specific relation to surrounding neoplastic cells (either compressed between two cells or stretched at the periphery of cell aggregates and sheets) (**Table 8).** Mitoses and other mimickers did not show such specific relationships with nearby cells. **Figure 3-2, Figure 3-3, and Figure 3-4** show how to differentiate between mitoses and different mimickers in WSI.

Table 8: Morphological features of mitotic figures versus mimickers in WSI.

Recorded criteria	Mitosis (n=964, 74%)	Apoptosis (n=121, 9%)	Hyperchromatic nuclei (n=178, 14%)	Lymphocytes (n=17, 1%)	Indeterminate Structures (n=20, 2%)	P value
Nuclear membrane						
Absent	964 (100)	121 (100)	0(0)	0 (0)	0 (0)	**<0.0001**
Present	0 (0)	0 (0)	4 (2)	17 (100)	0 (0)	
Inconclusive	0(0)	0 (0)	174 (98)	0(0)	20 (100)	
Hairy-like projection						
Absent	555 (58)	121 (100)	178 (100)	17 (100)	19 (95)	**<0.0001**
Present	409 (42)	0 (0)	0 (0)	0 (0)	1 (5)	
Outline or border of the hyperchromatic nuclei						**<0.0001**
Regular	109 (11)	113 (93)	171 (96)	17 (100)	6 (30)	
Irregular	855 (89)	8 (7)	7 (4)	0 (0)	14 (70)	
Nuclear symmetry						**<0.0001**
Symmetrical	738 (77)	91 (76)	130 (73)	17 (100)	8 (40)	
Asymmetrical	226 (23)	30 (24)	48 (27)	0 (0)	12(60)	
Homogeneity of chromatin						
Homogenous	877 (91)	77 (64)	173 (97)	15 (88)	17 (85)	**<0.0001**
Heterogeneous	87 (9)	44 (36)	5 (3)	2 (12)	3 (15)	
Cytoplasmic features						
Slightly eosinophilic	784 (81)	11 (9)	27 (15)	0 (0)	15 (75)	**<0.0001**
Deeply eosinophilic	81 (8)	101 (84)	5 (3)	0 (0)	3 (15)	
Clear	64 (7)	0 (0)	2 (1)	0 (0)	0 (0)	
Basophilic	22 (3)	0 (0)	1 (1)	0 (0)	0 (0)	
Rim or not seen	13 (1)	9 (7)	143 (80)	17(100)	2 (10)	
Relative cell size compared to nearby cells.						**<0.0001**
Similar	734 (67)	33 (27)	163 (92)	6 (35)	18 (90)	
Increased	210 (21)	3 (3)	13 (7)	0 (0)	2 (10)	

Decreased	20 (2)	85 (70)	2 (1)	11 (65)	0 (0)	
Relation to surrounding cells No specific relation Stretched at the periphery. Compressed between 2 cells	 964 (100) 0 (0) 0 (0)	 121(100) 0 (0) 0 (0)	 23 (13) 95 (53) 60 (34)	 17 (100) 0 (0) 0 (0)	 20 (100) 0 (0) 0 (0)	**<0.0001**

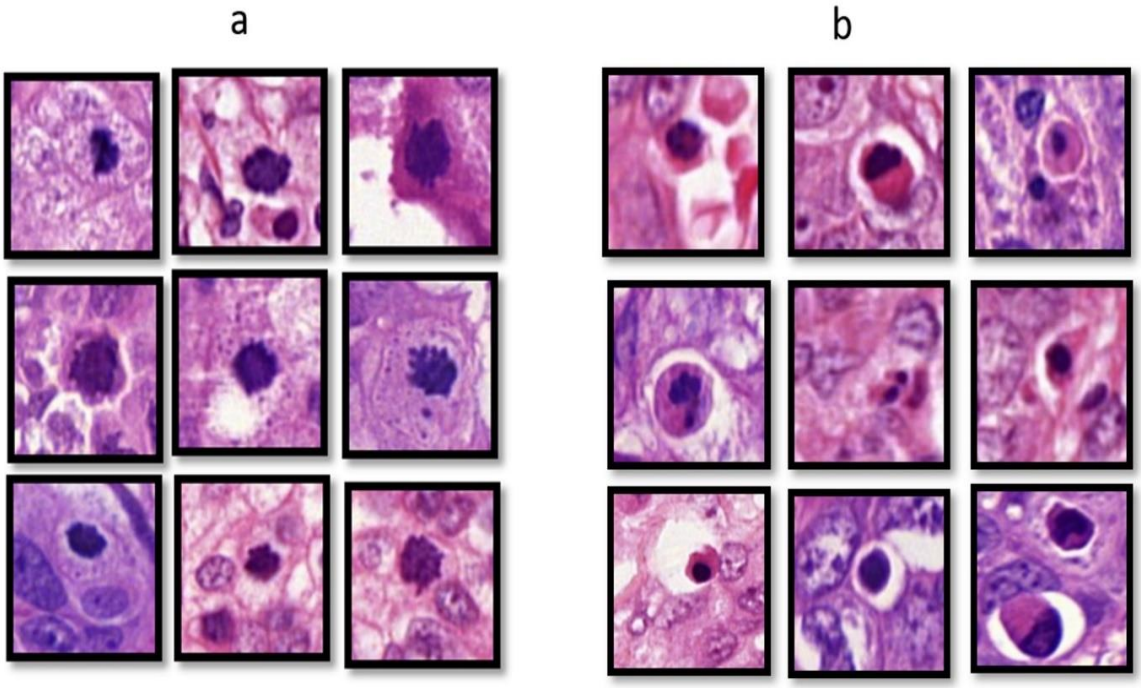


Figure 3-2: Differentiating between prophase and apoptosis in WSI at (40x)
a- different shapes of prophase show hairy-like projections /ragged borders, slightly eosinophilic cytoplasm, and relatively similar size compared to nearby neoplastic cells, b- different shapes of apoptosis show deeply eosinophilic cytoplasm, regular borders, relatively smaller size compared to nearby neoplastic cells, cytoplasmic retraction, and fragmentation of nuclei in some figures.

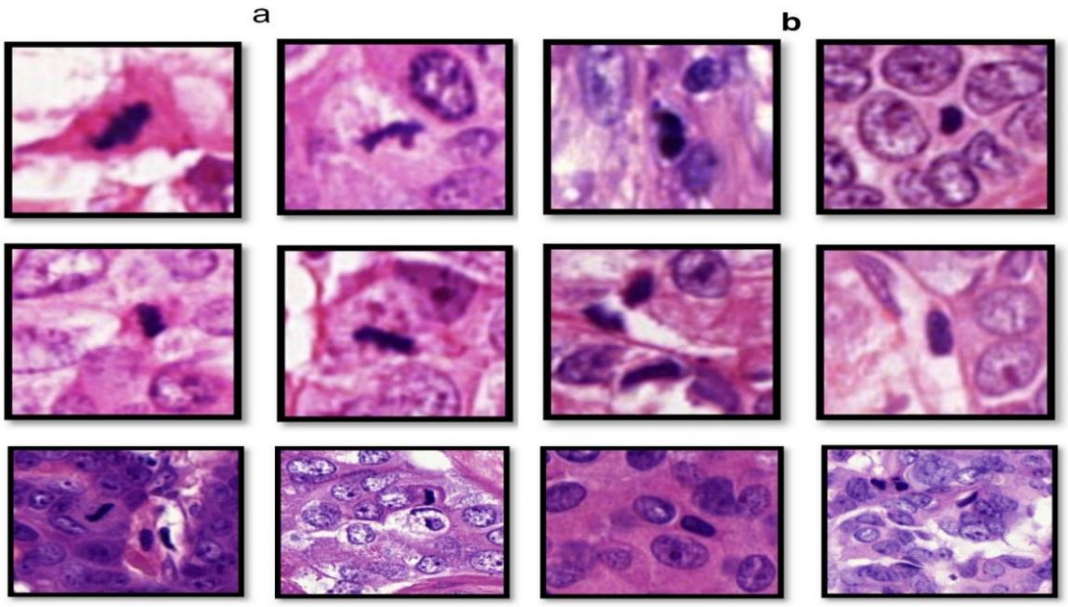


Figure 3-3: Differentiating between metaphase and elongated hyperchromatic nuclei in WSI at 40x
a- different shapes of metaphase show hairy-like projections/irregular outline, abundant cytoplasm, and no specific relation to surrounding cells, b- different shapes of elongated dark hyperchromatic nuclei show no visible cytoplasm, regular outline and have specific relations with surrounding cells (stretched at the periphery of cell nests or compressed between 2 cells).

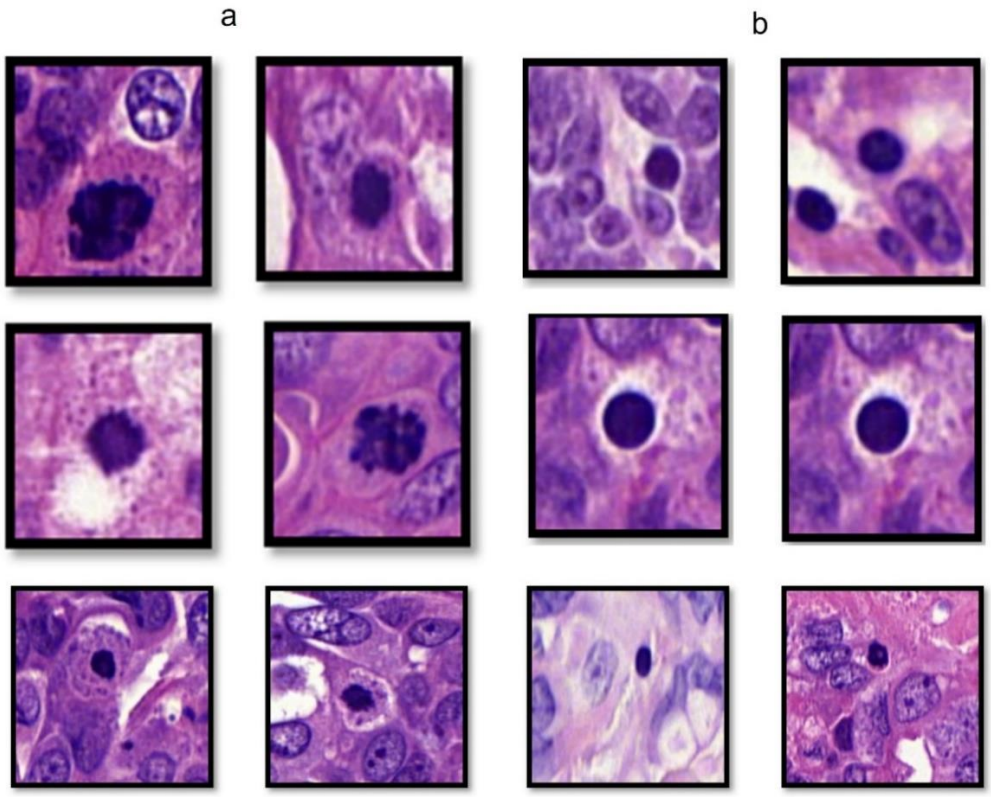


Figure 3-4: Differentiating between prophase and lymphocytes in WSI (40x) a- different shapes of prophase show irregular outline/hairy-like projections and abundant cytoplasm, b- different shapes of lymphocytes show smooth regular outline, no visible cytoplasm and are relatively smaller in size compared to nearby neoplastic cell.

3.6.3 Refined criteria for mitosis identification in WSI

A set of refined criteria was established to identify mitoses based on the difference between the morphology of mitotic figures in LM versus WSI (**Table 9**). Most of the mitotic figures show the presence of abundant eosinophilic cytoplasm with irregular nuclear outline. However, the presence of deep eosinophilic cytoplasm by itself could not reliably exclude mitosis, as it was observed in 8% of the identified mitotic figures. Hairy-like projection was not obvious in WSI, but it was replaced by irregular border. Most of the typical mitotic figures were symmetrical in shape while almost all atypical mitotic figures show asymmetry. All mitotic figures have no specific relationship to the surrounding tissues.

Table 9: Comparison between existing and refined criteria for mitosis identification

Existing criteria	Refined criteria
1- Absent nuclear membrane 2- Hairy-like extensions of nuclear material 3- Shape of chromatin mass a- Typical b- Atypical	1- Absent nuclear membrane 2- Hairy-like projection may not be visible in WSI. 3- Irregular nuclear outline (ragged border) 4- Shape of chromatin mass a- Typical (symmetrical) b- Atypical (mostly, asymmetrical) 5- Chromatin material is typically Homogeneous. 6- Cytoplasmic features: cytoplasm is commonly to be slightly eosinophilic and less frequently basophilic or clear. Deeply eosinophilic cytoplasm does not exclude mitosis though it is most frequently seen in apoptosis. Cytoplasmic granularity is a distinctive feature of mitosis. 7- Absence of specific relation to nearby neoplastic cells (not compressed between or stretched at the periphery of cell nests/sheets) 8- Similar size to surrounding malignant cells.

3.6.4 Validating the refined criteria for mitoses in WSI

To validate the refined criteria for mitotic figures identification in WSI, the concordance rate in identifying the various morphological features of mitosis in WSI was assessed and showed the highest agreement in recognising the following features: absence of nuclear membrane, ragged nuclear border, slightly eosinophilic cytoplasm and similarly sized to nearby neoplastic cells (**Table 10**).The interobserver concordance in identifying and counting true mitoses in 20 selected ROIs showed a kappa of 0.8 when using the refined criteria compared to kappa of 0.4 when using the existing criteria. There was much improvement in recognition of mitoses compared to applying the existing criteria in the same cases. Inter-observer concordance in the different mitotic phases was almost perfect concordance between 2 pathologists (the author and Dr Asmaa Ibrahim) after applying the refined criteria of mitosis especially in metaphase (**Table 11**).

Table 10: Concordance of recognising the recorded morphological features of mitoses

Recorded criteria	Inter-observer concordance			
	First observer (%)	Second observer (%)	Third observer (%)	Concordance rate
Nuclear membrane				
Absent	100	100	100	1.0
Present	0	0	0	1.0
Hairy-like projection				
Absent	58	70	60	0.6
Present	42	30	40	0.4
Nuclear outline				
Regular	10	2	1	0.4
Irregular	90	98	99	0.9
Nuclear symmetry				
Symmetrical	80	82	90	0.8
Asymmetrical	20	18	10	0.2
Homogeneity of chromatin				
Homogeneous	93	95	95	0.9
Heterogeneous	7	5	5	0.6
Cytoplasmic features				
Slightly eosinophilic	82	85	75	0.8
Deeply eosinophilic	8	9	10	0.9
Clear	7	6	12	0.8
Basophilic	2	0	3	0.3
Rim or not seen	1	0	0	0.3
Relative cell size compared to nearby neoplastic cells				
Similar	77	86	70	0.7
Increased	20	14	28	0.3
Decreased	3	0	2	0.2
Relation to surrounding cells				
No specific relation	100	100	100	1.0
Specific relation	0	0	0	1.0

Table 11: Interobserver concordance in mitotic phases identification

Mitotic phases	Intraclass correlation coefficient (ICC)
Prophase	0.95
Metaphase	0.98
Anaphase	0.97
atypical	0.80
Total	0.98

3.6.5 Comparing between mitotic counts in LM and in WSIs

As some mitotic features were not identified in WSI, the impact of losing some mitoses in WSIs compared to LM was assessed in 341 cases. A total of 5807 mitoses were examined.

The sensitivity of mitosis detection in WSIs was 87%, while the specificity 94%. There was a strong linear positive correlation between mitotic count in LM and WSIs ($r=0.95$, $p<0.0001$). However, there was a systematic under-estimation of mitoses in WSIs compared to LM by approximately 20% of all mitotic figures ($y=0.8$) **(Figure 3-5).** The mean mitotic count in LM, was 17 mitoses/ 3 mm2 (range 0-130), compared to a mean count of 13 mitoses / 3 mm2 (range 0 – 102) when the same areas were assessed, case by case, in WSI.

Using categorical scores for mitotic counts, WSI assessment showed a tendency to underestimate mitotic scores when compared to LM. 37% of cases with score 1 in WSI were assigned a mitotic score 2 on LM, while 27% of cases score 2 in WSI were assigned as score 3 with LM **(**

Table 12). Overall, a total number of 64 (18.8%) cases were underscored in WSIs when compared to LM. None of the WSI-based mitotic scores was higher than their corresponding scores in LM. 14% of mitoses was not recognised with confidence in WSIs due to some degree of haziness and loss of image focus.

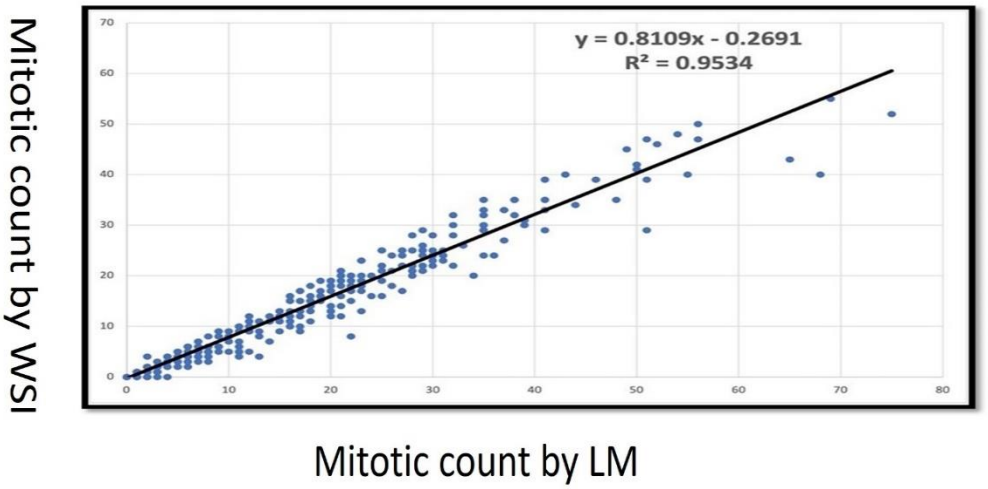


Figure 3-5: Scatter plot shows the correlation between mitotic counts with LM versus WSI.

Table 12: Comparison between mitotic scores on LM versus WSI using the existing mitosis cutoff

Mitotic scores on LM	**Mitotic scores using WSI**			**X^2 P value**
	Score 1 (n, %)	**Score 2 (n, %)**	**Score 3 (n, %)**	311.1 **<0.0001**
Score 1	144 (100)	0 (0)	0 (0)	
Score 2	37 (37)	63 (63)	0 (0)	
Score 3	0 (0)	27 (27)	70 (73)	

3.6.6 The prevalence of mitotic phases in BC

To further investigate the reasons contributing to the underestimation of mitotic counts in WSI and the impact of different mitotic shapes on the ability to recognise and distinguish them from mimickers, a total of 9068 mitoses were examined from the whole slide (n=125). These figures were used to assess the frequency of mitotic phases and shapes in BC. It was shown that 76% of the observed mitoses were typical mitoses, while 24% were atypical forms. The most common phase of mitosis was metaphase (48%), followed by prophase (36%). Anaphase and telophase (2%) were the least common phases, but they were the easiest phases to identify as true mitoses in WSIs. **Table 13** and **Table 14** show the prevalence of each phase of mitosis in both LM and WSI. Comparing the prevalence of different mitotic phases, prophase mitosis in WSIs showed 18% reduction

when compared with their respective counts in LM, while metaphase showed 8% reduction. Atypical mitoses showed 9% reduction and both anaphase and telophase showed about 2% reduction from their corresponding counts in LM. Chromatin asymmetry was observed in 91% of atypical mitoses. Most typical mitoses (89%) were relatively similar in size to nearby neoplastic cells, while 60% of atypical mitoses showed a relative increase in size. **Table 15** shows the morphological features of mitosis recorded by each mitotic phase.

Table 13: The prevalence of mitotic phases in LM

Type	Mean	range	Minimum-maximum
Metaphase	45	48	26.6-48.3
Prophase	37.38	44.85	15.15-60
Anaphase	2.2	25	0-25
Telophase	1.5	7.7	0-7.7
Atypical mitosis	14.3	57.6	0-57.6

Table 14*:* The prevalence of mitotic phases in WSI

Type	Mean	range	Minimum-maximum
Metaphase	45.88	53.37	19.35-72.73
Prophase	32.36	54.99	6.21-61.11
Anaphase	4.2	27.5	0-27.5
Telophase	2.3	13.6	0-13.6
Atypical mitosis	15.3	61.29	0-61.29

Table 15: Morphological features of mitosis recorded by each mitotic phase.

Features	Metaphase (n=418, 43%)	Prophase (n=278, 29%)	Anaphase (n=43, 4%)	Telophase (n=15, 2%)	Atypical (n=210, 22%)	P value
Nuclear membrane						
Absent	418 (100)	278 (100)	43 (100)	15 (100)	210 (100)	
Present	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	
Hairy outline						
Absent	245 (59)	141 (51)	23 (53)	8 (53)	138 (66)	**0.02**
Present	173 (41)	137 (49)	20 (47)	7 (47)	72 (34)	
Outline or border of the nuclei						**<0.0001**
Regular	56 (13)	38 (14)	11 (26)	5 (33)	3 (1)	
Irregular	362 (87)	240 (86)	32 (74)	10 (67)	207 (99)	
Nuclear symmetry						
Symmetrical	416 (99)	252 (91)	43 (100.0)	8 (53)	190 (91)	**<0.0001**
Asymmetrical	2 (1)	26 (9)	0 (0.0)	7 (47)	20 (9)	
Homogeneity of chromatin						**<0.0001**
Homogenous	404 (97)	224 (81)	43 (100)	15 (100)	191 (91)	
Heterogeneous	14 (3)	54 (19)	0 (0)	0 (0)	19 (9)	
Cytoplasmic features						
Slightly eosinophilic	350 (84)	211 (76)	27 (63)	15 (100)	181 (86)	**<0.0001**
Deeply eosinophilic	37 (9)	17 (6)	12 (28)	0 (0)	15 (7)	
Pale	9 (2)	6 (2)	4 (9)	0 (0)	5 (2)	
Basophilic	3 (1)	10 (4)	0 (0)	0 (0)	2 (1)	
Rim or not seen	19 (4)	34 (12)	0 (0)	0 (0)	7 (3)	

Relative cell size compared to nearby cells						
Similar	373 (89)	241 (87)	22 (51)	14 (93)	84 (40)	**<0.0001**
Increased	34 (8)	30 (11)	19 (44)	1 (7)	126 (60)	
Decreased	11 (2)	7 (2)	2 (5)	0 (0)	0 (0)	

3.6.7 Refining existing mitosis cutoff in WSI.

To refine the cutoff of mitosis counting to be optimal for BC grading using WSI, the microscopic mitotic count was plotted against patient outcome in 341 patients using X tile software. The generated cutoffs were 11-22 whereas score 1, (<12 mitoses), score 2 (12 to 22 mitoses) and score 3 (>22 mitosis). With plotting the digital mitotic count (using the same area size that was used in microscopic scoring) against BCSS, refined cutoffs were generated as followed, score 1: < 8 mitosis, score 2: 8-19 mitoses and score 3:>19 mitoses. Univariate survival analysis showed a significant association between LM mitotic score and BCSS (p=0.002). In addition, a significant association was observed between BCSS and digital mitotic count using the existing cutoff of mitosis (p=0.02). With applying the refined cutoff on WSI, and repeating the analysis, a strong association between BCSS and digital score was observed (p=0.006). Also, there was much improvement in score matching between the LM and WSI mitotic scores using the refined cutoff (**Table 16**). For validation of the refined cutoff, a large number of cases (n=2583) were scored for mitosis in 3 mm^2 in WSI, then the refined cutoff was applied then tested against patient's outcome and there was a strong association between mitotic scores in WSI and poor BCSS and short DMFS (both p<0.0001) (**Figure 3-6)**

Table 16: Comparison between Mitotic score with LM versus with WSI with the refined cutoff

Mitotic scores on LM	**Mitotic scores in WSI**			**x2** **P value**
	Score 1 **(n, %)**	**Score 2** **(n, %)**	**Score 3** **(n, %)**	
Score 1	134 (93)	10 (7)	0 (0)	511.4 **<0.0001**
Score 2	4 (4)	90 (90)	6 (6)	
Score 3	0 (0)	10 (10)	87 (90)	

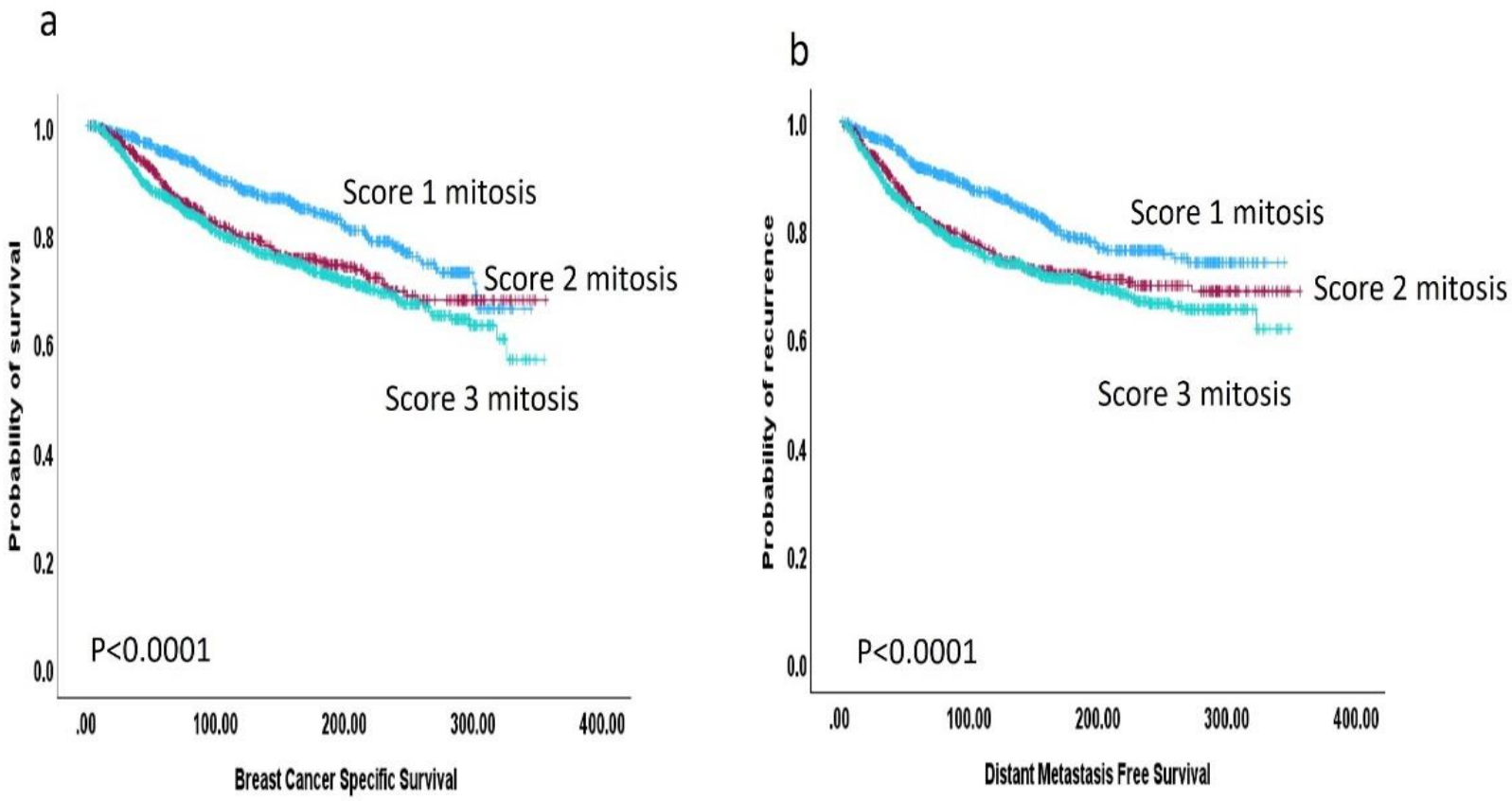


Figure 3-6: Kaplan Meir curves showing the association between digital mitotic scores and both BCSS (a) and DMFS (b) with applying the refined cutoff for mitosis.

3.7 Discussion

The practice of pathology started changing with the introduction of slide scanners which enable the acquisition of pathology information from glass slides and translate it into WSIs. WSI provides the possibility of viewing and manipulating pathology samples on a computer screen in a way comparable to LM [136]. In addition, WSI boasts many advantages over glass slides including easy image accessibility, sharing and annotating [405]. Although, some studies have shown that the diagnostic performance of WSI is comparable to that of a conventional microscope, implementing WSI in primary diagnostics is still in its infancy and needs validation and adaptation of current practice guidelines. Mitotic count is one of the independent prognostic factors in BC and is a well-established component of the Nottingham histological grade [63, 137].

The presence of certain morphological features such as clearly visible hairy-like projections of chromatin help to identify the mitoses in LM, which has the advantage of fine-tuning adjustment. However, these features cannot be identified easily in all mitoses and this issue is likely to be more pronounced in WSIs. Previous studies indicated low to moderate concordance rates in detecting mitoses using LM [398, 406].

The guidelines criteria for mitoses identification and counting, which are currently in use are derived from LM practice and whether they can be applied as such to WSIs or need further adjustments remains to be investigated. The lack of defined criteria for assessing mitoses digitally

may lead to inconsistent reporting which would have clinical implications [106, 136].

AI will be the future technology to help diagnosis and prognosis of BC. This study`s findings can be applied to facilitate BC grading through accurate mitotic counts in WSI.

A set of refined morphological criteria for identifying true mitoses and differentiating them from mimickers in WSIs was proposed. Although hairy-like projections were considered the cornerstone feature in identifying mitoses in LM [149, 407], these projections are not obvious in approximately 60% of mitotic figures in WSIs. Assessing mitoses in WSI in previous literature depended on existing criteria to define mitosis in LM, which include absent nuclear membrane and clear hairy-like projections at the outline of the nucleus [136, 407]. However, the lack of fine-tuning facility in whole slide scanners makes identifying these fine details a challenging task. As a result, numerous mitoses may go undetected in WSI, leading to potentially significant underscoring [136].

Refined criteria can also increase the reliability of AI-based algorithms for mitosis detection by improving the quality of the training sets. Using the refined criteria could improve overall BC grading and inter-observer agreement as shown in this study. Underestimation of BC grade could occur if the existing LM-based mitosis cutoffs are used in mitosis counting using WSI. This is caused by decreased the sensitivity of mitosis identification in WSI.

Several mitotic phases, particularly anaphase and telophase, were better detected in WSI as the two chromatin poles were easily appreciated. Prophase was difficult to detect on LM and was less likely to be picked up in WSI [408].

The challenges in differentiating true mitoses from other mimickers such as apoptotic bodies, lymphocytes or other dark hyperchromatic dense tumour cells are increasing with the use of one focal plane scanners. Most of the mitoses showed absent hairy-like projections but were still identifiable as true mitoses due to irregular nuclear outline which may reflect the condensed nuclear material and the absence of nuclear membrane. Ragged (irregular) border was evident in most mitoses in WSI, while most mimickers showed regular nuclear outlines.

It was observed that almost all mitoses showed abundant cytoplasm, while its absence favours mimickers. The various shades of eosinophilic cytoplasm were categorised as slightly eosinophilic (which represented the cytoplasm of most of mitoses) and deeply eosinophilic (which was compatible with most apoptoses, and which was seen in only a few mitoses). In addition, some mitoses showed clear cytoplasm, and few had basophilic or amphophilic cytoplasm. However, these mitoses usually show other characteristic features including absent nuclear membrane or ragged nuclear borders, which help identify them. Cytoplasmic granularity was a unique feature observed in mitoses only and could result from the fragmentation of the Golgi apparatus or ribosomal activity during mitosis [409].

Candidate figures lacking one or more criteria of the previously mentioned criteria should be examined with caution before being categorised as mitoses. Although some studies recommend excluding challenging figures when using WSIs such as prophase which can be easily confused with other mimickers [400, 410] and telophase which may be counted as two mitoses [155], this may affect the mitotic score and hence overall grading. By applying the refined criteria, prophase identification may become less difficult, and improvement of inter-observer concordance was observed in the identification of prophase after applying these criteria.

A systematic reduction in mitosis count in WSI compared to LM was observed with an overall rate of 20% when evaluating mitoses within a specified area. Notably, 10% of cases in which mitotic counts were around the cutoff showed down scoring in WSI. The lower accuracy of identifying mitoses in WSI may be due to the limitation of whole slide images caused by missing the fine-tuning capability (z-axis) in WSI with some mitoses becoming in focus and easily identifiable and some become out of focus and can either be missed or not identified as mitotic figures. Z-stacking or multilayer scanning is a valuable method to overcome this issue, but it requires considerably larger sized files and extended scanning times [155]. Based on the previous, the existing cutoff of mitosis using LM would lead to an underestimation of the mitotic grade in WSI. Therefore, refined cutoffs for mitosis were generated in WSI to reflect the degree of reduction in mitotic count between the LM and WSI

using the same area size. A large cohort of patients was used to validate the refined cutoff in WSI and marked improvement of the association of mitotic scores in WSI and patient outcome was noticed after applying the refined cutoff. This can improve the grading of BC in WSI.

Conclusion: With the refined criteria for mitosis identification, recognition of true mitoses in WSI can be improved. Although there is a tendency to underscore mitotic counts in WSIs, this may be addressed by adjusting mitotic scores for BC grade. With the refined mitosis cutoffs, scoring mitoses using WSI in BC can be as reliable and reproducible as when using a microscope. Further development of software and image quality will improve the use of WSI in routine pathology practice.

Chapter 4 The role of atypical mitosis in BC

4.1 Introduction

During mitosis, the chromosomes condense and attach to spindle fibres that pull one copy of each chromosome to opposite sides of the cell [411]. In malignant tumours, mitosis can occur more frequently compared with normal tissues because of increased cell-cycle activity [412]. In addition to the common features of the absence of the nuclear membrane and hairy extension of nuclear materials, typical mitosis is characterised by a specific arrangement of the chromatin [413]. This includes either clotted (prophase), in plane (metaphase/anaphase), or in separate clots (telophase) [136]. However, an atypical mitotic figure is defined as anything other than the typical form of normal mitosis including multipolar, ring, dispersed, asymmetrical, anaphase bridge and lag-type mitoses [143]. Atypical mitosis refers to the presence of unusual, dysregulated, and random assembly of the nuclear materials within the dividing cells which results in abnormal morphology of mitoses [138]. It has been shown that multipolarity and lagging chromosomes are often easily detected, but anaphase bridges are not easily identified [414].

Although counting mitosis is an overall score of typical and atypical mitosis without differentiation, the presence of atypical mitotic figures is generally acknowledged to be a feature of malignancy [415]. In multipolar mitosis, for example, it is obvious that with the division of each of these cells into three daughter cells, this will result in a grossly unequal distribution of chromosomes [141]. Increasing evidence reveals that errors in mitosis can also promote the direct and indirect acquisition of

DNA damage and chromosome breaks [416]. It has been shown that atypical mitosis reflects the underlying genetic and chromosomal abnormalities such as chromosomal instability and aneuploidy [412, 417].

Atypical mitosis provides prognostic value in certain tumours including urothelial carcinomas [147, 418]. In pancreatic cancer, the existence of multipolar mitosis was an independent prognostic factor [419]. Furthermore, the presence of atypical mitosis was associated with a shorter survival in non-epithelioid mesothelioma [420].

In BC, a previous study showed that the presence of atypical mitosis is associated with poor prognosis [147]. However, it utilised a limited number of cases. Therefore, this study aimed to evaluate the clinal and biological value of atypical mitosis in BC using large cohorts.

4.1.1 Hypothesis and aims.

In this chapter it was hypothesised that atypical mitosis provides additional prognostic significance in BC and its molecular subtypes. In addition, it was hypothesised that the genetic pathways associated with atypical mitosis have important role in the progression of BC. Therefore, the aim of this chapter is to:

1- Quantify atypical mitosis in 2 large cohorts (n= 1878) of BC patients using WSI to assess the relationship between atypical mitosis and disease outcome in BC.

2- Explore the molecular changes and pathways associated with atypical mitosis using the publicly available RNA sequencing data (RNA-Seq).

4.2 Patients and Methods

This study was conducted on two retrospective cohorts of well-characterised primary invasive BC specimens. These cohorts include an internal cohort (Nottingham cohort) and the publicly available cohort; The Cancer Genome Atlas Program (TCGA).

4.2.1 Characteristics and significance of atypical mitosis

Full-face sections of the Nottingham cohort (n=846) were used for the evaluation the clinical value of atypical mitosis. For this purpose, three to four tumour sections per case were histologically reviewed, and the representative section, defined as the section with the highest tumour grade and burden, was selected to be included in the study. The clinicopathological data and outcome data were available for this cohort (**Table 17**). This cohort of BC patients had uniform adjuvant treatment based on NPI and ER status and none of the patients received neoadjuvant therapy. The median follow-up time was 138 months. Freshly cut, 4-µm full-face H&E sections were prepared from the selected FFPE blocks and digitally scanned at 40× magnification using a high-throughput automated scanner (Panoramic 250 Flash III: 3D-Histech, Budapest, Hungary). WSIs were viewed by Case Viewer software (version 2.2.0.85; 3D-Histech) on a full-screen panel (21-inch screen with 1366 × 768 pixels resolution) at 40x magnification.

Table 17: Clinicopathological characteristics of the Nottingham cohort

Variables	No (%)
Tumour size	
≤ 2cm	487 (58)
> 2cm	359 (42)
Histologic tumour grade	
Grade 1	75 (9)
Grade 2	191 (22)
Grade 3	580 (69)
Histologic tumour types	
No special type (NST)	639 (75)
Lobular	53 (6)
Other special types	116 (14)
NST mixed	38 (5)
Molecular subtypes	
Luminal	380 (44)
Triple negative	324 (83)
HER2 enriched	116 (13)
Lymph node status	
Negative	533 (63)
Positive	312 (37)
Lymhovascular invasion	
Absent	576 (68)
Present	270 (32)
Nottingham prognostic index	
Good prognostic group	160 (19)
Moderate prognostic group	524 (62)
Poor prognostic group	161 (19)

4.2.2 **Identification and counting of atypical mitosis in WSI.**

Typical and atypical mitosis were quantified separately, and the overall mitotic count was calculated over an area of 3 mm^2 in hotspots. Typical mitoses were identified by their morphology (prophase, metaphase, anaphase, and telophase) [149, 421]. Atypical mitoses were defined as mitoses with any morphological appearance other than the typical forms and include (i) lag atypical mitosis defined as a rod-like structure with one or more unattached chromosomes at one side or both sides of metaphase; (ii) spindle multipolarity which shows three attached poles (tripolar), four attached poles (tetrapolar) or more than 4 poles (multipolar); (iii) dispersed mitosis which is represented by multiple dispersed chromosomes with no specific shapes; (iv) polar asymmetry in which there are two separated masses with unequal size; (v) anaphase bridge that is characterised with a string of chromosome attached to one poles of anaphase; (vi) ring mitosis which is an unusual feature in which chromosomes are displaced to the periphery of the cells (**Figure 4-1** and **Figure 4-2**). Ambiguous structures such as cells with dark hyperchromatic nuclei and/or apoptotic cells were excluded. It was observed that the distribution of typical and atypical mitoses is not uniform and that some cases show high overall mitotic count but with low number of atypical mitoses while others showed high proportion of atypical mitosis. Therefore, the ratio between atypical and typical mitosis was calculated to assess its clinical value. Atypical-to-typical mitoses ratio was calculated by the number of atypical mitoses divided by the number of the typical mitoses per 3 mm^2.

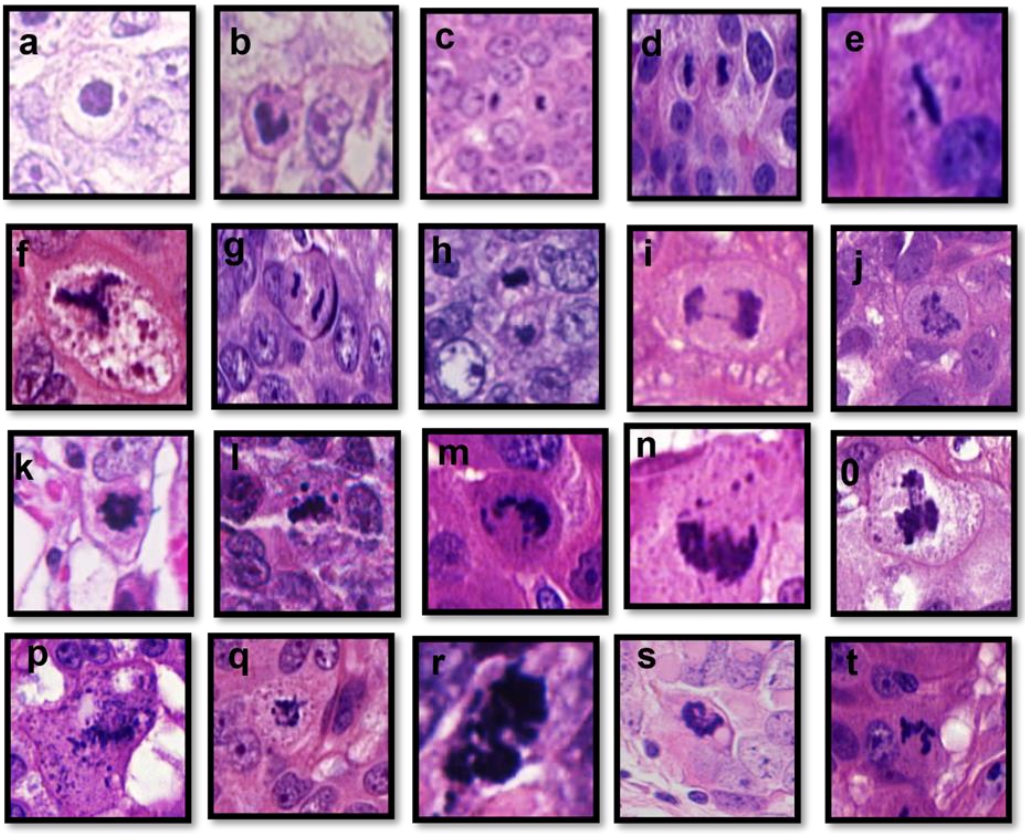


Figure 4-1: Morphological features of typical and atypical mitosis (40×)
a- prophase shows round shape with hairy-like appearance, b- metaphase shows curvilinear structure with slightly eosinophilic cytoplasm, c- anaphase shows two pulled away uniform chromatin masses, d- telophase shows two uniform completely separate cells, e- lag atypical mitosis shows two unattached chromosomes at each side of metaphase plate, f- tripolar atypical mitosis shows three poles with hairy outline and slightly granular eosinophilic cytoplasm, g- polar asymmetry shows either two pulled-away chromatid unequal in size, enlarged size and slightly eosinophilic cytoplasm or h- two non-uniform separated cells, i- anaphase bridge shows a string of chromatin extending from one pole of the anaphase to the other connected poles, j- dispersed mitosis shows non-clumped multiple chromosomes with slightly eosinophilic cytoplasm and enlarged size, k- ring atypical mitosis shows ring-like structure with hairy outline, l- lag atypical mitosis shows one un-attached chromosomes at one side of metaphase plate, m-t other forms of atypical mitoses showing asymmetry of the chromatin masses within the cells.

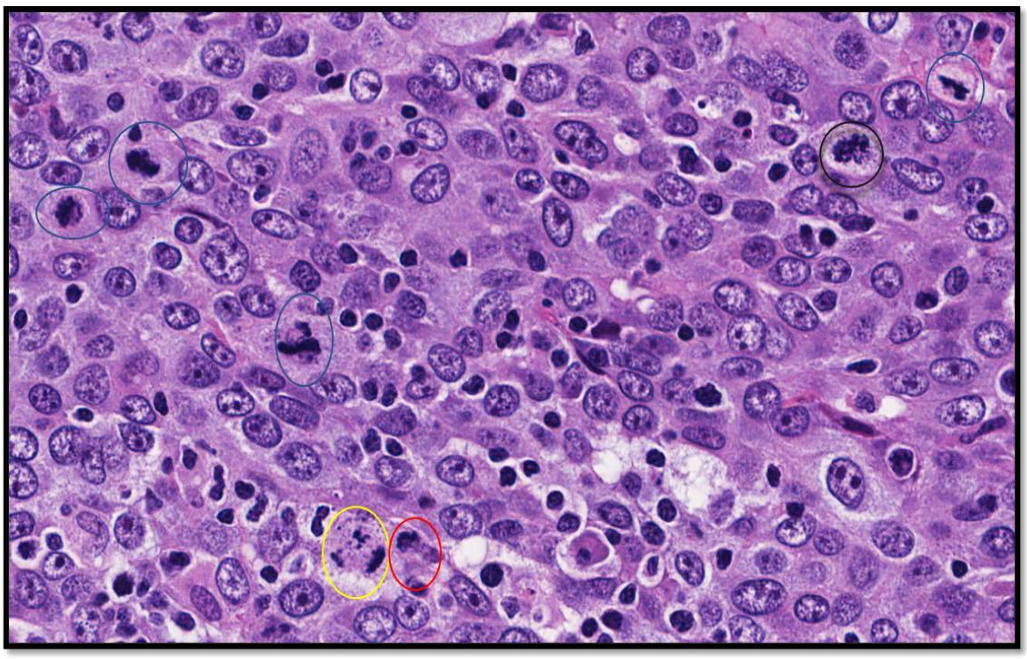


Figure 4-2: Representative image showing different types of atypical mitotic figures in one field (20x).

4.2.3 Prognostic and predictive value of atypical mitosis in BC

Atypical mitosis, overall mitoses, atypical-to-typical mitoses ratio were compared with the clinicopathological parameters of BC. In addition, atypical mitosis scores were tested against patients` outcome. The role of atypical mitosis in each of the molecular classes was assessed and the association of atypical mitosis with adjuvant therapy was evaluated.

4.2.4 Deciphering the molecular changes linked with atypical mitoses.

For deciphering the potential molecular changes linked with atypical mitoses, 1032 primary invasive BC specimens from TCGA-BRCA cohort (cBioPortal.org) were used [422]. Most of the specimens had online available H&E digitalised image prepared from FFPE tissue blocks. In

this cohort, atypical and typical mitoses was counted at 40x magnification using the same methodology as described above. Differentially expressed genes (DEGs) linked with atypical mitosis were identified using the DESeq2 implementation in R statistical tool [423, 424]. This part was done by Prof Nigel Mongan. Significant differentially expressed genes were identified based on log2 fold change (≥±1) combined with q-value (<0.05). The web-based gene set enrichment analysis tool (WebGestalt) was used to explore significantly enriched pathways and gene ontologies (GO) based on the identified DEGs [394]. The common DEGs were identified between atypical and overall mitosis using the Venny version 2.0 tool [425].

4.2.5 Statistical analysis

Statistical software 27 (SPSS, Chicago, IL, USA) was used for statistical analysis. Chi-square was used for analysis of categorical data. For statistical analysis, atypical mitosis, the overall mitotic counts and atypical-to-typical mitoses ratio were categorised based on the median, into low and high groups, and were correlated with the clinicopathologic parameters. The Kaplan–Meier survival method and the log-rank test were used for survival curves. For all tests, $P < 0.05$ (two-tailed) were considered statistically significant.

4.3 Results

4.3.1 Characteristics of atypical mitoses in BC

The median number of typical mitoses 17 (range 0-120 mitoses) while the median number of atypical mitoses was 4 (range 0-103 mitoses) (

Figure 4-3). The median number of overall mitosis (both typical and atypical combined) was 21 (range 0-153 mitoses per case). Atypical mitoses were present (1 atypical mitosis or more) in 77% (n=649/846) of the cases. However, when the cohort was dichotomised into high and low groups based on the median number of atypical mitoses (n=4), 48% of cases showed high atypical mitosis score. The mean of atypical-to-typical mitoses ratio was 0.4 (median=0.27; range 0-4).

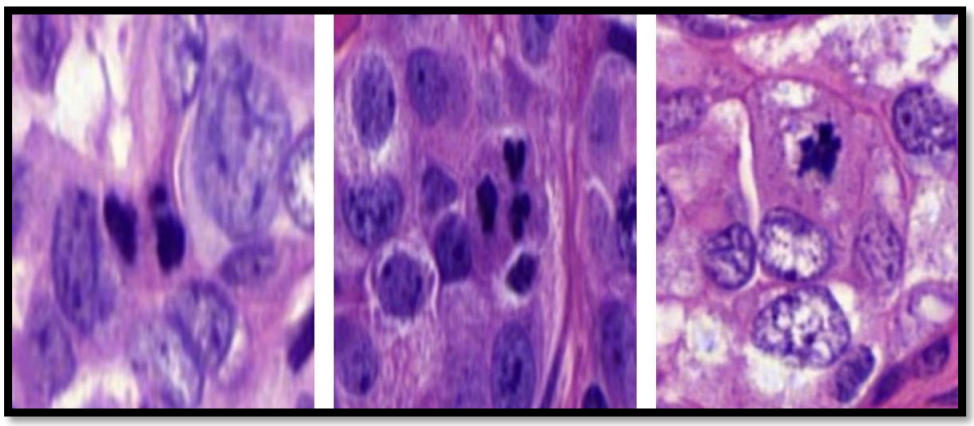


Figure 4-3: Unfamiliar shapes of atypical mitosis

4.3.2 Correlation with other clinicopathological parameters

There was a significant association between high overall mitosis and parameters characteristic of aggressive tumour behaviour including larger tumour size (p<0.0001), TNBC (p<0.0001) and poor NPI (p<0.0001) (**Table 18**). Similarly, high atypical mitoses showed a significant association with high grade (p<0.0001), larger tumour size (p<0.0001), NST tumour type (p<0.0001), poor NPI group (p<0.0001) and TNBC phenotype (p<0.0001) (**Table 19**). In addition, high atypical-to-typical mitoses ratio showed similar associations with

clinicopathological parameters **(Table 20).** However, there was no association between either overall mitosis, atypical mitosis or atypical-to-typical mitoses ratio and LN status or LVI.

4.3.3 Correlation of atypical mitosis with overall mitosis.

There was a positive correlation between atypical mitosis and the overall mitosis in the whole cohort (r=0.81, p<0.0001) (F). With classification of the whole cohort into different molecular classes, there was less significant association between atypical mitosis and the overall mitosis in the TNBC phenotype (r=0.6, p<0.0001) (F).

Table 18: The relationship between overall mitoses and clinicopathological parameters

Parameters	Low mitoses≤21	High mitoses>21	X^2 *P*-value
Tumour size			28. 2
≤ 2cm	284 (67)	143 (33)	**<0.0001**
> 2cm	203 (48)	216 (52)	
Tumour grade			
Grade 1	75 (99)	1 (1)	280.3
Grade2	172 (90)	18 (10)	**<0.0001**
Grade 3	180 (31)	400 (69)	
Histologic types			
NST	251 (39)	388 (61)	135.5
Lobular	35 (83)	7 (17)	**<0.0001**
Other special types	88 (90)	10 (10)	
Mixed tumours	53 (79)	14 (21)	
Molecular subtype			
Luminal	298 (78)	82 (22)	230.8
Triple negative / basal	71 (21)	253 (79)	**<0.0001**
HER2+	45 (39)	71 (61)	
Lymph node status			
Negative	269 (51)	264 (49)	0.002
Positive	158 (51)	154 (49)	1
Lymphovascular invasion			
Absent	269 (48)	287 (52)	0.1
Present	138 (51)	132 (49)	0.8
Nottingham prognostic index			
Good prognostic group	149 (93)	11 (7)	145.5
Moderate prognostic group	221 (42)	303 (58)	**<0.0001**
Poor prognostic group	57 (35)	104 (65)	

HER2, human epidermal growth factor receptor 2. Significant P values are in **bold.**

Table 19: The relationship between atypical mitosis and clinicopathological parameters

Categories	Low atypical mitoses ≤ 4	High atypical mitoses >4	X^2 P-value
Tumour size			
≤ 2cm	289 (59)	198 (41)	22.4
> 2cm	154 (43)	205 (57)	**<0.0001**
Tumour grade			
Grade 1	73 (97)	2 (3)	229.3
Grade2	168 (88)	23 (12)	**<0.0001**
Grade 3	202 (35)	378 (65)	
Histologic types			
No special type (NST)	268 (42)	371 (58)	115.4
Lobular	44 (83)	9 (17)	**<0.0001**
Other special types	102 (88)	14 (12)	
Mixed NST	29 (76)	9 (24)	
Molecular subtype			
Luminal	293 (77)	87 (23)	117.5
Triple negative	92 (28)	232 (72)	**<0.0001**
HER2 +	44 (45)	72 (55)	
Lymph node status			
Negative	285 (54)	248 (46)	0.6
Positive	158 (51)	154 (49)	0.43
Lymphovascular invasion			
Absent	305 (53)	271 (47)	0.3
Present	138 (51)	132 (49)	0.6
Nottingham Prognostic index			
Good prognostic group	149 (93)	11 (7)	132.3
Moderate prognostic group	231 (44)	293 (56)	**<0.0001**
Poor prognostic group	63 (39)	98 (61)	

HER2, human epidermal growth factor receptor 2. Significant P values are in bold.

Table 20: Relationship atypical-to-typical mitoses ratio and clinicopathological parameters

Categories	Low ratio ≤0.27	High ratio >0.27	X2 P-value
Tumour size			
≤ 2cm	267 (55)	220 (45)	10.2
> 2cm	157 (44)	202 (56)	**<0.0001**
Tumour grade			
Grade 1	66 (88)	9 (12)	126.3
Grade2	142 (74)	49 (26)	**<0.0001**
Grade 3	216 (37)	364 (63)	
Histologic types			
No special type (NST)	269 (42)	371 (58)	71.0
Lobular	34 (64)	19 (36)	**<0.0001**
Other special types	93 (80)	23 (20)	
Mixed NST	28 (74)	10 (26)	
Molecular subtype			
Luminal	260 (68)	120 (32)	93.4
Triple negative	108 (33)	216 (67)	**<0.0001**
HER2 +	45 (47)	71 (53)	
Lymph node status			
Negative	273 (51)	260 (49)	0.6
Positive	151 (48)	161 (52)	0.4
Lymphovascular invasion			
Absent	304 (53)	272 (47)	5.1
Present	120 (44)	150 (56)	**0.02**
Nottingham Prognostic index			
Good prognostic group	133 (83)	27 (19)	86.1
Moderate prognostic group	226 (43)	298 (57)	**<0.0001**
Poor prognostic group	65 (40.4)	96 (59.6)	

HER2, human epidermal growth factor receptor 2, ratio, atypical-to-typical mitoses ratio, Significant P values are **in bold.**

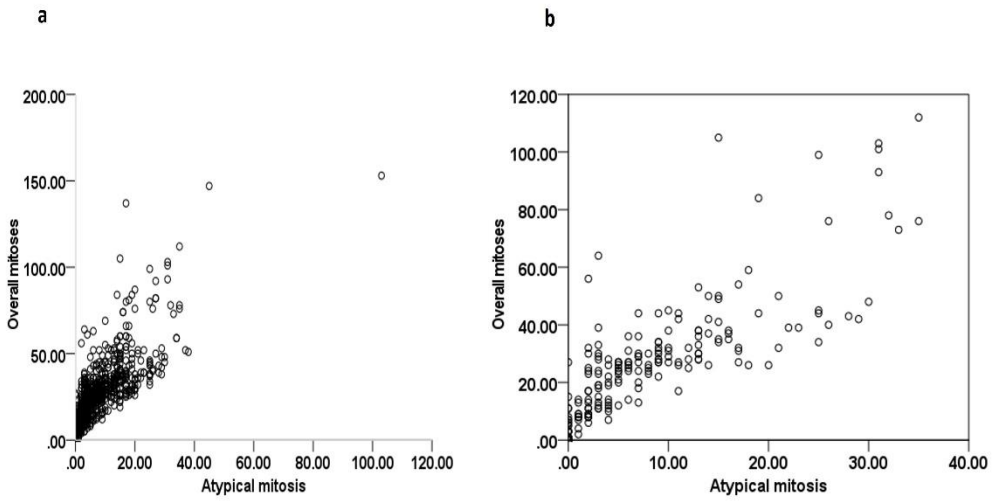


Figure 4-4: Scatter blots showing the correlation between atypical mitoses and overall mitoses in the whole cohort (a) and in triple negative breast cancer patients (b)

4.3.4 Outcome analysis

There was a significant association between the overall mitoses score and patient outcome in terms of shorter BCSS ($p < 0.0001$) and DMFS ($p < 0.0001$) (**Figure 4-5**). Although there was no association between atypical mitosis and patient outcome (**Figure 4-6**), high atypical-to-typical mitoses ratio showed a significant association with poor outcome ($p = 0.013$) (**Figure 4-7**). When the molecular classes were considered, the overall mitoses score showed a significant association with shorter BCSS in luminal BC ($p = 0.001$) but not in either TNBC or in the HER2 subtypes. Similarly, a significant association was observed between high atypical mitoses and both shorter BCSS and DMFS in the luminal BC subtype (p values = 0.006 and 0.03, respectively) (**Figure 4-8**) but not in the TNBC, nor the HER2 subtypes. However, when the atypical-to-typical mitoses ratio was applied, a significant association with outcome was observed not only in luminal BC, but also in TNBC subtype in terms of both BCSS and DMFS (**Figure 4-9**). When the whole cohort was stratified based on the adjuvant chemotherapy treatment, there were significant associations between overall mitoses, and the atypical-to-typical mitoses ratio and poor outcome ($p = 0.001$ and 0.028, respectively) in the chemotherapy naïve group but not in chemotherapy treated patients ($p = 0.9$ and 0.2, respectively). However, when the analysis was carried out on various molecular subtypes, high atypical-to-typical mitoses ratio, showed shorter BCSS in TNBC patients who received chemotherapy ($p = 0.003$). This association was not observed using the overall mitoses was ($p = 0.9$) (**Figure 4-10)**

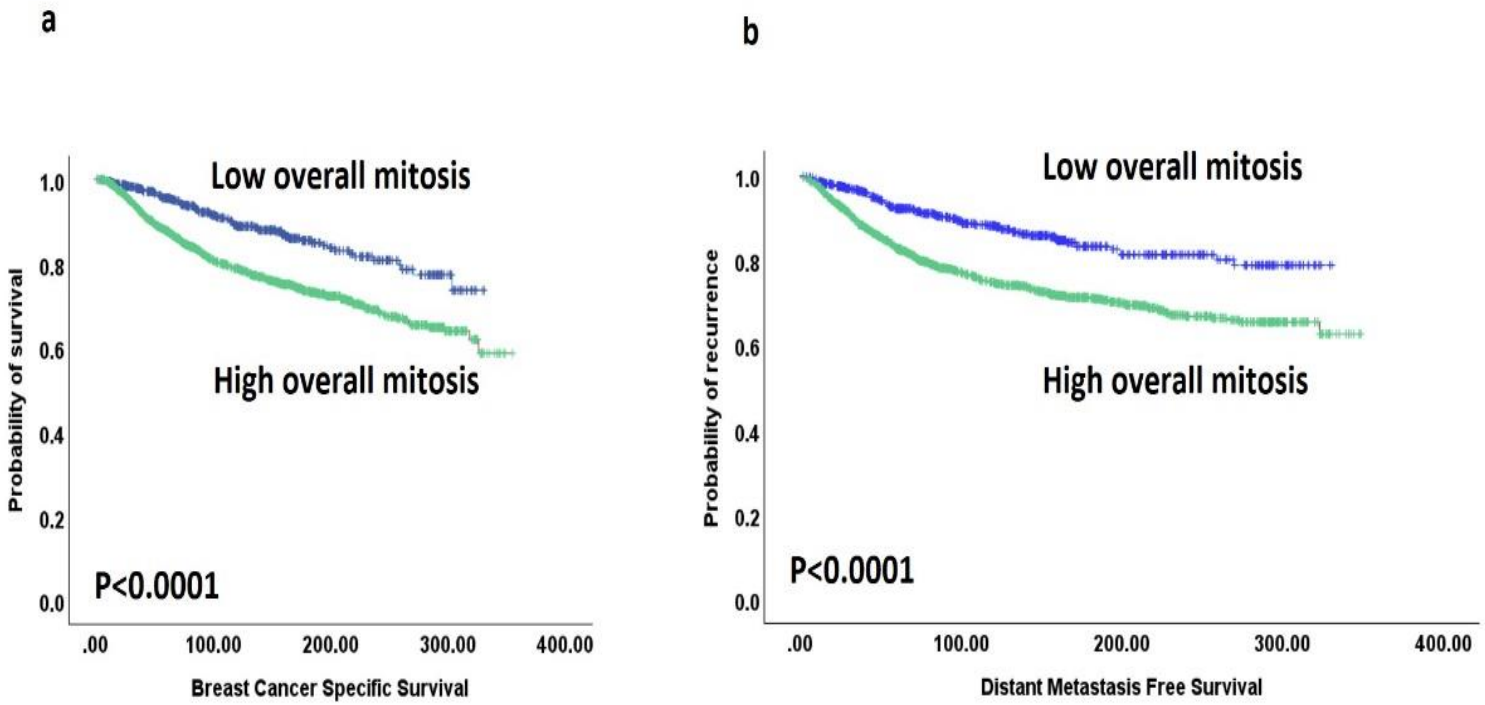


Figure 4-5: Kaplan Meier curves showing the associations of overall mitosis with (a) BCSS and with (b) DMFS.

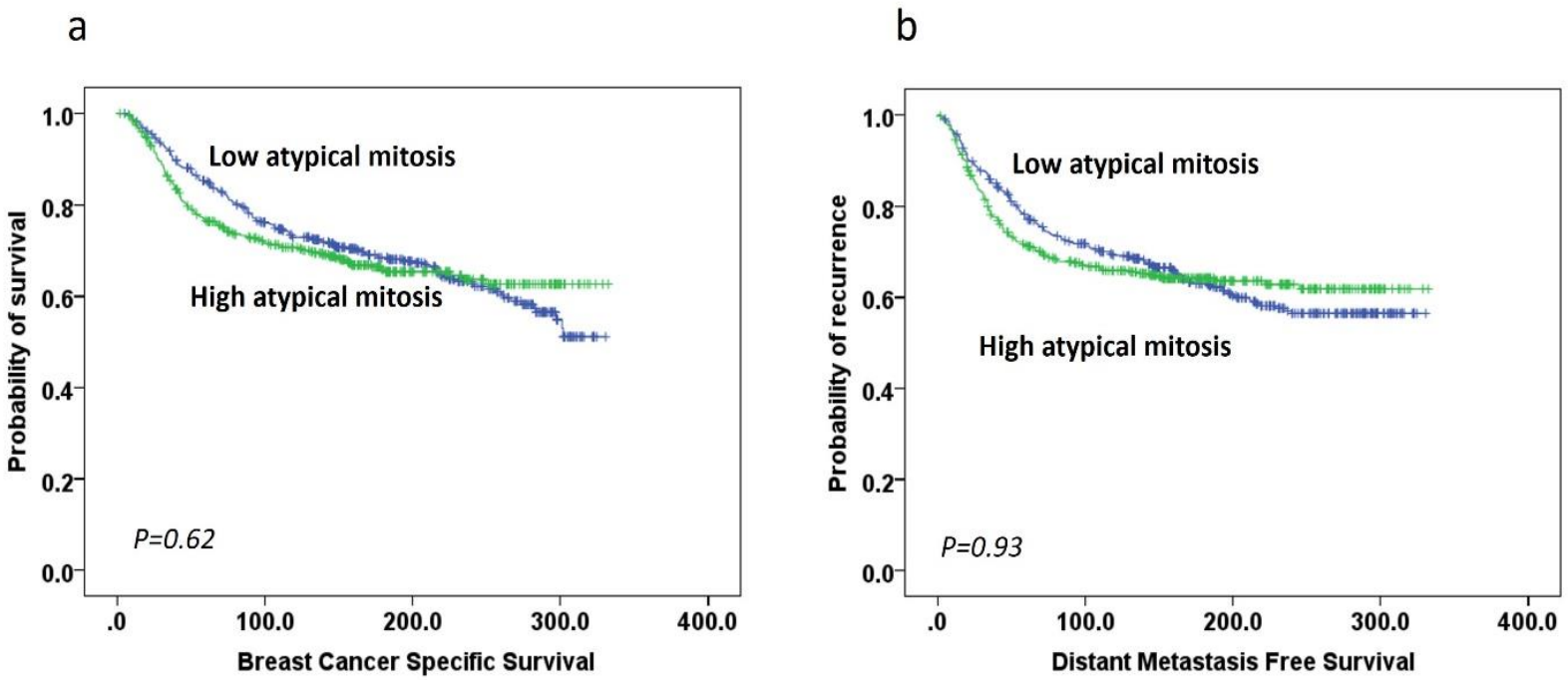


Figure 4-6: Kaplan Meier curves showing the associations of atypical mitosis with (a) BCSS and with (b) DMFS.

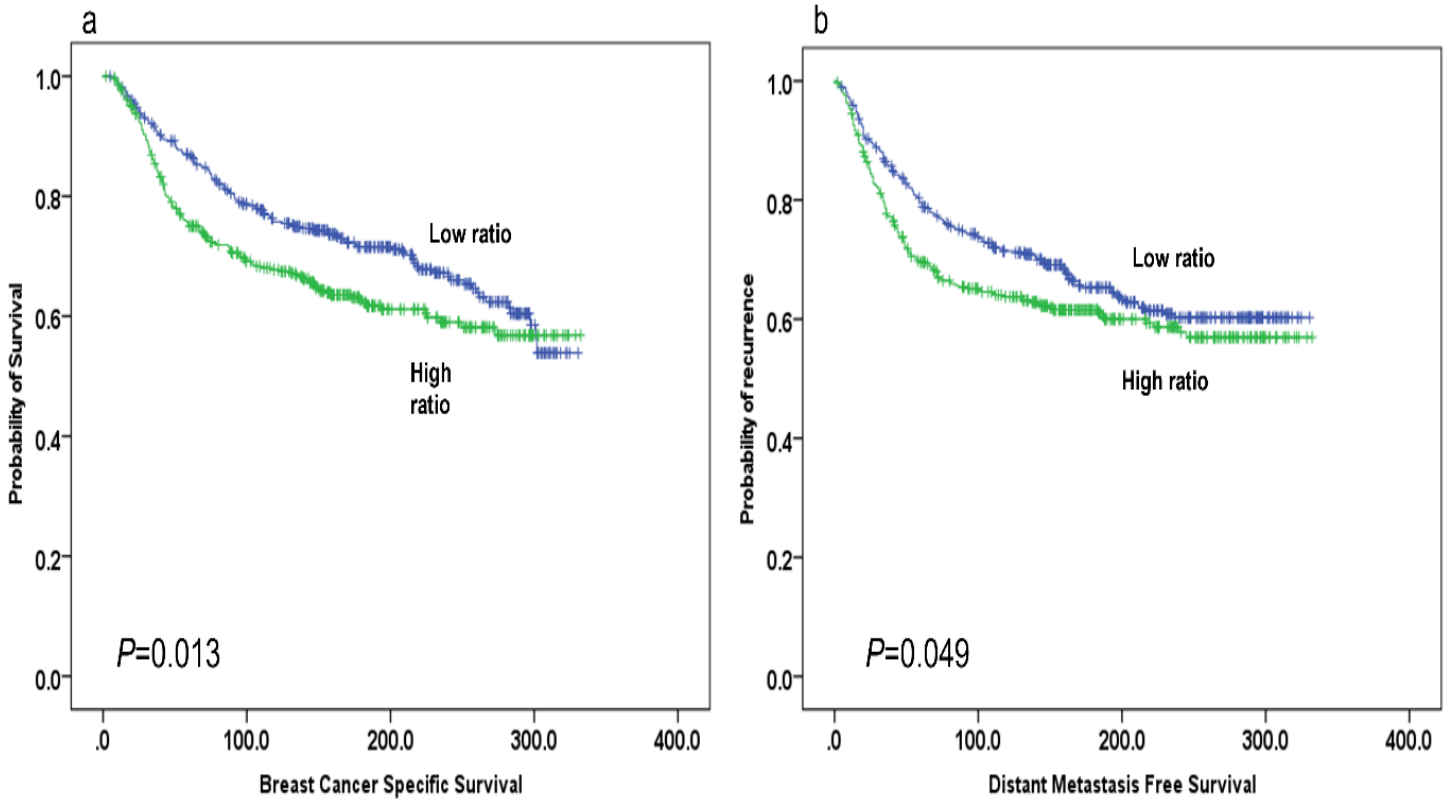


Figure 4-7: Kaplan Meier curves showing the associations of atypical-to-typical mitoses ratio with (a) BCSS and with (b) DMFS.

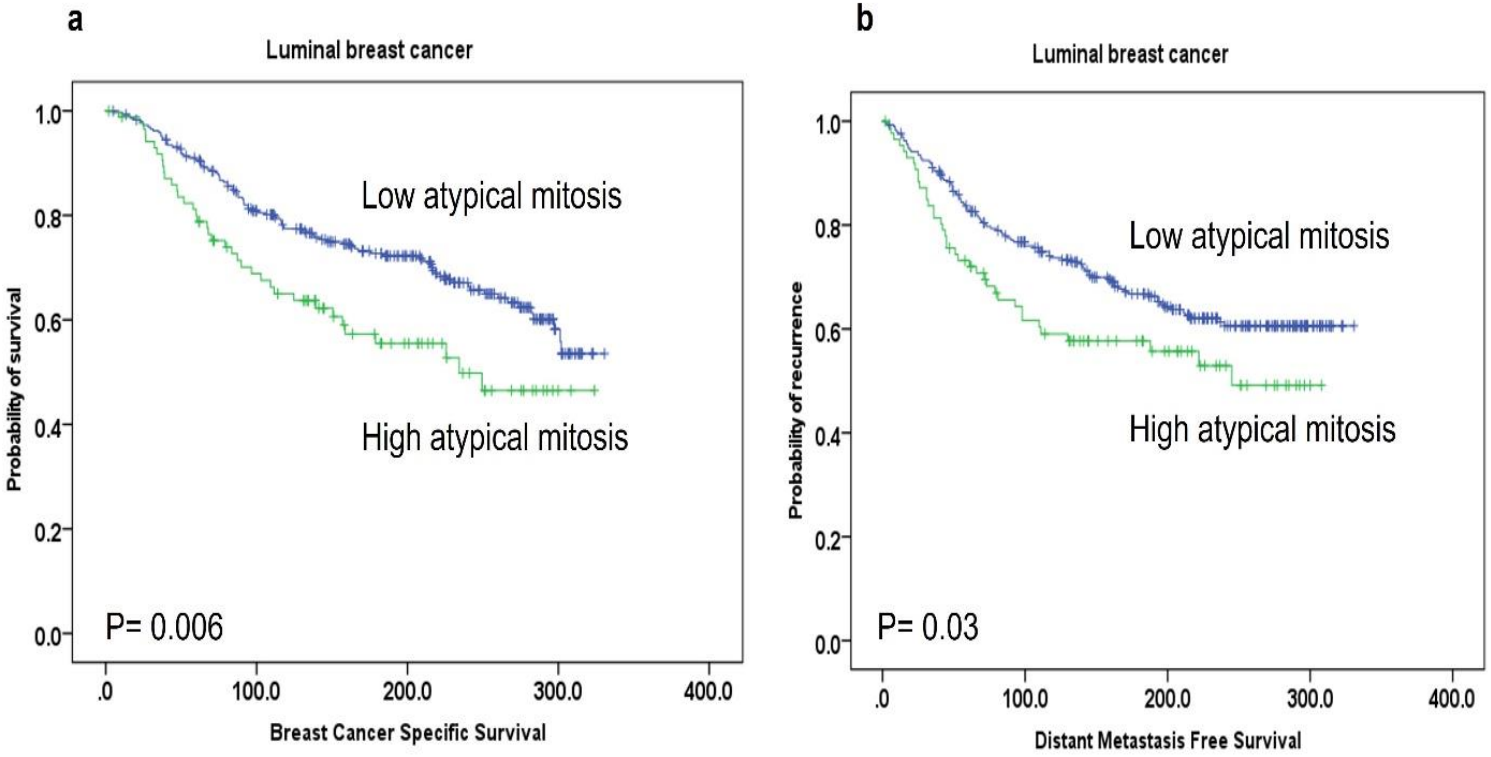


Figure 4-8: Kaplan Meier curves showing the associations of atypical mitoses with (a) BCSS and with (b) DMFS in luminal BC.

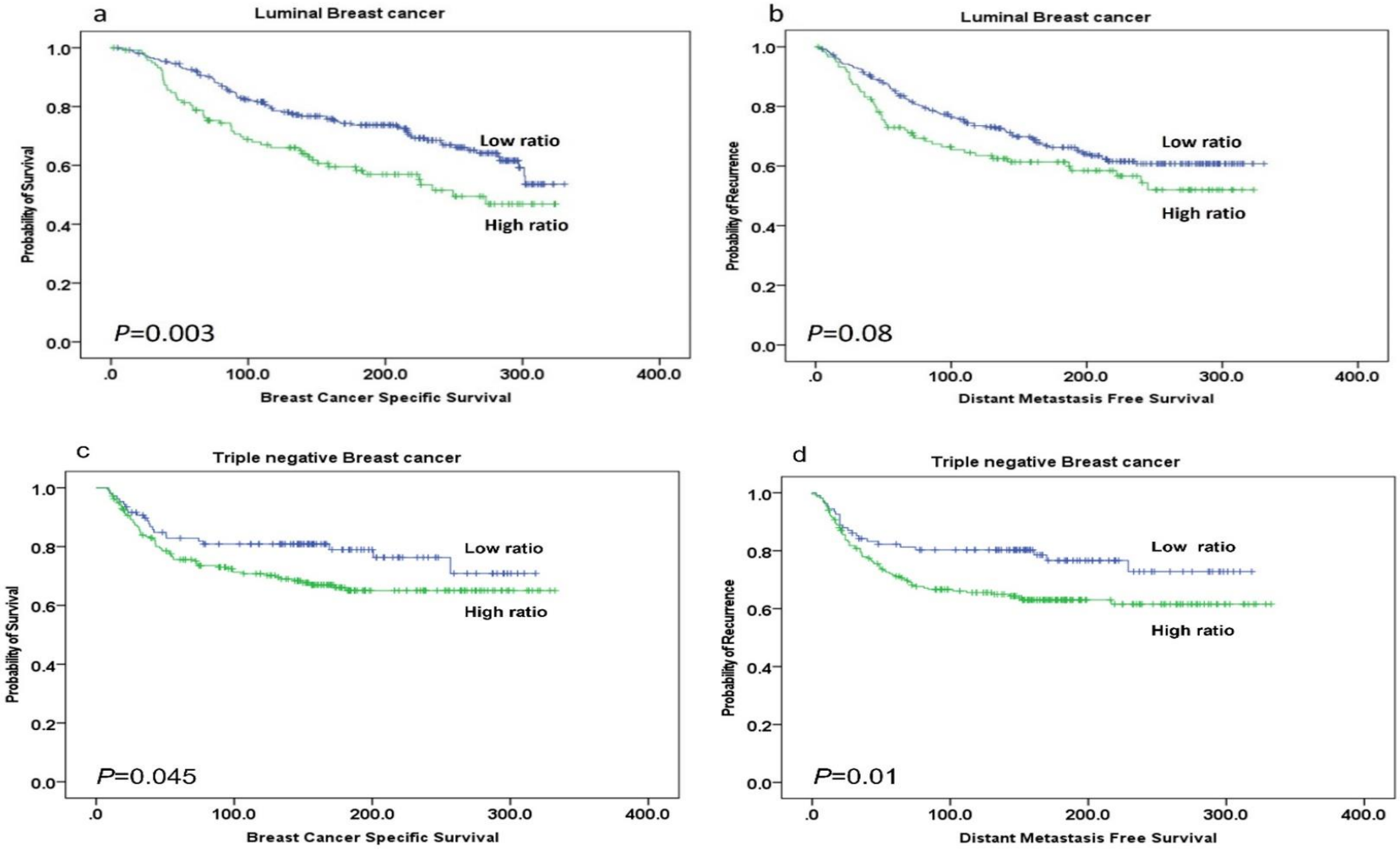


Figure 4-9: Kaplan Meier curves showing the associations of atypical-to-typical mitoses ratio with BCSS in luminal BC (a, b) while (c, d) show association of high ratio with both poor BCSS and DMFS in TNBC.

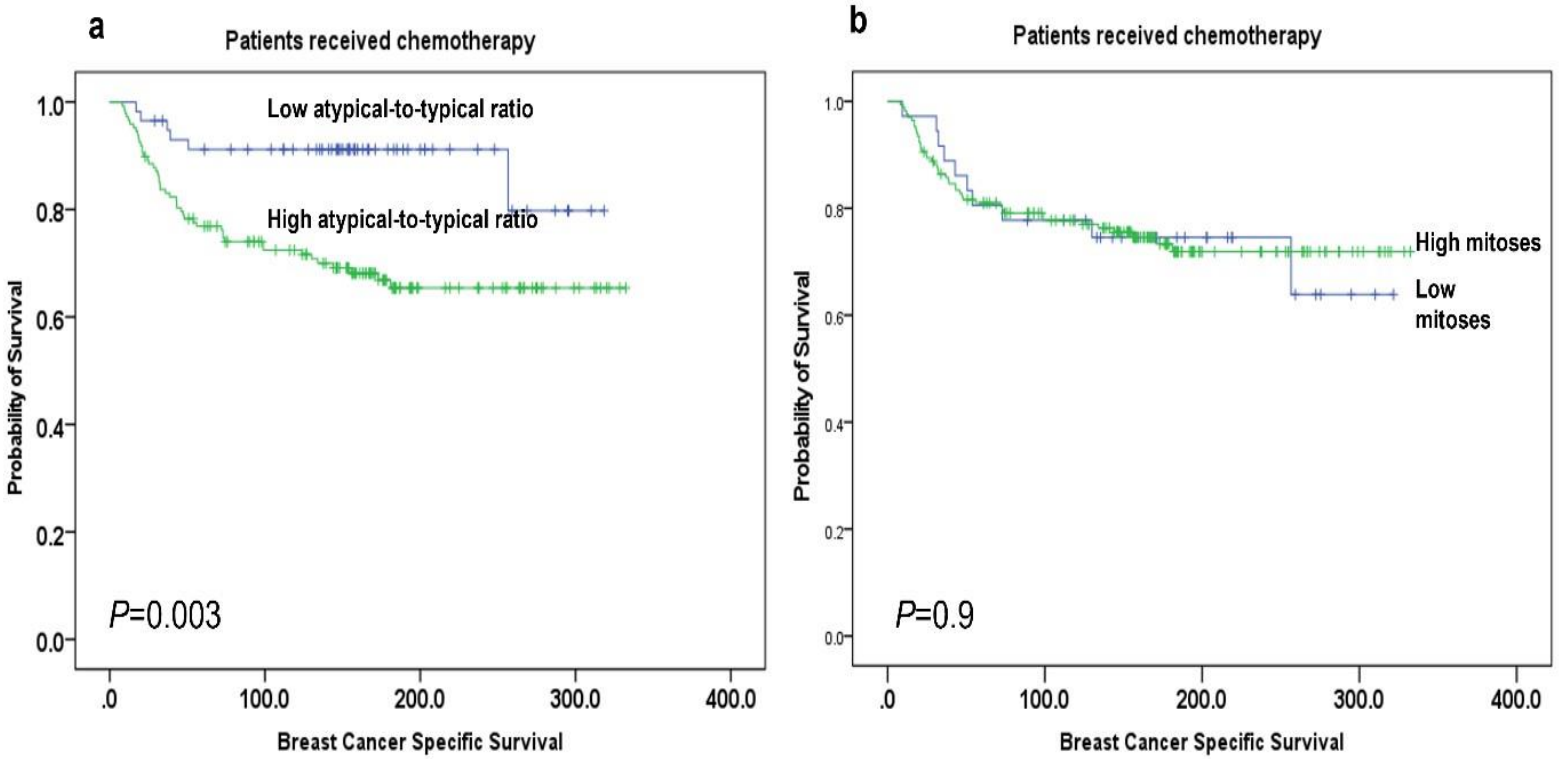


Figure 4-10: Kaplan Meier curves showing a significant association between atypical-to-typical mitoses ratio and BCSS in TNBC patients who did receive chemotherapy (a), but no association was observed in terms of overall mitoses (b).

4.3.5 Atypical mitoses and the transcriptomic profiling

The median of overall mitoses in the TCGA-BRCA cohort, was 18 (range 0–148 mitoses per case) while the median of atypical mitoses was 2 (range 0–97 mitoses). The median of typical mitoses was 13 (range 0–116) and the mean atypical-to-typical mitoses ratio was 0.3 (median = 0.2; range 0 to 5). In terms of patient outcome, there was a significant association between atypical-to-typical mitoses ratio and shorter overall survival (p=0.007). However, no association was found between atypical mitosis and patient outcome (p=0.28) (Figure 4-11).

DGE analysis identified a set of genes (n = 3247) that were linked with atypical mitoses, where 2494 genes were significantly associated with high atypical mitoses, 753 genes were associated with low score cases. Regarding typical mitoses, 1864 genes were significantly associated with high typical mitoses while 952 genes were associated with low typical mitoses (**Table 21**). There were about 60% (1623) genes commonly up-regulated across atypical and typical mitoses and 574 genes were commonly down-regulated) (**Figure 4-12** and **Figure 4-13**). The genes that showed higher expression in cases with high atypical mitoses were involved in specific biological pathways including, regulation of chromosomal organisation, localisation, and segregation, double strand DNA breakdown repair, regulation of cell cycle, and cell cycle check points as shown in (**Table 22**).

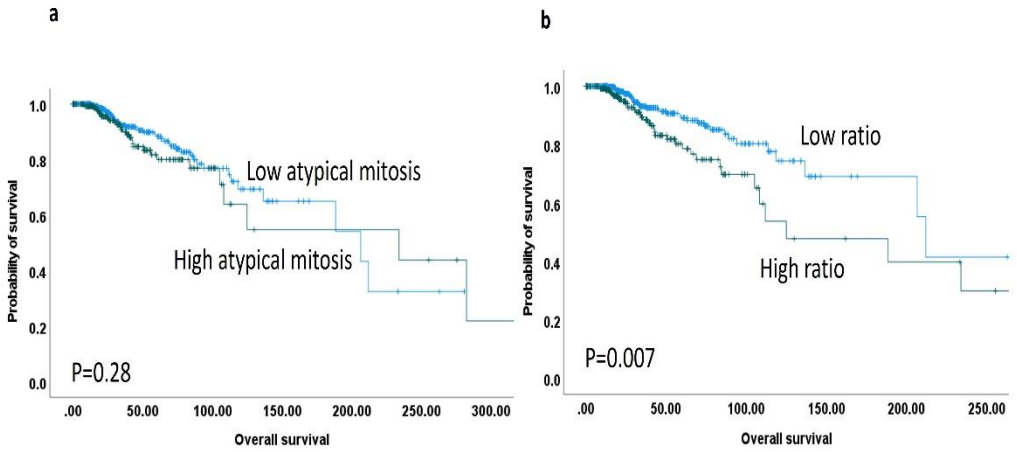


Figure 4-11: Kaplan Meier curves showing a significant association between atypical-to-typical mitoses ratio and overall survival in the TCGA cohort patients (b), but no association was observed in terms of atypical mitosis (b).

Table 21: Differential gene expression linked with the different mitotic scores.

parameters	Total genes	Upregulated genes	Downregulated genes
Atypical mitoses	3249	2495	753
Typical mitoses	2816	1864	952
Overall mitoses	3224	2307	717

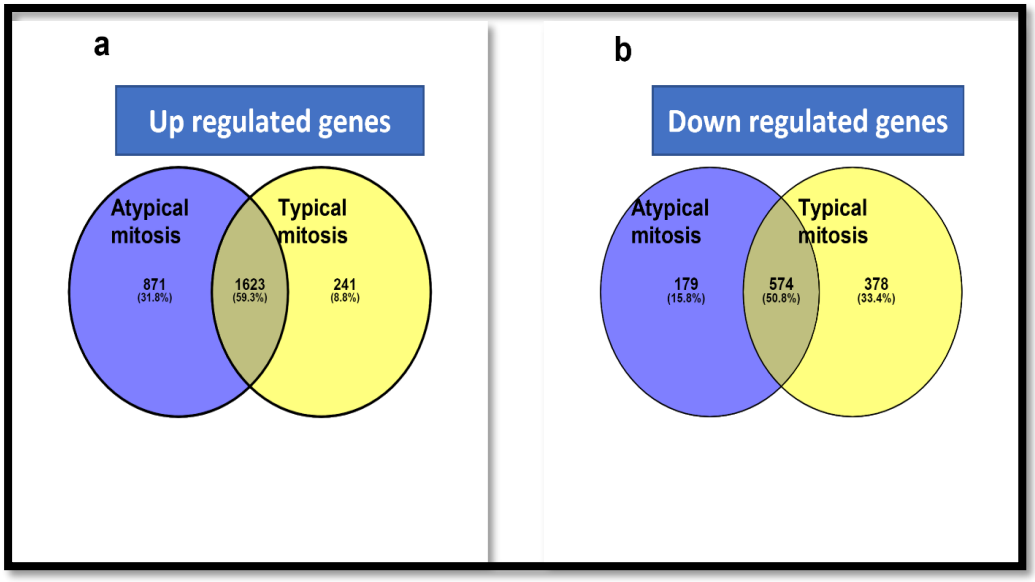


Figure 4-12: Venn diagram of obtained DEG

(a) the overlap between the up-regulated DEGs associated with atypical and typical mitoses while (b) the overlap between the down-regulated DEGs associated atypical and typical mitoses.

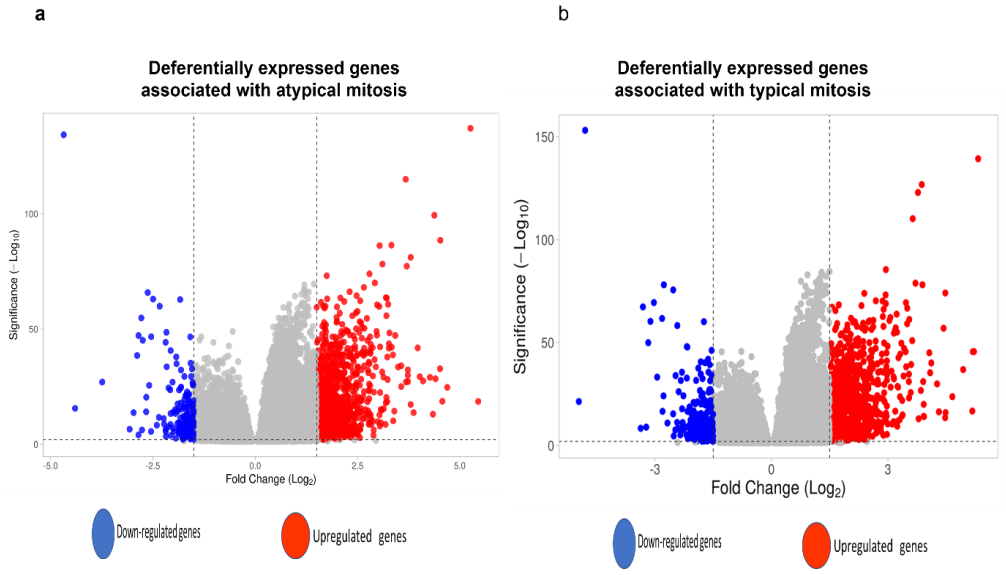


Figure 4-13: (a) Volcano plot obtained from DEG analysis

It compares the genes associated with high versus low atypical mitosis in TCGA cohort while (b) compares genes associated with high versus low typical mitosis in the same cohort.

Table 22: The Biological process that are significantly associated with high atypical mitoses.

Gene Set	Description	Size/ Overlap	Adjusted P value	Gene symbol
GO:0033044	Regulation of chromosome organization	18/18	**0.0001**	HIST1H1A-PADI2- AURKB-CDC20- TTK-PLK1
GO:0006302	Double-strand break repair	9/8	**0.001**	TRIP13- CDC45-CDCA5- EXO1-RAD54L- RPA4-AUNIP
GO:0048285	Organelle fission	41/32	**0.004**	UBE2C-TTK-KIF2C- ASZ1-CEP55-TPX2-CDCA8

GO:0050000	Chromosome localization	9/9	**0.004**	KIF2C- CEP55- FAM83D- NDC80 CDCA8- KIF14- TERB2
GO:0007059	Chromosome segregation	30/25	**0.003**	AURKB- CDC20- BIRC5- SKA1- TTK- FAM83D- KIF2C- PLK1
GO:0045930	Regulation of cell cycle	14/13	**0.01**	AURKB- CDC20- TTK-TICRR- STK33- AURKA- NDC80
GO:0044772	Mitotic cell cycle phase transition	35/27	**0.02**	MCM10- AURKB- CDC20- UBE2C- FOXM1- TTK- MELK- TICRR- ORC6- CDC25A
GO:0010721	Cell cycle checkpoint	19/18	**0.04**	BCL11A-DLL3- EPHA7-GAL- ISL2-KLK8-LHX2- LIN28A-MELTF- MT3

For the Gene Set Enrichment Analysis (GSEA) method, the category size is determined by the number of annotated genes in the category and the reference gene list. The category overlap is calculated by the overlapping differentially expressed genes (fold change (≥±1) combined with adjusted p-value (<0.05).

4.4 Discussion

Chromosome segregation errors during cell division can cause genetic abnormalities in neoplastic cells and these abnormalities can be

observed morphologically as atypical mitotic figures [138]. Atypical mitotic figures are characterised by abnormalities in the mitotic spindle symmetry and abnormal sister chromatid separation. Atypical mitosis is associated aggressive tumour behaviour [400, 412, 426]. However, detailed characterisation of atypical mitoses in BC remained to be defined. In this chapter, atypical mitoses were assessed in routine histological specimens and the prognostic value of atypical mitosis was shown to be similar to or better than other costly molecular prognostic parameters [412].

In this study, the cohort was enriched with cases rich in mitotic figures including BC cases that show mitotic grading scores 2 and 3. To avoid comparing BC with high atypical mitoses against BC with low overall mitoses scores that may bias the results as the findings may represent the mitotic activity rather than the atypicality of mitosis, atypical mitoses and the atypical-to-typical mitoses ratio were assessed.

There were strong associations between high atypical mitoses score and parameters characteristic of aggressive tumour behaviour indicating that not only the overall mitotic score but also the atypical mitoses score, and notably the atypical-to-typical mitoses ratio can provide important prognostic value in BC.

In the whole cohort, atypical mitoses did not show association with patient outcome though, when the cohort was stratified into molecular sub-classes, atypical mitoses classified luminal BC subtypes into two distinct prognostic groups. Atypical-to-typical mitoses ratio showed prognostic

stratification in the whole cohort, the luminal BC class and the TNBC subtype. The overall mitotic score was not prognostic in TNBC. This can indicate that TNBC has high proliferative activity and the atypical-to-typical mitoses ratio reflects the combination of the proliferative activity and the genomic instability of these tumours. When the cohort was stratified based on adjuvant chemotherapy and the different molecular classes, a significant improvement was observed in the outcome of TNBC patients with low atypical-to-typical mitoses ratio when they are treated with chemotherapy compared to those with high ratio. In patients who did not receive chemotherapy, the outcome of those with low atypical-to-typical mitoses ratio was similar to that with those with the high ratio while when chemotherapy was given, patients with low ratio showed significantly longer survival compared to those with high ratio. Such observation was not identified when the overall mitoses was considered in TNBC patients who received chemotherapy. These findings suggested that TNBC patients with low atypical-to-typical mitoses ratio are good candidates to chemotherapy while with high-risk patients with high atypical-to-typical mitoses ratio should be offered additional therapy.

In this study, 77% of the DEG identified were higher in tumours harbouring atypical mitoses and were associated with biological mechanisms linked to chromosomal localisation and segregation, centrosome spindle and microtubule and DNA breakdown. These findings are consistent with other studies linking atypical mitoses association with genomic instability, telomere dysfunction and chromosomal abnormalities [427]. Previous studies have reported that

chromosomal instability and telomere dysfunction are the main contributors to atypical mitoses in malignancy and that atypical mitosis can be considered a morphological marker of chromosomal instability [428, 429]. Additionally, It was proposed that presence of atypical mitotic figures is caused by specific centrosomal alterations in each neoplasm, and may considered as a distinct event that is possibly involved in carcinogenesis [430]. Aneuploidy is one of the common characteristic features of cancer cells and is related to poor clinical outcomes [431]. The underlying mechanism for aneuploidy in cancer may be a defect in the mitotic process that is required to segregate duplicated chromosomes into daughter cells [432]. Atypical mitosis could provide significant morphological signature of underlying aneuploidy of tumour cells. In addition, it can be a promising potential candidate for targeted therapy.

Conclusion

BC tumours with atypical mitoses have molecular features associated with an aggressive phenotype, higher risk of metastasis and chemotherapy resistance. The use of quantitative assessment of atypical mitoses in the context of overall mitotic number may be considered a prognostic indicator in BC.

Chapter 5 Assessment and clinical value of Ki67 in BC

5.1 Introduction

Although, proliferative activity in BC is routinely evaluated by a count of mitoses [392, 433], the addition of Ki67 immunohistochemical staining is an effective method of assessing proliferation and provides extra prognostic and predictive information [200].

Because mitotic count varies with fixation and methodology, there might be some merit to recommend Ki67 as a more reliable test for proliferative rate [434, 435]. However, some international committees do not recommend its use in routine practice due to insufficient clinical evidence and lack of standardisation and assessment method reproducibility [436-438]. For example, it is still unclear whether analysis of Ki67 should be performed in hotspots, in the tumour periphery or as an average of the whole tumour section [439, 440]. Visual assessment of Ki67 is considered the standard method for Ki67 evaluation, but this approach is time- and cost-intensive. Despite the presence of several methods for Ki67 assessment , there is a consensus that any nuclear and/or mitotic spindle staining should be considered in the final score regardless the intensity or the pattern of the staining [200, 441]. Occasional cytoplasmic or cell membrane staining of Ki67 can be observed but is not included in the final Ki67 score. Staining usually has several intensities ranging from faint to strong, and various patterns including homogenous, coarse-granular, fine granular/dotted, nuclear membrane, subnuclear membrane condensations, nucleolar, and a mixture of all the aforementioned patterns [182, 442]. With the era of digital pathology, AI-assisted

interpretation helps to provide an accurate, rapid quantification of Ki67 [206, 443]. Several studies have proposed automatic analysis for Ki67 assessment and concluded that it is a potentially effective method [280, 444, 445]. Also, a high degree of consistency for Ki67 between AI counting and the gold standard (visual counting) was observed [443].

In BC, the Ki67 labelling index is an essential tool used to classify luminal (ER-positive) tumours into two molecular surrogate subclasses: luminal A and B, which have different outcomes and therapeutic approaches [214, 446, 447]. In luminal BC, there is a wide span of proliferative activity whereas Ki67 provides significant prognostic value and identify a certain group of these tumours [36, 448]. Cutoffs to distinguish “Ki67 high” from “Ki67 low” in luminal BC vary from 1% to 29% which it is one of the perceived limitations of Ki67 use in the clinical setting. So, consistent cutoff is crucial when considering the prognosis for BC patients [215]. Also, the Ki67 cutoff should be changeable according to the context of its use either diagnostic or prognostic or in the different molecular classes [36]. Therefore, efforts should be made to develop a standard methodology and accepted cutoffs that will promote our ability to properly assess this marker. Recently, digital pathology has made great progress in image acquisition and digital analysis. This makes AI comparable to visual evaluation under the microscope regarding Ki67 interpretation [443]. By comparing the difference between AI-assisted and manual interpretation, the clinical applicability of AI-assisted interpretation is going to provide a scientific theoretical basis for an accurate and individualised treatment of BC patients [378, 443].

5.1.1 Hypothesis

In this chapter, it was hypothesised that the appropriate assessment of Ki67 can improve its use in clinical practice. Also, it was hypothesised that the underlying genetic pathways associated with Ki67 could be associated with clinical value in in early-stage BC. In addition, it was hypothesised that high diversity of Ki67-positive tumour cells could add prognostic value in BC.

5.1.2 Aims

This chapter aims to 1) compare between various methods of visual Ki67 assessment considering practicality, time effectiveness, reproducibility, the prognostic and predictive value utilising a large cohort of luminal ER-positive, 2) identify the optimal scoring method that is clinically valid and applicable, 3) characterise the quantitative differences between BC cells in the cell cycle and those in mitosis 4) explore the underlying genetic pathways associated with Ki67 through assessing the biological pathways associated with *MKI67* gene and 5) compare the manual scores of Ki67 with their corresponding automated scores.

5.2 Patients and methods

5.2.1 Study cohort.

This study was conducted on a large cohort (n=3004) of BC. This cohort was divided into sub-cohorts for addressing each aim. The clinical characteristics of the study cohort were summarised in the method chapter. At least 3-4 H&E-stained sections from each case were histologically reviewed to select one representative section containing the highest tumour burden and histological grade for assessment of cellular proliferation as described in chapter 2.5.5.

5.2.2 Ki67 staining:

Full-face sections were prepared from the selected FFPE tissue blocks. The IHC staining was performed automatically using clinically validated, DAKO Cytomation EnVision+ detection system according to the standard protocols. Briefly, sections were deparaffinised, re-hydrated and microwaved in citrate buffer (pH 6.0) for the antigen retrieval. Primary antibody (anti-human Ki67 monoclonal antibody MIB1, DAKO, Denmark) was applied and incubated for 30 minutes.

5.2.3 Comparing between various methods of eyeballing Ki67 assessment

To compare between various methods of visual Ki67 assessment, a sub-cohort (n=1662) of ER positive, HER2 negative BC with LN stage 1 (negative LN) or stage 2 (tumour metastasis in 1-3 regional LNs) was utilised. Ki67 was assessed using visual examination of the stained

sections using light microscopy (ECLIPSE Ni-U; Nikon Instruments Inc, Tokyo, Japan) with 10x ocular subjective lens (eyepieces) and 40x objective lens. Invasive tumour cells were only assessed while areas of necrosis, poor section quality, inviable tissue, and positivity within ductal DCIS or inflammatory cells were excluded. Sections were evaluated, and the percentage of positivity of Ki67 expression was assessed using the following methods:1) Quantification of Ki67 positive cells per fixed number of tumour cells (n=1000 cells) within the hotspot and the percentage was calculated. A hotspot was defined as the area with the highest density of Ki67 positive tumour cells, compared to the surrounding tumour tissue areas. Hotspots were identified at 10x magnification, and then 40x magnification was used to score Ki67 [449]. 2) Average estimation of Ki67 expression within the hotspot, whereas the average percentage of Ki67 positive tumour cells, in relation to all tumour cells within the hotspot, was subjectively estimated (without cellular count). 3) Average Ki67 expression within the whole slide. Here the average estimation of the proportion of Ki67 positive cells in the whole tumour area including hot and cold spots areas was carried out. The whole slide was scanned at low power (5x or 10x) magnification and then the mean percentage of Ki67 positive was assessed at multiple high-power views covering the different areas of the tumour [200].

5.2.3.1 Assessment of intensity and pattern of Ki67 expression

In a subset of cases (n=415), the intensity of Ki67 expression was calculated and classified into strong, moderate, or weak staining compared to the intensity within the positive tissue control. The staining

pattern was classified into 1- diffuse homogenous staining of the nucleoplasm, 2- granular pattern, which refers either to stained nucleoli or granules of distinct size dispersed throughout the nucleoplasm and, 3- mixed pattern that is strongly stained granules against a diffusely positive background of a lower staining intensity (**Figure** 5-1) [184, 185]. Finally, a combined score of the various intensities and percentage of Ki67 expression was used to estimate Histo score (H-score) of Ki67 to achieve a numerical continuous score ranging between 0 and 300 [450]. Also, a Quick score (The sum of score for the proportion of stained cells and the score for the intensity of staining) was calculated to get a total score, range of 0-8 [451]. The scoring time, in minutes, was recorded for each method. In addition, multiple cutoffs that were previously described to categorise Ki67 into distinct high and low proliferating tumours, including 10% [452, 453], 14% [213], 20% [454] and 30% [209] were used to compare the performance of different scoring methods. A second experienced pathologist (Dr Michael Toss) independently scored 25% of the cohort and the level of inter-observer agreement was analysed to assess the degree of inter-observer agreement in each of the scoring methods.

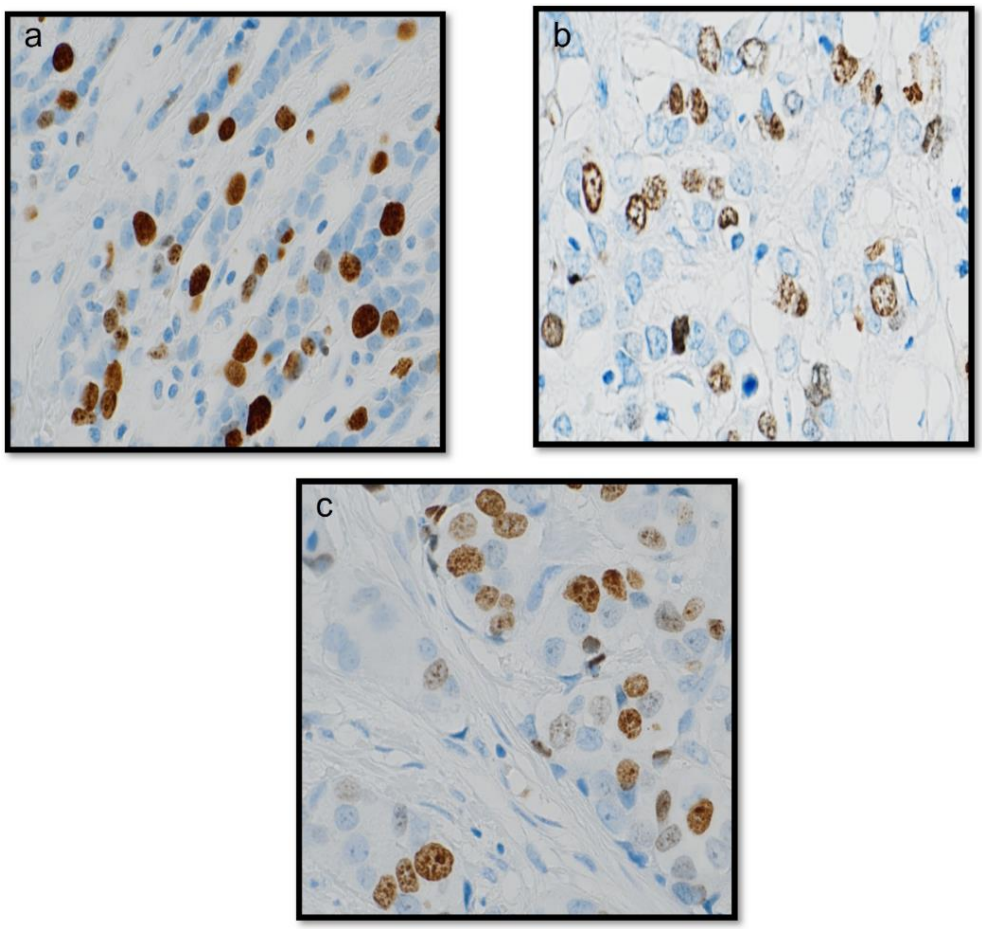


Figure 5-1: Patterns of Ki67 staining
a- homogenous staining of the nucleoplasm, b- granular pattern shows either stained nucleoli or granules of distinct size dispersed throughout the nucleoplasm, c- mixed pattern of homogenous and granular ones.

5.2.3.2 Assessment of heterogeneity in Ki67 expression

To assess the heterogeneity of Ki67 expression, the difference (Δ Ki67) in Ki67 expression between the hotspots and average methods was calculated [186, 449]. The difference was defined here as the heterogeneity index and was used to dichotomise the cohort into two groups, low vs high Ki67 heterogeneity index groups, based on the median. Ki67 heterogeneity index was correlated with clinicopathologic parameters and outcome data.

5.2.3.3 Scoring time

For each method of scoring, the scoring time was calculated and recorded in minutes whereas a timer was started at picking of the slides. Then the mean +SD of scoring time was calculated and compared in each method. 200 cases were used to assess the scoring time.

5.2.4 Evaluation of the quantitative differences between BC cells in the cell cycle and those in mitosis

To assess the quantitative differences between BC cells in the cell cycle and those in mitosis, a sub-cohort (n=621) of invasive BC specimens including all molecular classes (luminal, TN, HER2 enriched BC) was used. Full-face sections were prepared from the selected FFPE block per case. Mitotic counts in H&E stained slides were performed on FFPE tissue.

5.2.4.1 Mitosis counting

Counting mitoses was performed at 400× magnification using a light microscope equipped with 10× ocular and 40× (0.63) objective. Hotspots areas were selected based on being the most cellular area, and mostly located at the peripheral invasive front of the tumour. Fields with the highest mitotic counts were selected to commence the counts. Only cells with evident morphology of mitosis were counted [149]. Ambiguous cells with a hyperchromatic nucleus or cells suspected of apoptosis were excluded.

5.2.4.2 Evaluation of BC cells in cell cycle and in mitosis

Ki67 was assessed by three methods including 1) Ki67 in hotspots, 2) the average Ki67 in the whole slide as described earlier (BC cells in cell cycle) [200] and 3) positive- Ki67 mitotic cell which was calculated by the number of positive- Ki67-stained mitotic figures in 1000 cells in hotspot areas (BC cells in mitosis). To also assess the ability of Ki67 to stain mitosis effectively, the negative- Ki67 mitotic cells in 1000 malignant cells in the same area were quantified. The relationship between mitosis and Ki67 was assessed by comparing mitotic scores (1, 2 & 3) in H&E against Ki67 (hotspots and average).

5.2.4.3 Evaluation of the proportion of BC cells in cell cycle entering mitosis phase

To evaluate the proportion of BC cells in the cell cycle entering the mitosis phase, positive-Ki67 mitotic cells in 1000 cells in hotspot were counted and this count was converted into a percentage (percentage of mitosis in 1000 cells = number of mitoses per 1000 cells/1000x100). Consequently, if we consider that Ki67 represents the total number of BC cells in the cell cycle, the proportion of cells entering mitosis phase = (the percentage of positive-Ki67 mitotic cells in 1000 cells/percentage of positive-Ki67 in 1000 cells x 100).

For statistical purposes, the absolute mitotic counts were converted into score 1, score 2, and score 3 using the existing cutoffs according to the CAP recommendations, where, 0-11 mitoses per 10 HPFs, was equivalent to score 1, 12-22 mitosis for score 2, while score 3 was

considered when > 22 mitoses per 10 HPF [455]. To define the different proliferative activity of BC cells, the whole cohort was categorised into four groups based on positive- Ki67 mitotic count and Ki67 in hotspots: group 1=low mitosis, low Ki67, group 2= high mitosis, high Ki67, group 3= low mitosis, high Ki67, and group 4= high mitosis, low Ki67. Furthermore, the relationship between these groups and available clinicopathologic parameters was assessed. In addition, the relationship between the hotspot for Ki67 and the hotspot for mitosis was evaluated.

5.2.5 Assessment of the best cutoff for Ki67 in luminal BC.

To determine the optimal cutoff for Ki67 in luminal BC, the whole cohort (n=3004) was used. Several cutoffs of Ki67 were compared in terms of patient's outcome including the common cutoffs (10%, 20%, 14% and 30) that were used in literature. In addition, X tile software and ROC curve were used to define the best cutoff according to BCSS and DMFS.

5.2.6 Assessment the genetic pathways associated with Ki67.

To assess the DEGs associated with Ki67, TCGA-BRCA cohort (cBioPortal.org) [23] was used and the DEGs linked to *MKI67* were explored. RNA sequencing was carried out with the Illumina HiSeq platform, and the data was processed using previously described methods [456]. The resulting sequencing reads were aligned to the human hg19 genome assembly using MapSlice [457]. RSEM was used to quantify gene expression for the transcript models corresponding to the TCGA GAF 2.13 and was normalised within samples to a fixed upper quartile. DESeq2 R statistical tool [458] was used for differential analysis

of count data, using shrinkage estimation for dispersions and fold changes to improve stability and interpretability of estimates. Significant differentially expressed genes were identified based on log2 fold change (≥±1) combined with adjusted p-value (<0.05). The web-based gene set enrichment analysis tool (WebGestalt) was used to explore significantly enriched pathways and GO based on the identified DEGs [394]. The common DEGs were identified between low and high *MKI67* using the Venny version 2.0 tool [425]. Prof Nigel Mongan did the bioinformatic analysis.

5.2.7 Assessment of automated Ki67 scores in BC

AI-based tools were used to quantify Ki67 using a large subset of luminal BC patients (n=2233). A total of 1393 was used to train the algorithm. A DL model was trained to classify cells in the input images into four types: Ki67 positive tumour cells, Ki67 negative tumour cells, Ki67 positive non-tumour cells and Ki67 negative non-tumour cells. A large number of cases were used to train the algorithm, whereas regional and cellular Ki67 annotation were performed, as described in detail in chapter 2.6.1.1.2. Then, the tissue regions were divided into small regions (patches) of size 252×252 pixels at 40× magnification, and each patch was processed independently by the DL model to identify relevant cells. The proposed simultaneous cell detection and classification network (Multi-MapDe) that used in this study is a significant enhancement of a previously proposed cell detection network (MapDe) [459] and this network enables the model to detect all cell types simultaneously.

5.2.7.1 Ki67 digital (automated) score

The percentage Ki67 expression in invasive BC after masking all areas of DCIS was calculated using the DL- model, which predicted tumour cells with Ki67 positivity and final score for each case was calculated.

5.2.7.2 Ki67 co-localisation (Ki67CL score)

Ki67CL score was generated to measure the spatial co-localisation (degree of overlap) between Ki67 positive tumour and Ki67 negative tumour cells [460]. The Ki67CL score was calculated based on spatial clustering of tumour cells in the WSI regions (**Figure 5-2).**

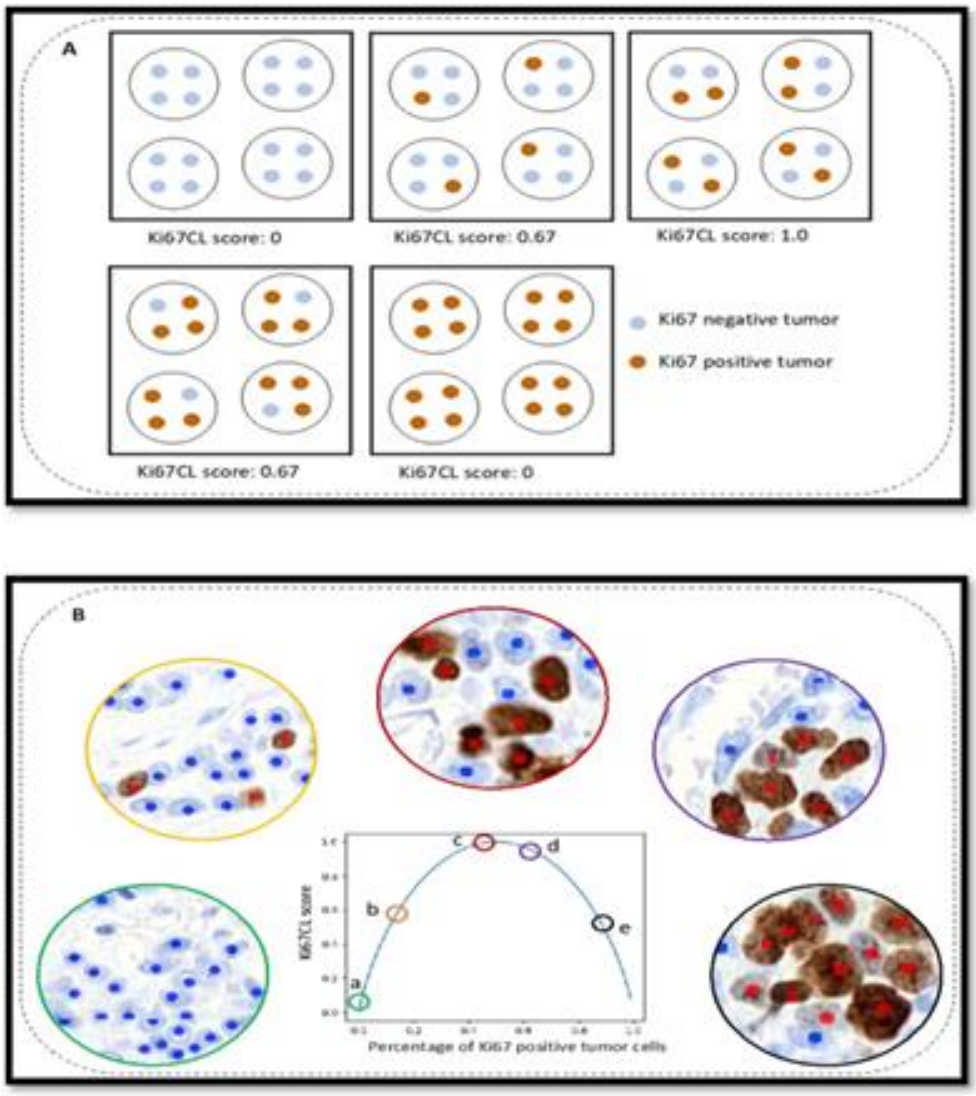


Figure 5-2: Illustration of Ki67CL score
A- each black square represents one WSI and each black circle represents one cluster. B- Ki67CL score varies when percentage of Ki67 positive tumour cells changes from 0 to 100%.

5.2.8 Statistical analysis

Statistical software 27.0 (SPSS, Chicago, IL, USA) was used for statistical analysis. Chi-square was used for analysis of categorical data. Mann-Whitney test was used for assessing non-parametric data. Ki67 expression levels were categorised based on the median [215, 461] which were 12%, 10%, 10% in quantification method, average estimation within hotspot and the average in whole slide methods, respectively. For statistical purposes, the pattern of Ki67 expression was categorised into granular and other types. Also, the H-score of Ki67 expression was categorised into low and high groups based on median (H-score 50) while Quick score was divided into a low proliferation group ≤ 4 and high proliferation group > 4. Inter-class correlation coefficient test was used to assess concordance among observers. For statistical analysis, the percentage of positive Ki67 cells was also categorised as low (<10%), intermediate (10% to 20%), or high (>20%) [462]. For statistical analysis, the proportion of cells in the cell cycle entering the mitosis phase was categorised based on the median, into low and high groups, and were correlated with the available clinicopathological parameters. For all tests, $P < 0.05$ (two-tailed) was considered statistically significant. The Kaplan–Meier survival method and the log-rank test were used for survival curves. Cox regression models was used for the multivariate analysis. Estimated HR and their 95% confidence interval (95% CI) were calculated. For all tests, $P < 0.05$ (two-tailed) were statistically significant.

5.3 Results

5.3.1 Comparing methods of eyeballing Ki67 assessment

5.3.1.1 Inter-observer agreement of Ki67 expression assessment

There was consistent agreement between observers regarding Ki67 assessment methods with Kappa values of 0.9 and 0.8 obtained for Ki67 quantification and average estimation in the hotspot, respectively. Evaluation of average Ki67 scoring per the whole section had a Kappa value of 0.7. There was also a strong positive correlation between Ki67 expression using the quantification method and the average estimation method within hotspot ($r=0.8$, $p <0.0001$) and between quantification method and the average in the whole slide ($r=0.9$, $p <0.0001$).

5.3.1.2 Distribution of Ki67 expression in BC with the different scoring methods

The mean of Ki67 percentage using quantification method per 1000 cells was 16% (range 1-98%) while the mean of Ki67 scores with estimation of the average within hotspot was 14% (range 1-95%). Ki67 proportion in average of the whole slide showed a mean 12% (range 1-95%). **Figure 5-3** shows examples of low and high expression levels of Ki67 in luminal BC. When Ki67 was dichotomised using different cutoffs, the percentage of high proliferating tumours was higher in the quantification method than both average methods. For example, using 14% as a cutoff, there was 41% of cases had high proliferation rate using the quantification method compared to 33% of cases in the average evaluation of Ki67 in the whole slide. Higher percentages of the tumours that showed high proliferation

index were consistently noticed in the counting of Ki67 positive cells/1000 tumour cells using different cutoffs (**Table 23**). 82% of the cohort showed homogeneous and mixed pattern of Ki67 staining while 18% showed pure granular pattern. Regarding the Ki67 intensity, 38% of the cases showed strong expression whereas 62% showed moderate or weak expressions. When the percentage of Ki67 expression was combined with the intensity of staining, low H-score ≤ 50 was observed in 87% of cases. 36% of the study cases showed low Quick score.

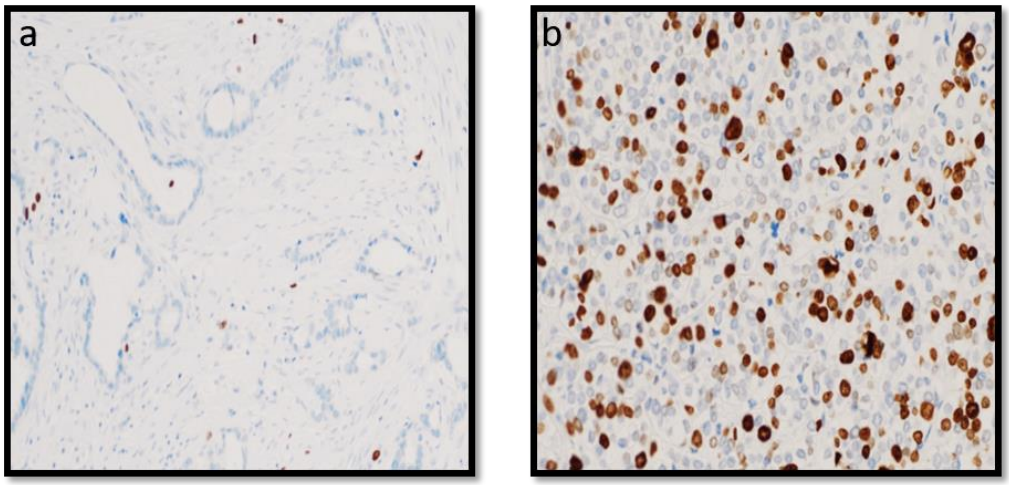


Figure 5-3: a) BC case showing very low Ki67 expression while b) another case showing high Ki67 expression.

Table 23: **Low versus high Ki67 expression groups with different methods of scoring using multiple cutoffs.**

Scoring methods	Cutoffs			
	10% (n, %)	**14%** (n, %)	**20%** (n, %)	**30%** (n, %)
Counting per 1000 cells				
Low	713 (45)	926 (59)	1196 (76)	1416 (90)
High	870 (55)	657 (41)	387 (24)	167 (10)
Average estimation within the hotspot				
Low	900 (57)	1079 (68)	1312 (82)	1452 (92)
High	683 (43)	504 (32)	271 (18)	131 (8)
Average estimation in whole slide				
Low	835 (53)	1052 (67)	1314 (83)	491(94)
High	748 (47)	531 (33)	269 (17)	92 (6)

5.3.1.3 Scoring time

The mean ± SD of scoring time for quantification 1000 cells method of Ki67 was 6.6 ± 1.2 minutes (median = 7 minutes, range 5-12 minutes). At the beginning of the study, quantification of 1000 cells took around 12 minutes, then after getting more experience in counting, the time was decreased gradually. The mean ± SD of time for average estimation in the hotspot method were 2.0 ± 0.7 minutes while subjective estimation in the whole slide was 1.2 ± 0.6 minutes. The difference between the time of Ki67 scoring with the quantification method and the average estimation in hotspot method was statistically significant (p=0.002).

5.3.1.4 Assessment of Ki67 in clinical routine practice

In the clinical setting, such counting method can be considered as time-consuming. Therefore, the tumour cell density and the number of cells per specific areas were considered in tumours with variable cellularity aiming to reduce the time needed for the actual cell count of every case.

To test the reliability of this simplified method, 100 cases were scored after adjustment of the range of cells per HPF in tumours with variable cells density. This reduced the time to 1-3 minutes per case and showed high concordance with the actual count method (K=0.9).

In this method, pathologists need to choose hotspot area at 10x, after calibration and training, pathologists can calibrate their HPF and estimate the number of cells per HPF in tumours with different cell density and assess the approximate number of HPFs that are required to count 1000 cells per case with actual cell count in every case. Then they can visually estimate the average percentage of Ki67 positive cells in the 1000 cells without actual count of the cells (**Figure 5-4**). This method would be more objective with standardisation of the 1000 cells. It will also be time-saving task.

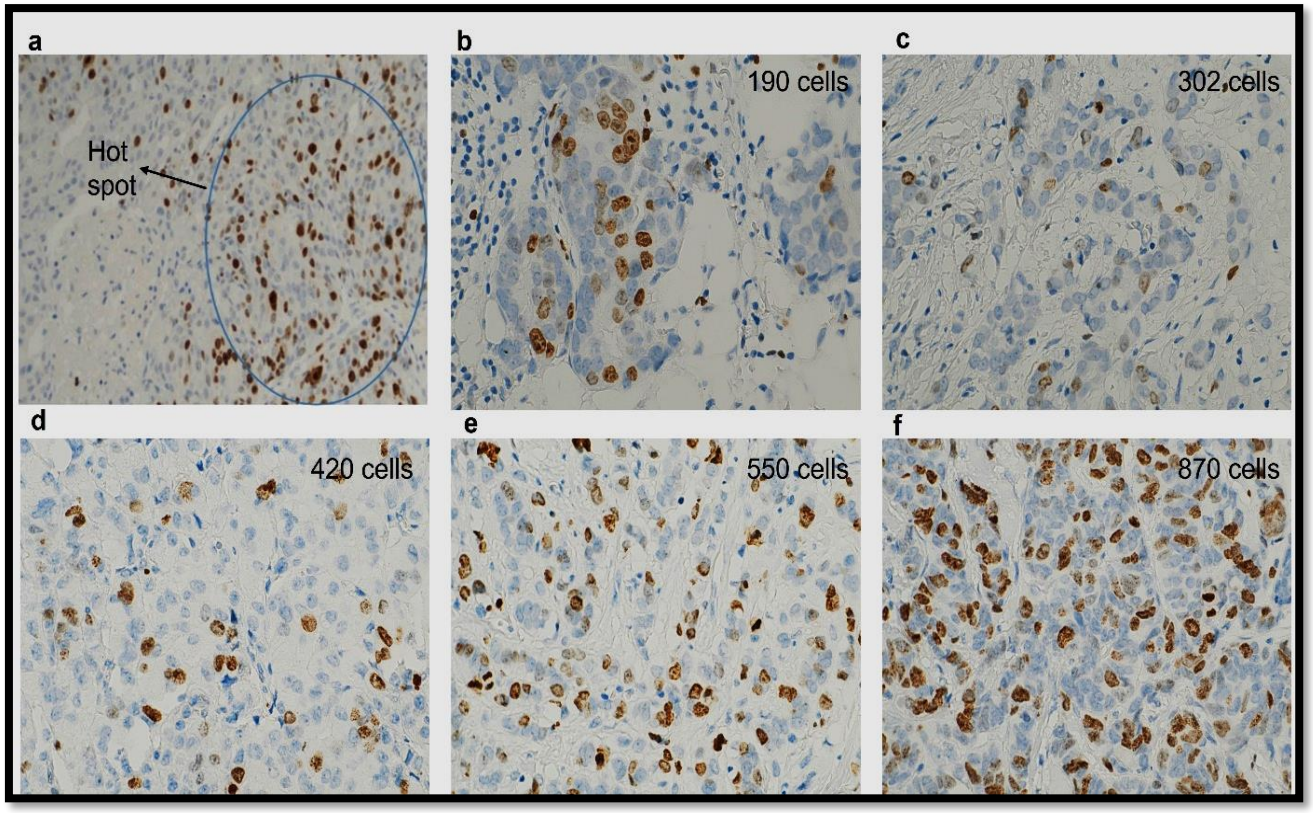


Figure 5-4: Assessment of Ki67 in clinical practice
a- choosing hotspot area at 10x. (b-f) examples of different tumours densities, which affect the number of fields to count 1000 cells either in 1 HPF in very high-density tumour or in more than 5 HPFs in low density tumour.

5.3.1.5 Correlation between the Ki67 and clinicopathological parameters in luminal BC

There was a significant association between high Ki67 expression using quantification of cells in hotspot and younger age at diagnosis, larger tumour size, grade 3 tumours and poor NPI (all, p <0.0001). Similar findings were observed regarding Ki67 scores using the estimation of Ki67 average within hotspot and the average in whole slide **(Table 24).**

Furthermore, granular pattern of staining showed significant associations with high tumour grade, NST tumour type, LVI and poor NPI (p=0.009). High quick score was associated with high grade and NST tumour type. Additionally, high Ki67 H-score showed strong associations with larger tumour size, grade 3 tumours, NST tumour type LVI, LN metastasis and poor NPI (all, p <0.0001).

Table 24: The relationship between Ki67 expression in luminal BC assessed by various methods and clinicopathological parameters

Variables	Number (%)	Objective counting in hotspot		X^2 p-value	Subjective estimation within hotspot		X^2 p-value	Subjective estimation in whole slide		X^2 p-value
		Low (n, %)	High (n, %)		Low (n, %)	High (n, %)		Low (n, %)	High (n, %)	
Age at diagnosis (years)										
<50	390 (25)	189 (49)	201 (51)	21.5	172 (44)	218 (56)	34.3	164 (42)	226 (58)	23.8
≥50	1193 (75)	737 (62)	456 (38)	**<0.0001**	728 (61)	465 (39)	**<0.0001**	671 (56)	522 (44)	**<0.0001**
Menopausal state										
Premenopausal	452 (29)	228 (50)	224 (50)	16.9	210 (46)	242 (54)	28.0	196 (43)	256 (57)	22.4
Post-menopausal	1131 (71)	698 (62)	433 (38)	**<0.0001**	690 (61)	441 (39)	**<0.0001**	639 (57)	433 (43)	**<0.0001**
Tumour size (cm)										
≤2	1054 (67)	676 (64)	378 (36)	41.3	663 (63)	391 (37)	47.0	613 (58)	441 (42)	37.1
>2	529 (33)	250 (47)	279 (53)	**<00001**	237 (45)	292 (55)	**<00001**	222 (42)	307 (58)	**<00001**
Tumour grade										
Grade 1	358 (23)	313 (87)	45 (13)	371.0	301 (84)	57 (16)	347.0	286 (80)	72 (20)	314
Grade 2	812 (51)	525 (65)	287 (35)	**<0.0001**	515 (66)	297 (34)	**<0.0001**	474 (58)	338 (42)	**<0.0001**
Grade 3	413 (26)	88 (15)	325 (85)		84 (20)	329 (80)		75 (18)	338 (82)	
Histologic tumour types										
No special type (NST)	836 (53)	392 (47)	444 (53)	129.2	385 (46)	451 (54)		347 (42)	489 (48)	
Lobular	187 (12)	129 (69)	58 (31)	**<0.0001**	128 (68)	59 (32)	101.0	123 (66)	64 (34)	108.0
Other special types	334 (21)	273 (82)	61 (18)		254 (76)	80 (24)	**<0.0001**	243 (73)	91 (27)	**<0.0001**
NST mixed	226 (14)	132 (58)	94 (42)		133 (59)	93 (41)		122 (54)	104 (46)	
Lymph node status										
Negative	1098 (69)	679 (56)	419 (44)	16.5	660 (60)	438 (40)	15.5	611 (56)	487 (44)	12.1
Positive	485 (31)	247 (51)	238 (49)	**<0.0001**	240 (49)	245 (51)	**<0.0001**	224 (46)	261 (54)	**<0.0001**
Lymphovascular invasion										
Absent	1246 (79)	782 (63)	464 (37)	43.8	770 (62)	476 (38)	58.1	711 (47)	535 (53)	43.7
Present	337 (21)	144 (43)	193 (57)	**<0.0001**	130 (39)	207 (61)	**<0.0001**	124 (37)	213 (63)	**<0.0001**
Nottingham prognostic index										
Good prognostic group	765 (48)	585 (76)	180 (24)	215.2	573 (75)	192 (25)	213.4	527 (31)	238 (69)	178.6
Moderate prognostic group	721 (46)	320 (44)	401 (56)	**<0.0001**	307 (43)	414 (57)	**<0.0001**	294 (41)	427 (59)	**<0.0001**
Poor prognostic group	97 (6)	21 (22)	76 (78)		20 (21)	77 (79)		14 (14)	83 (86)	

5.3.1.6 Outcome analysis

In univariate analysis, high Ki67 expression showed a significant association with poor outcome in terms of BCSS and DMFS (**Figure 5-5)** using all scoring methods (all, p<0.0001). However, the proportion of Ki67 positive cells within 1000 tumour cells showed the highest prediction of poor outcome (highest HR and the lowest p value) regardless the cutoff used to define high and low proliferation index (**Table 25**).

In addition, in multivariate analysis, the hotspot quantification of Ki67 positive cells among 1000 tumour cells provided a statistically significant, prognostic value independent of the other two methods in all cutoffs applied to categorise the scores (**Table 26**). The other two methods were not significant except for the average estimation of Ki67 within the hotspot that showed weak independent association with BCSS when 10% cutoff was used.

Regarding adjuvant therapy, there was a significant association between high Ki67 and poor outcome in patients who received endocrine therapy regardless of the method of scoring (**Figure 5-6**). However, in multivariate analysis, the quantification of Ki67 positive cells among 1000 tumour cells in hotspot method showed statistically significant prognostic value independent of the other two methods.. No association was observed between Ki67 and outcome in patients treated with combined endocrine therapy and chemotherapy.

Notably, there was strong association between Ki67 granular expression pattern and shorter DMFS (p=0.009) but not BCSS (p=0.05) (**Figure 5-7**). Tumours that showed high Ki67 expression and more granular pattern

showed the worst outcome in terms of DMFS (p=0.009) compared to low Ki67 expression with homogeneous pattern (**Figure 5-8**).

Neither H-score nor Quick Ki67 scores showed association with patient outcome apart from shorter DMFS that was shown in cases with higher Ki67 H. score (**Figure 5-9).**

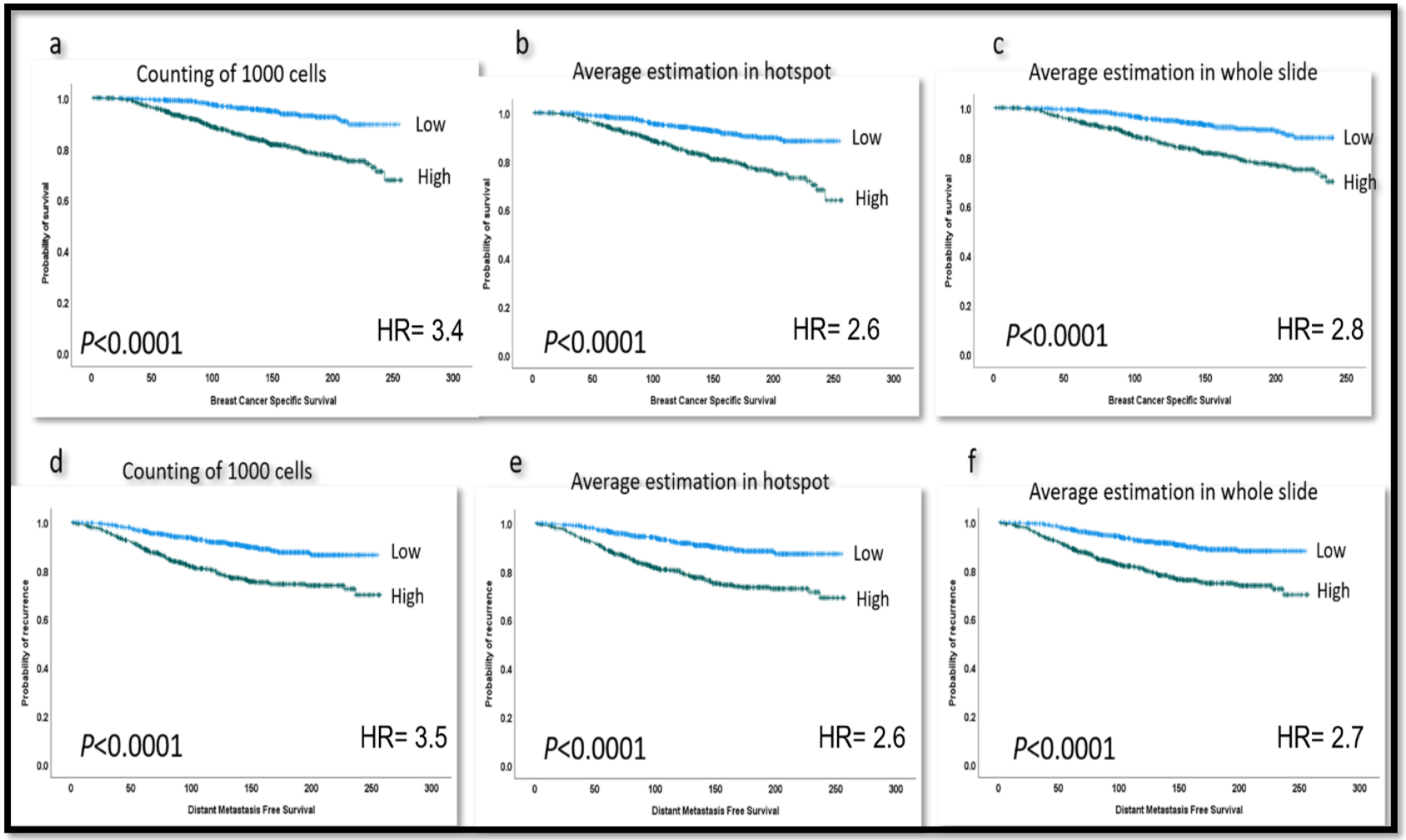


Figure 5-5: Kaplan Meier plots showing the associations of BCSS with Ki67 using all scoring methods (a-c) and the same findings were observed in terms of DMFS (d-f)

Table 25: Univariate Cox regression analysis for predictors of breast cancer-specific survival according to the different methods of Ki67 assessment with different Ki67 cutoffs

Cut-off	10%			14%			20%			30%		
Scoring Method	HR	95%CI	*p*-value	HR	95%CI	*p*-value	HR	95%CI	*p*-value	HR	95%CI	*p*-value
Counting Ki67 positive cells per 1000 invasive tumour cells	3.4	2.4-4.8	**1.2×10^{-11}**	2.8	2.1-3.7	**2.7×10^{-11}**	2.4	1.8-3.1	**7.6×10^{-9}**	2.5	1.7-3.5	**0.000001**
Average estimation within the hotspot	2.6	1.9-3.6	**6.5×10^{-10}**	2.1	1.5-2.8	**9.8×10^{-7}**	2.1	1.6-2.9	**0.000003**	2.3	1.5-3.3	**0.00006**
Average estimation in whole slide	2.8	2.0-3.8	**1.6×10^{-10}**	2.7	2.0-3.6	**1.7×10^{-11}**	2.2	1.6-2.9	**0.000001**	1.9	1.2-3.1	**0.007**

HR, hazard ratio; CI, confidence interval. Significant *p* values are in bold.

Table 26: Multivariate Cox regression analysis for predictors of breast cancer-specific survival according to the different methods of Ki67 assessment with different Ki67 cutoffs

Cutoff	**10%**			**14%**			**20%**			**30%**		
Scoring Method	Hazard ratio	95% CI	P-value	Hazard ratio	95% CI	P-value	Hazard ratio	95% CI	P-value	Hazard ratio	95% CI	P-value
Counting Ki67 per 1000 invasive tumour cells	2.80	1.500-5.20	**0.002**	1.8	0.90-3.20	0.06	1.8	1.1-3.1	**0.02**	2.5	1.4-4.3	**0.001**
Average estimation within the hotspot	1.40	0.97-2.20	0.073	1.2	0.76-1.70	0.56	1.3	0.83-2.0	0.30	1.4	0.8-2.5	0.300
Average estimation in whole slide	0.93	0.50-1.60	0.059	1.3	0.75-2.40	0.19	1.8	0.6-1.7	0.97	0.5	0.3-1.1	0.085

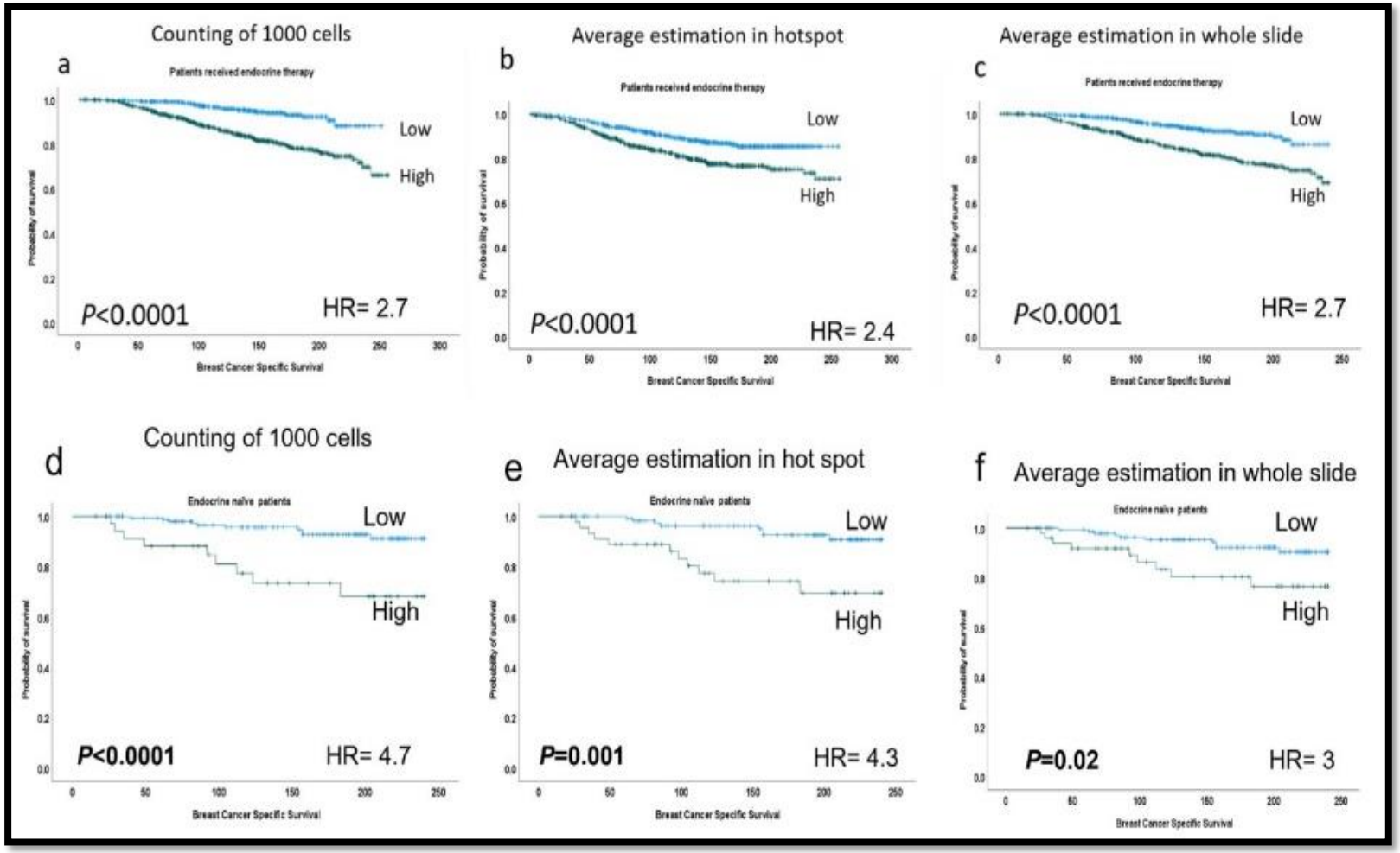


Figure 5-6: Kaplan Meier curves show the association of Ki67 with different scoring methods
Association with BCSS in luminal BC patients received endocrine therapy using 10% cutoff (a-c). Similar findings were observed in endocrine therapy naïve patients using the same cutoff (d-f).

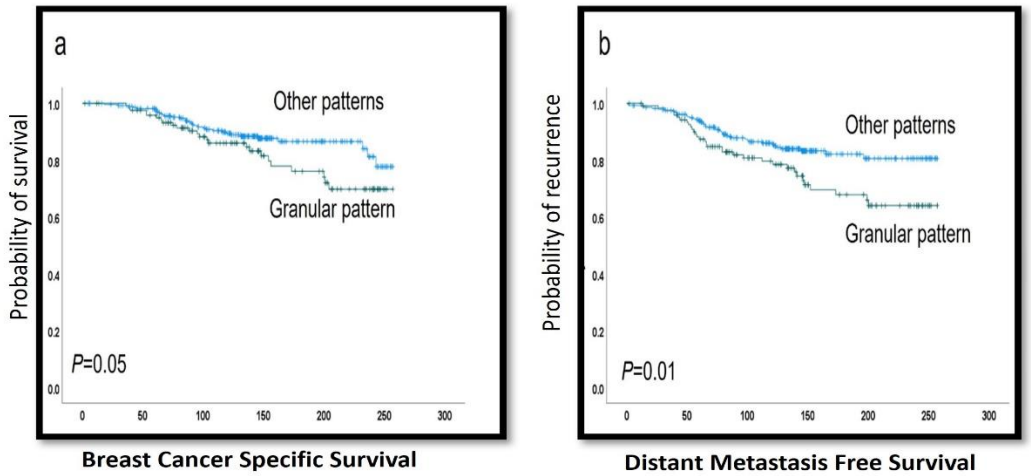


Figure 5-7: Kaplan Meier curves show the association of Ki67 granularity and outcome
No association was observed between granularity pattern of Ki67 expression in luminal BC in terms of BCSS (a) while a significant association was observed in terms of DMFS (b)

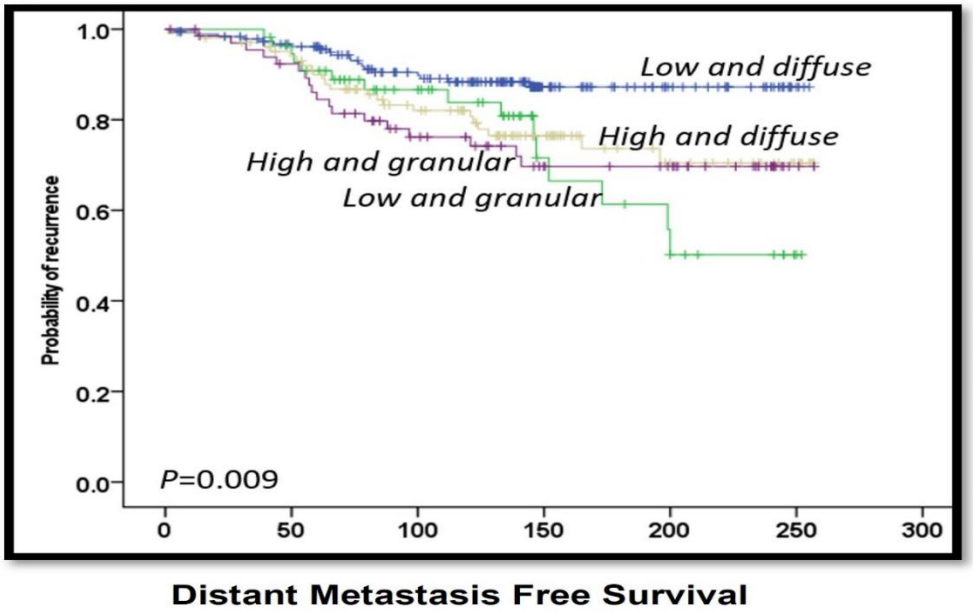


Figure 5-8: Kaplan Meier curve shows the association of high Ki67 expression with granular pattern and shorter DMFS in luminal BC.

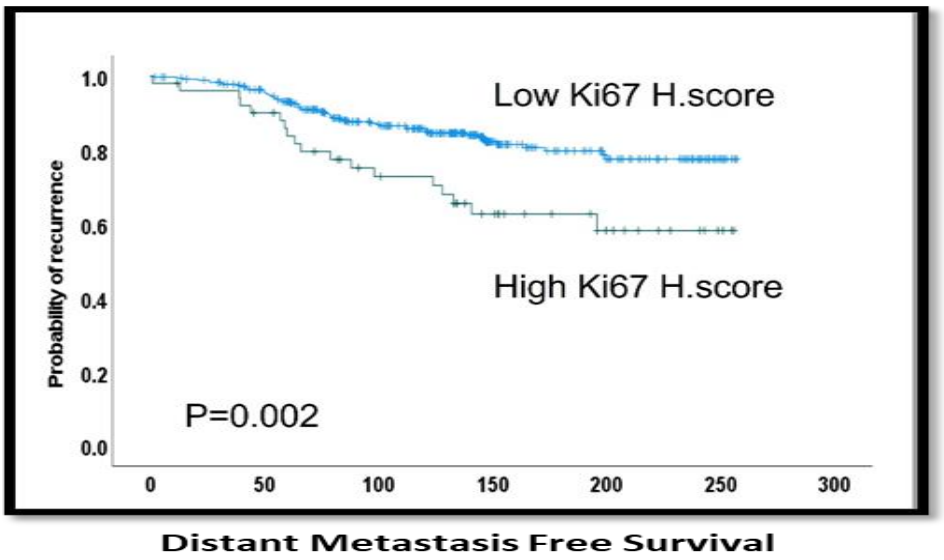


Figure 5-9: Kaplan Meier curve shows a significant association between high Ki67 H. score and short DMFS**.**

5.3.2 Clinical impact of Ki67 intra-tumoural heterogeneity

The mean heterogeneity index between hotspot and whole slide was 2.8% (range 0-66%). There was a significant association between high Ki67 heterogeneity index and larger tumour size, higher tumour grade, NST tumour type, presence of both LVI and LN metastasis, and poor NPI (**Table 27**). Tumours with higher Ki67 expression had significantly higher heterogeneity values ($p<0.001$) (**Figure 5-10**). Additionally, high heterogeneity index showed strong association with shorter BCSS and DMFS (both $p <0.0001$) in the whole cohort and when the analysis was confined to patients who received endocrine therapy (**Figure 5-11**).

Table 27: Relationship between Ki67 heterogeneity index and clinicopathological parameters

Variables	Heterogeneity index		X^2
	Low (n, %)	High (n, %)	P-value
Age (years)			
<50	192(49)	198 (51)	6.4
≥50	675 (57)	518 (43)	**0.01**
Menopausal state			
Premenopausal	235 (52)	217 (48)	1.97
Post-menopausal	632 (56)	499 (44)	0.16
Tumour size (cm)			
≤2	623 (59)	431 (41)	23.9
>2	244 (46)	285 (54)	**<0.0001**
Tumour grade			
Grade 1	263 (73)	95 (27)	117.7
Grade 2	460 (57)	352 (43)	**<0.0001**
Grade 3	144 (35)	269 (65)	
Histologic tumour types			
No special type (NST)	411 (49)	425 (51)	32.2
Lobular	104 (56)	83 (44)	**<0.0001**
Other special types	225 (67)	109 (33)	
NST mixed	127 (56)	99 (44)	
Lymph node status			
Negative	623 (57)	475 (43)	5.6
Positive	244 (50)	241 (50)	**0.01**
Lymphovascular invasion			
Absent	714 (57)	532 (43)	15.2
Present	153 (45)	184 (55)	**<0.0001**
Nottingham prognostic index			
Good prognostic group	500 (65)	265 (35)	72.3
Moderate prognostic group	334 (46)	387 (54)	**<0.0001**
Poor prognostic group	33 (34)	64 (66)	

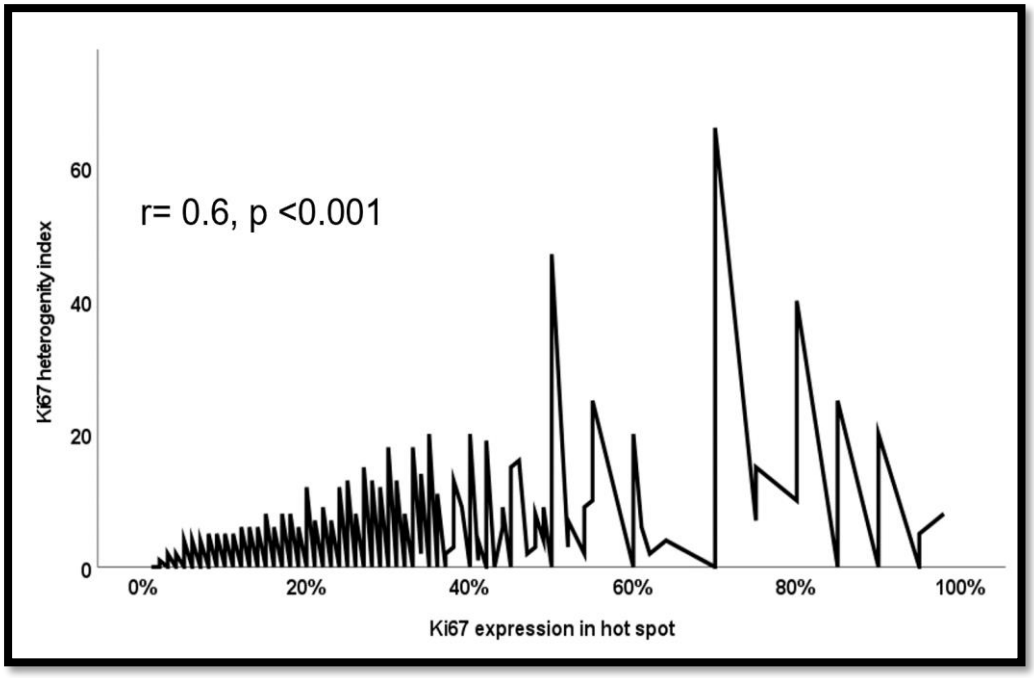


Figure 5-10: Line graph shows the correlation between Ki67 heterogeneity index and the Ki67 expression levels

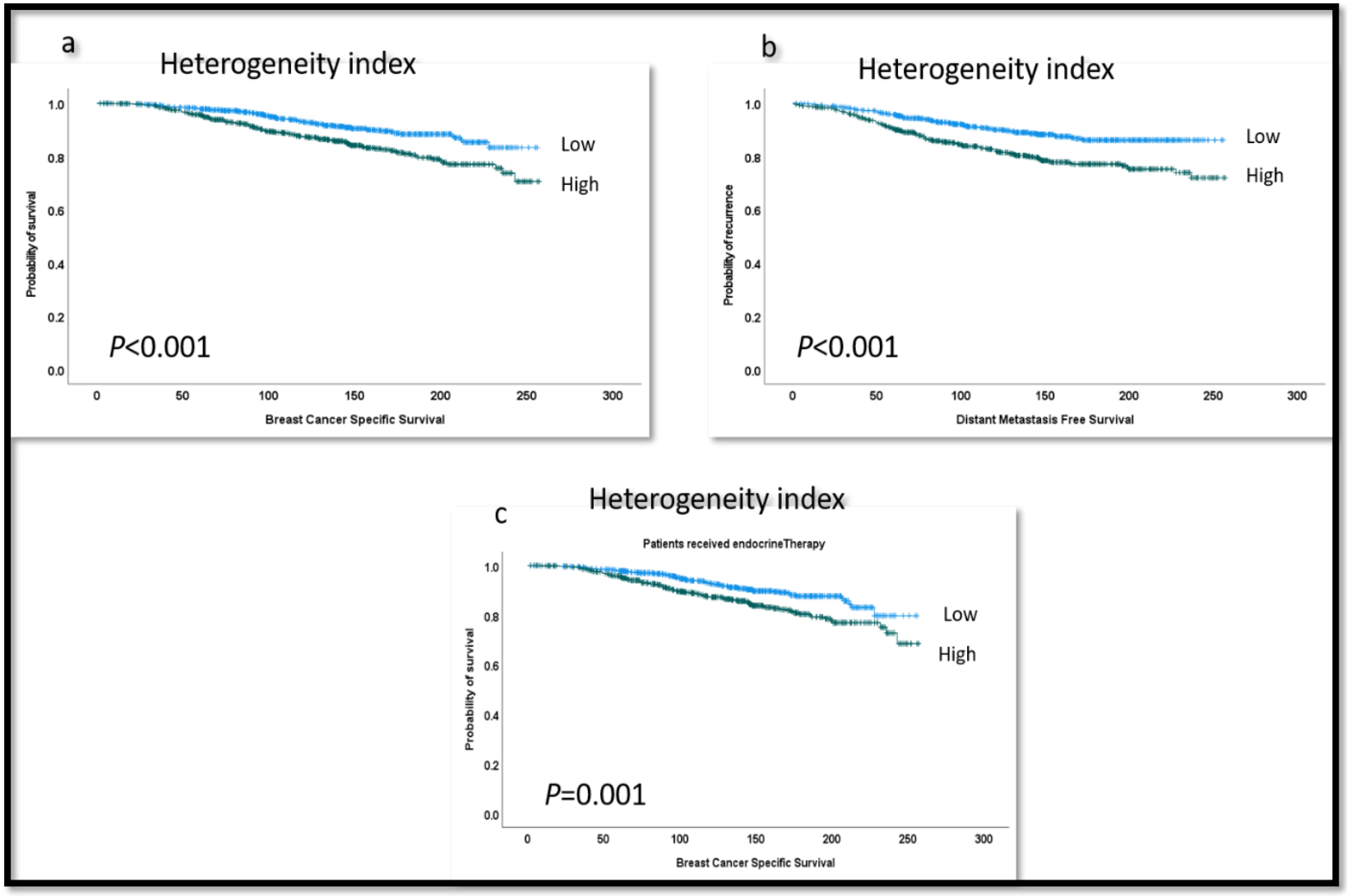


Figure 5-11: Kaplan Meier plots show significant associations between Ki67 heterogeneity index and shorter BCSS (a), DMFS (b) in the whole cohort and BCSS in patients who received endocrine therapy (c).

5.3.3 The quantitative differences between BC cells in cell cycle and in mitosis

The mean and standard deviation (SD) of BC cells in the cell cycle, (Ki67-positive cells in the whole invasive tumour) were 24% ± 22%, the median was 17%, (range <1% to 95%), whereas the mean and SD in the hotspots were 31% ± 27%, the median was 21%, (range <1% to 100%).

Of the 621 informative cases assessed, 50% showed high proliferative activity, with a high proportion of cells in the cell cycle when hotspots were assessed. However, when using the average method (in the whole slide), this proportion dropped to 40%. **Table 28** shows the distribution of cycling cells and the cells in mitosis.

There was a strong association between cycling cells and cells in the mitosis regardless of the area of assessment (in hotspots or the average) ($p<0.0001$) (**Table 29**, **Figure 5-12 and Figure** 5-13).

The mean and SD of Ki67-positive mitotic cells was 12.99±13.4 mitoses, (range 0-65 mitoses) (**Figure 5-14** and **Figure 5-15**).

Table 28: Distribution of Ki67 and mitotic scores

Categories	Number	%
Ki67 expression in hotspot		
Low	170	27%
Intermediate	138	23%
High	313	50%
Ki67 expression in average		
Low	218	35%
Intermediate	157	25%
High	246	40%
Ki67- positive mitotic cells		
Low	316	51%
High	305	49%
Mitotic counts in H&E		
Score 1	264	42%
Score 2	123	20%
Score 3	234	38%

Ki67-positive cells: low (<10%), intermediate (10 to 20%); high (>20%). Mitotic scores (score 1: 0-11 mitoses; score 2: 12-22 mitoses; score 3: >22 mitoses). Ki67 positive mitotic cells: low (<10 mitoses), high (>10 mitoses).

Table 29: Relationship between mitotic scores and scores of cycling (Ki67-positive) cells (low, intermediate, high)

Variables	Mitosis scores			X^2 P-value
Proportion of cycling cells	**Score 1** n (%)	**Score 2** n (%)	**Score 3** n (%)	
In hotspots				
Low	133 (78)	19 (11)	18 (11)	215.8
Intermediate	80 (58)	35 (25)	23 (17)	**<0.0001**
High	51 (16)	69 (22)	193 (62)	
Average				
Low	17 (78)	25 (12)	22 (10)	267.2
Intermediate	67 (43)	50 (32)	37 (25)	**<0.0001**
High	26 (11)	48 (20)	172 (69)	

Significant p values are in bold.

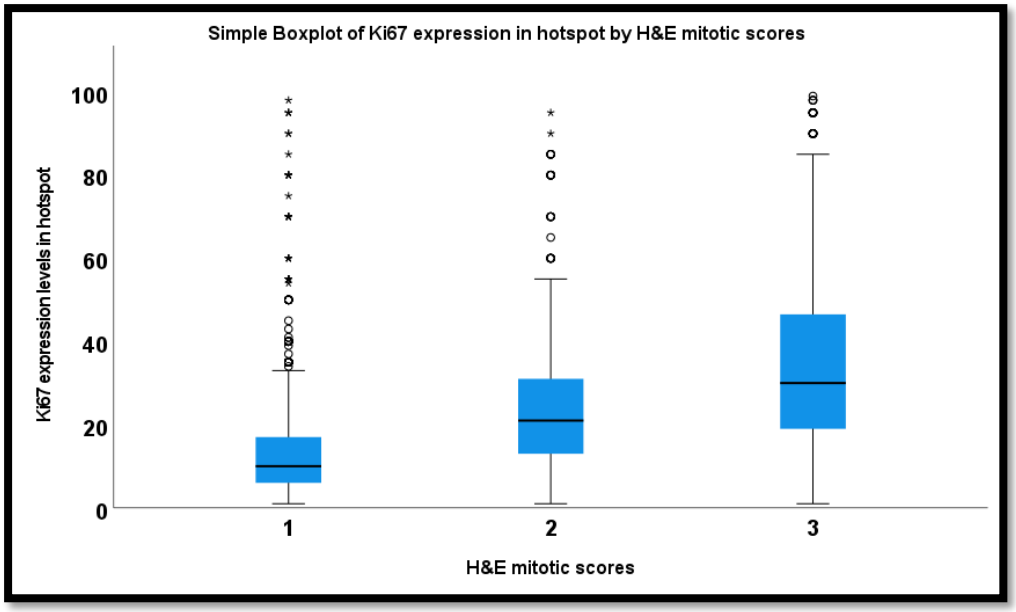


Figure 5-12: Box plot shows the relationship between mitotic scores in H&E and Ki67 positive cells.

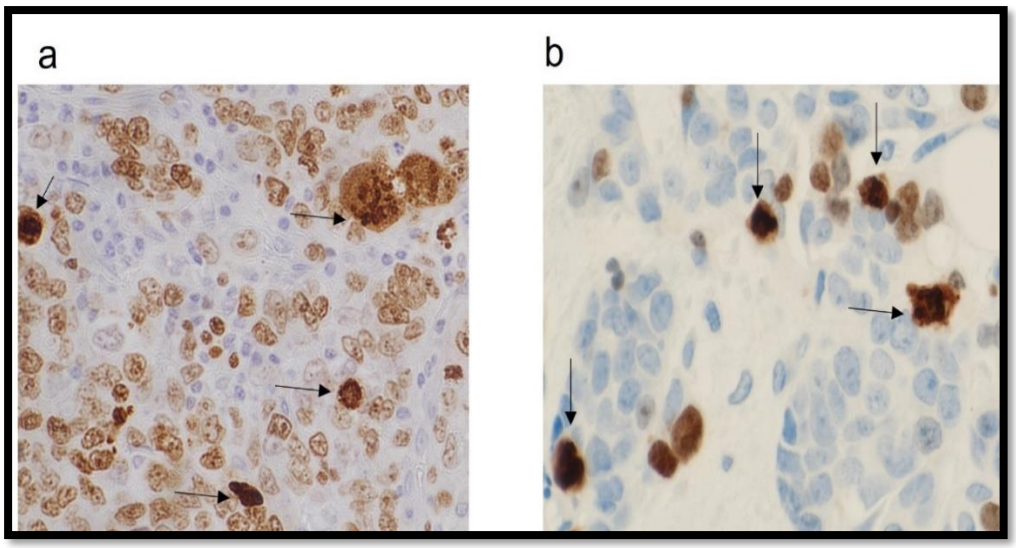


Figure 5-13: Relationship between hotspot of Ki67 and mitosis (40x)
a- overlapped hotspots for mitosis (black arrows) and Ki67, b- Mitosis hotspot (black arrows) but cold area for Ki67.

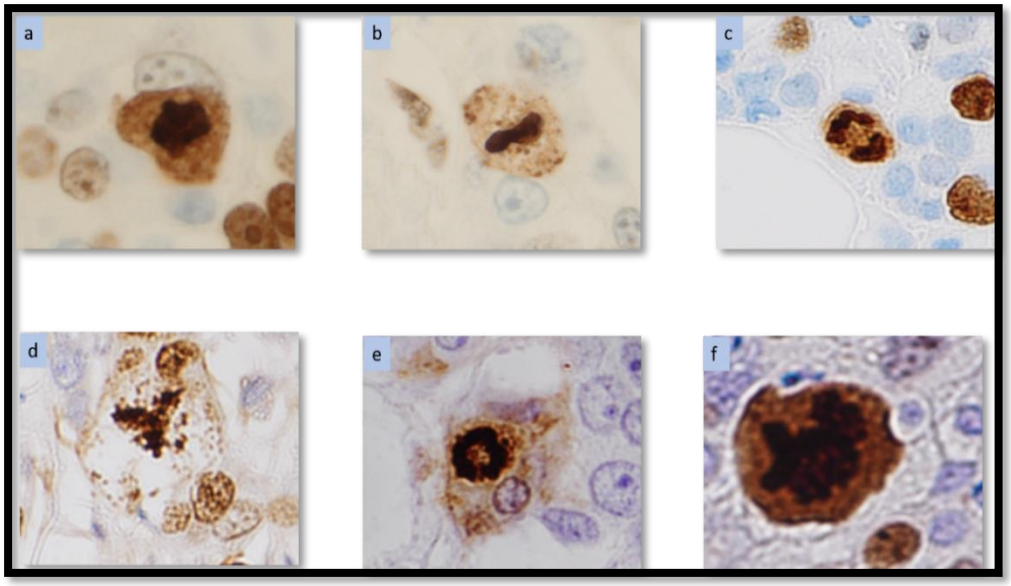


Figure 5-14: Different phases of Ki67-positive mitotic cells (40x)
a-prophase shows slightly irregular outline, b- metaphase shows rod-like structure, c- anaphase shows two uniform poles pulled away, d-tripolar atypical mitosis shows three poles, e-ring atypical mitosis f- multipolar atypical mitosis.

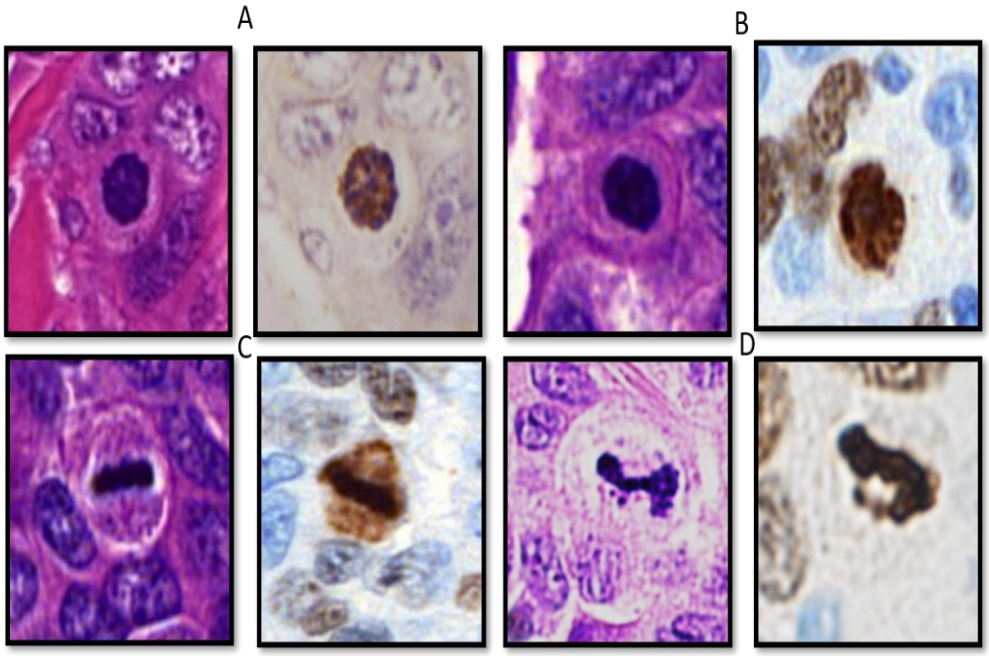


Figure 5-15: Matched mitotic figures between H&E and Ki67(40x)
A- prophase shows fine hairy-like projection, and positive for Ki67. B-prophase shows irregular outline, and positive for Ki67. C- metaphase shows fine hairy projection and positive for Ki67. D- lag atypical mitosis shows not attached chromatid to metaphase and positive for Ki67.

5.3.3.1 Proportion of BC cells in cell cycle entering mitosis.

The mean (±SD) % of BC cells in the mitosis was (5.3 ±7.2%), the median was 3.7% and (range 0 to 73%). There was a significant association between mitosis proportion and tumour grade whereas 67% of grade 3 tumours showed high mitosis proportion ($p<0.0001$). There was also a significant association between low mitosis proportion and tumours with special histological types ($p<0.0001$) (**Table 30**). There was a significant association between high mitosis proportion and TNBC ($p<0.0001$) (**Figure 5-16)**. In addition, a significant association was observed between high mitosis proportion and LN metastasis ($p<0.02$), lymphovascular invasion ($p=0.004$) and higher NPI ($p<0.0001$).

The absolute number of Ki67 negative mitotic cells as defined by morphology, in 1000 cells was evaluated, and the mean count was 1 mitosis (ranged from 0 to 22 mitoses). When the absolute number was converted to a percentage, the mean percentage of Ki67-negative mitotic cells was 0.6%, (ranged 0% to 2.2%). Consequently, Ki67 can effectively stain 99% of mitotic figures, and Ki67-positive mitotic cells could represent almost all mitotic figures.

Table 30: Relationship between proportions of cycling BC cells entering mitotic phase and clinicopathological diameters.

Categories	Low mitosis proportion Number (%)	High mitosis proportion Number (%)	X2 P-value
Tumour grade			
Grade 1	60 (81)	14 (19)	98.12
Grade2	137 (68)	63 (32)	**<0.0001**
Grade 3	113 (33)	234 (67)	
Histologic types			
NST no special type	184 (43)	243 (57)	29.1
Lobular	26 (70)	11 (30)	**<0.0001**
Other special types	20 (80)	5 (20)	
Mixed tumours	80 (61)	52 (39)	
Molecular subtype			
Luminal A	102 (69)	46 (31)	49.0
Luminal B	123 (55)	100 (45)	**<0.0001**
Triple negative / basal	64 (33)	129 (67)	
HER2 enriched	21 (37)	36 (63)	
Lymph node status			
Negative	218 (53)	192 (47)	5.2
Positive	92 (44)	119 (56)	**0.02**
Lymphovascular invasion			
Absent	226 (54)	193 (46)	8.3
Present	84 (42)	118 (58)	**0.004**
Nottingham prognostic index			
Good prognostic group	115 (75)	39 (25)	63.2
Moderate prognostic group	168 (46)	196 (54)	
Poor prognostic group	27 (26)	76 (74)	**<0.0001**

Significance P values are in bold.

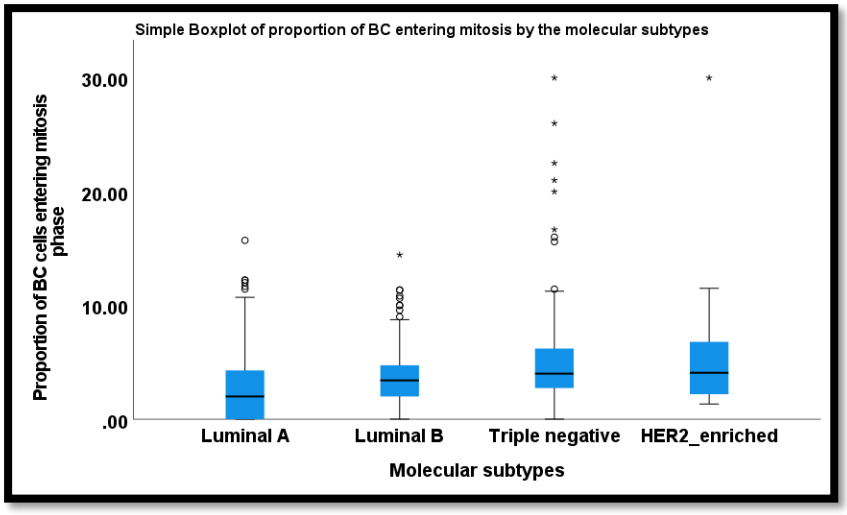


Figure 5-16: Box plot shows the relationship between proportion of BC cells entering mitosis phase and molecular subtypes.

5.3.3.2 Cell cycle and mitosis proportion groups

The whole cohort was stratified into four proliferative groups based on Ki67-positive mitotic counts and Ki67-positive cells in 1000 cells in hotspots using the same area. 42% of cases (260/621) had low proliferative activity and low mitotic activity (low mitosis/low Ki67) while 40% (251/621) had high mitosis/high Ki67. In contrast, 18% (110/621) showed discrepant results; 10% showed low mitosis / high Ki67 and 8% showed high mitosis/low Ki67.

Low mitosis/low Ki67group showed a significant association with grade 1 tumours, lobular carcinoma type, luminal A tumours and LN negativity ($p<0.001$). On the other hand, high mitosis/high Ki67 group showed a significant association with grade 3 tumours ($p<0.0001$).

Patients with discordant cell cycle and mitotic counts were assessed as separate groups and correlated with their corresponding clinicopathologic parameters. The low mitosis with high Ki67 group showed significant association with the luminal B subtype ($p<0.0001$). In contrast, high mitosis with low Ki67 group was significantly associated with grade 3 tumours, triple negative subtype, and higher NPI ($p<0.0001$) (**Table 31**).

Table 31 : Relationship between the discrepant mitosis/ Ki67 groups and clinicopathologic parameters.

Categories	Low mitosis/high Ki67 Number (%)	High mitosis/low Ki67 Number (%)	X^2 P value
Tumour grade			
Grade 1	3 (100)	0 (0)	25.6
Grade 2	35 (85)	6 (15)	**<0.0001**
Grade 3	25 (38)	41 (62)	
Histologic types			
No special type (NST)	45 (54)	39 (46)	4.01
Lobular	3 (75)	1 (25)	0.3
Other special types	0 (0)	1 (100)	
Mixed tumours	15 (71)	6 (29)	
Molecular subtype			
Luminal A	0 (0)	1 (100)	28.2
Luminal B	50 (78)	14 (22)	**<0.0001**
Triple negative / basal	9 (26)	26 (74)	
HER2 enriched	4 (40)	6 (60)	
Lymph node status			
Negative	39 (62)	24 (38)	1.3
Positive	24 (51)	23 (49)	0.3
Lymphovascular invasion			
Absent	40 (57)	30 (43)	0.001
Present	23 (58)	17 (42)	0.9
Nottingham prognostic index.			
Good prognostic group	15 (88)	2 (12)	12.49
Moderate prognostic group	41 (58)	30 (42)	**<0.0001**
Poor prognostic group	7 (32)	15 (68)	

Significant p values are in bold.

5.3.4 Assessment of the best cutoff of Ki67 expression in luminal BC.

To explore the optimal cutoff to categorise the proliferative activity in luminal BC, a comparative analysis was performed using different common cutoffs of Ki67. Using X tile, it was found that 10% of Ki67 expression is the best cutoff to stratify luminal BC patients into 2

prognostic risk groups. In addition, area under the curve showed that 10% cutoff showed the highest sensitivity and specificity (**Figure 5-17**). When multivariate analysis was considered in weighting the different methods of Ki67 with different cutoffs, 10% cutoff using counting Ki67 positive cells per 1000 invasive tumour cells in hotspot method showed independent association with poor outcome with the highest HR in terms of BCSS and DMFS.

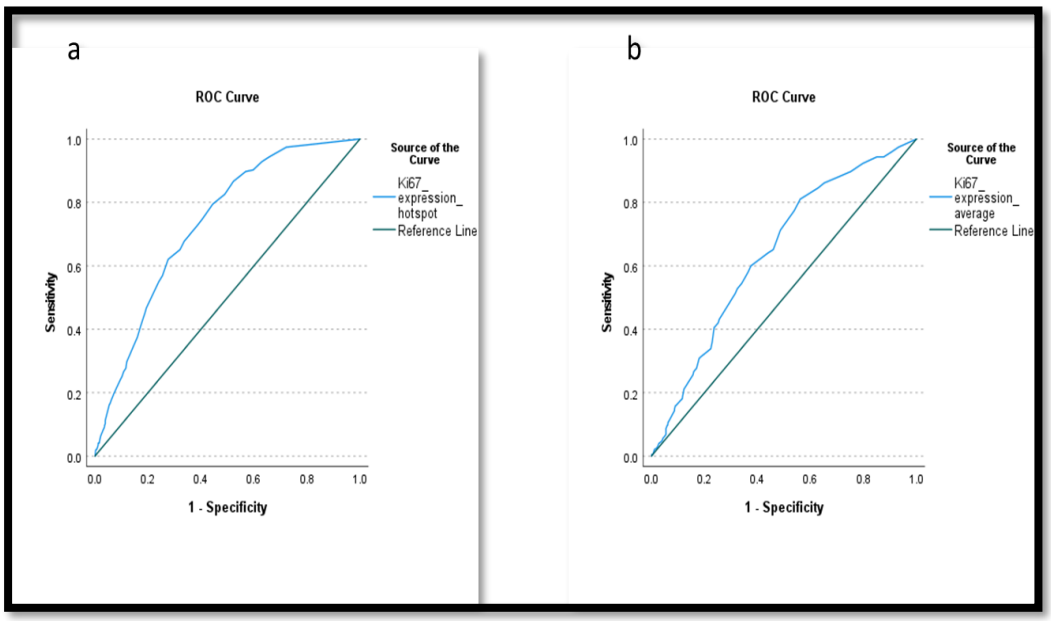


Figure 5-17: ROC curves show area under the curve of 0.7 for Ki67 expression levels in hotspot and it is 0.64 for average Ki67expression levels in the whole slide**.**

5.3.5 Genetic pathways associated with *MKI67*.

DGE analysis identified a set of genes (n = 7592) that were significantly associated with *MKI67* mRNA gene, while 5489 genes were significantly associated with high *MKI67,* 2103 genes were associated with low *MKI67*. The genes that showed higher expression in cases with high *MKI67* were involved in specific biological pathways including, regulation of cell cycle phase transition, stem cell proliferation, vascular endothelial growth factor production, neural precursor cell proliferation, chromosome segregation, DNA-binding transcription activator activity, RNA polymerase II-specific, DNA replication and damaged DNA binding (**Table 32**). On the other hand, the genes that showed higher expression in cases with low *MKI67* were associated with regulation of neurological system process, anatomical structure homeostasis, cAMP signalling pathway and PPAR signalling pathway.

Table 32: Gene Ontology shows the common pathways that are significantly associated with high *MKI67.*

Gene Set	Description	Size/ overlap	Adjusted P value	Gene symbol
hsa03030	DNA replication	14/14	**0.005**	MCM2-MCM4-RPA4-MCM7-FEN1-RFC4-DNA2
hsa04110	Cell cycle	47/30	**0.04**	Cyclin E1- TTK-BUB11-SMC1B-CCNA2-ORC1-CDDC6-PLK1
GO:0001228	DNA-binding transcription activator activity	65/35	**0.0001**	POU4F1-DMRT1-SOX11-TLX1-RFX4-MYBL2
GO:0003684	damaged DNA binding	12/12	**0.004**	NEIL3-POLQ-BLM-AUNIP-HMGB2-BRCA1-REQL4- MSH2
GO:0005874	Microtubule	59/28	**0.009**	KIF1A-FSD1-CAPN6-TPX2-BRIC5-SKA-ASPM
GO:1901987	Regulation of cell cycle phase transition	75/41	**<0.0001**	Cyclin E1- FAM83D-SIX3-E2F8-TICRR-ANLN-TTK
GO:0010573	Vascular endothelial growth factor production	7/7	**<0.0001**	IL1A - BRCA1- PTGS2-CYP1B1- CXCL17-HPSE- GATA4
GO:0007059	Chromosome segregation	89/58	**<0.0001**	HORMAD1-CCNE1-FAM83D-SGO1-CEP55-TOP2A

The category size is calculated based on the number of overlapping genes between the annotated genes in the category and the reference gene list for the "GSEA" method. Significant p values are in **bold.**

5.3.6 Assessment of the Ki67 automated scores

The mean +SD of the Ki67 automated score was 22.0+14.56. 40% of patients showed low Ki67 automated scores while 826 (60%) showed high scores. The mean + SD of Ki67 CL score was 0.377+0.118; while 695 (51%) of patients had low Ki67CL score, 692 (49%) showed high Ki67CL score. High Ki67 automated score showed association with parameters characteristic of aggressive tumour behaviour including high grade and poor NPI ($p<0.0001$). Similarly, high Ki67CL score showed a significant association with high tumour grade ($p<0.0001$), lymphovascular invasion ($p= 0.009$) and poor NPI ($p< 0.0001$). Head-to-head comparison of manual Ki67 assessment versus automated Ki67 assessment showed similar correlations with clinicopathological parameters as shown in (Table 33).

Table 33: Relationship between Ki67 expressions in luminal BC assessed by manual versus automated scoring and clinicopathological parameters.

Variables	Manual scoring of Ki67		X^2	Automate scoring of Ki67		X2
	Low (n, %)	High (n, %)	P-value	Low (n, %)	High (n, %)	P-value
Age						
<50	249 (45)	308 (55)	10.1	269 (47)	305 (53)	3.4
≥50	863 (53)	781 (47)	**0.001**	868 (51)	824 (49)	0.07
Menopausal state						
Premenopausal	274 (46)	324 (54)	7.2	284 (46)	333 (54)	5.3
Post-menopausal	651 (52)	778 (48)	**0.007**	863 (52)	814 (48)	**0.02**
Tumour size						
≤2cm	882 (55)	726 (45)	45	887 (54)	769 (46)	31
>2cm	240 (39)	376 (61)	**<00001**	257 (40)	378 (60)	**<0.0001**
Tumour grade						
Grade 1	394 (77)	116 (23)	368	400 (76)	130 (24)	319.2
Grade 2	652 (53)	588(47)	**<0.0001**	652 (51)	625 (49)	**<0.0001**
Grade 3	79 (17)	398 (83)		95 (20)	392 (80)	
Histologic tumour types						
No special type (NST)	502 (40)	758 (60)	160	528 (41)	768 (59)	130.1
Lobular	171 (63)	99 (37)	**<0.0001**	149 (53)	132 (47)	**<0.0001**
Other special types	120 (85)	22 (15)		121 (80)	31 (20)	
NST mixed	332 (60)	223 (40)		349 (62)	216 (38)	
Lymph node status						
Negative	860 (52)	808 (48)	2.9	852 (49)	872 (51)	0.93
Positive	265 (47)	294 (53)	0.08	295 (52)	275 (48)	0.3
Lymphovascular invasion						
Absent	996 (53)	876 (47)	34	988 (51)	939 (49)	7.8
Present	129 (36)	226 (64)	**<0.0001**	159 (43)	208 (57)	**0.005**
Nottingham prognostic index						
Good prognostic group	758 (63)	452 (37)	171.3	750 (60)	502 (40)	119.1
Moderate prognostic group	349 (38)	566 (62)	**<0.0001**	378 (40)	566 (60)	**<0.0001**
Poor prognostic group	18 (18)	84 (82)		23 (23)	79 (77)	

5.3.6.1 Automated scores and outcome analysis

Patients with high Ki67CL score had worse BCSS and short DMFS than patients with low score (both p <0.0001) (**Figure 5-18**). Similar findings were shown in the association between automated score and patient outcome. Ki67 expression using manual assessment also showed strong association with patient survival. **Figure 5-19** shows Kaplan Meier curves of manual versus automated Ki67 and patient outcome. Notably, Ki67 using both methods stratified patients who received endocrine therapy only into two risk groups (p<0.0001). However, in patients who received both endocrine and chemotherapy, there was no association between Ki67 and patient outcome using either automatic, CL or manual Ki67 scoring (p=0.21, 0.18 and 0.10, respectively).

In a multivariate Cox regression analysis including Ki67 scores (manual score, automated score and Ki67CL), tumour size, tumour grade and LN status, All Ki67 score showed independent association with outcome (**Table 34**, **Table 35** and **Table 36**). However, with incorporating all Ki67 scores with the clinicopathologic parameters, manual Ki67 scores and Ki67CL scores showed independent association with patient outcome while the automated scores lost the association.

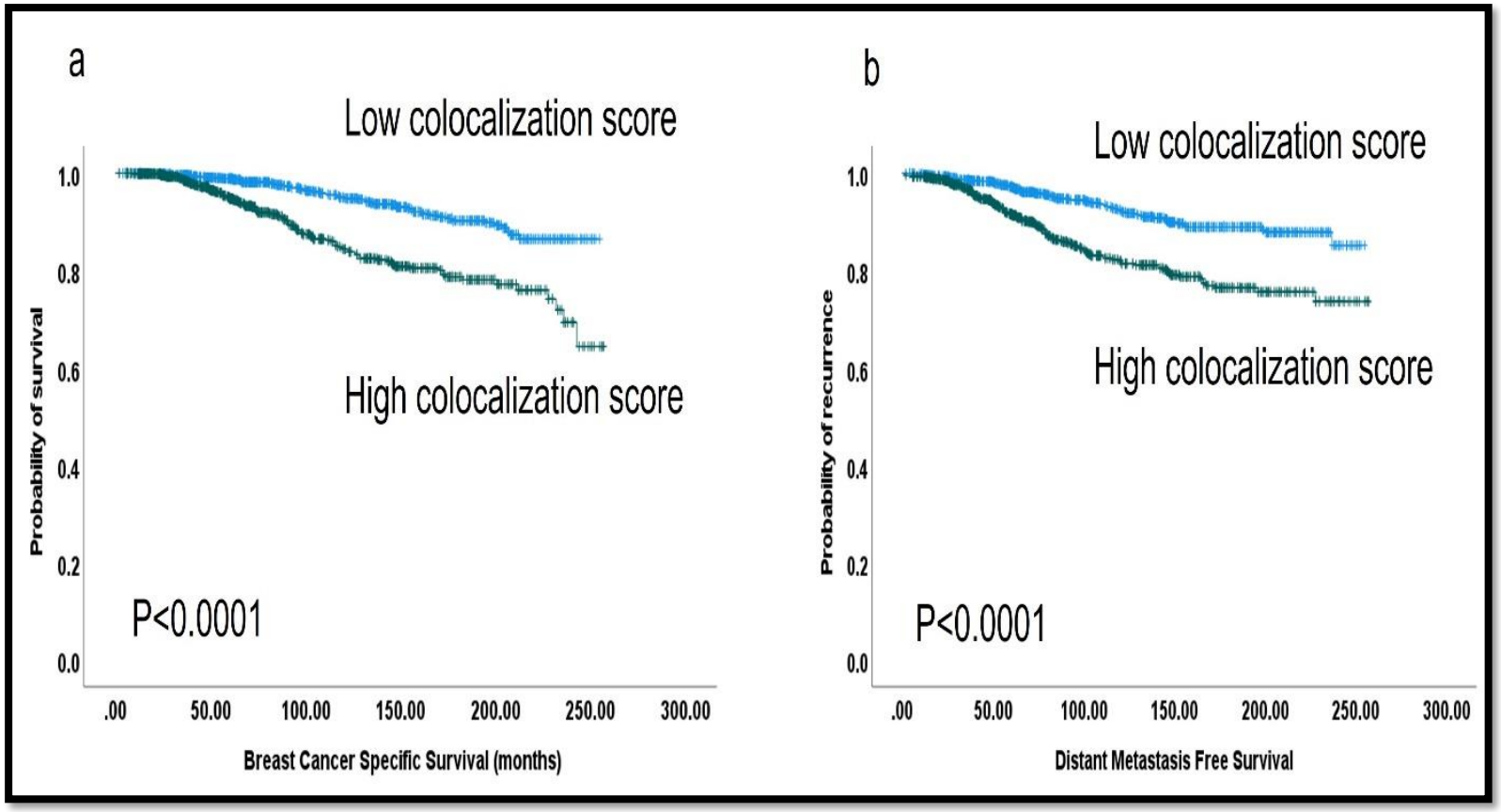


Figure 5-18: Kaplan Meier curves show of the association of high localisation score with poor survival in terms of both BCSS (a) and DMFS (b).

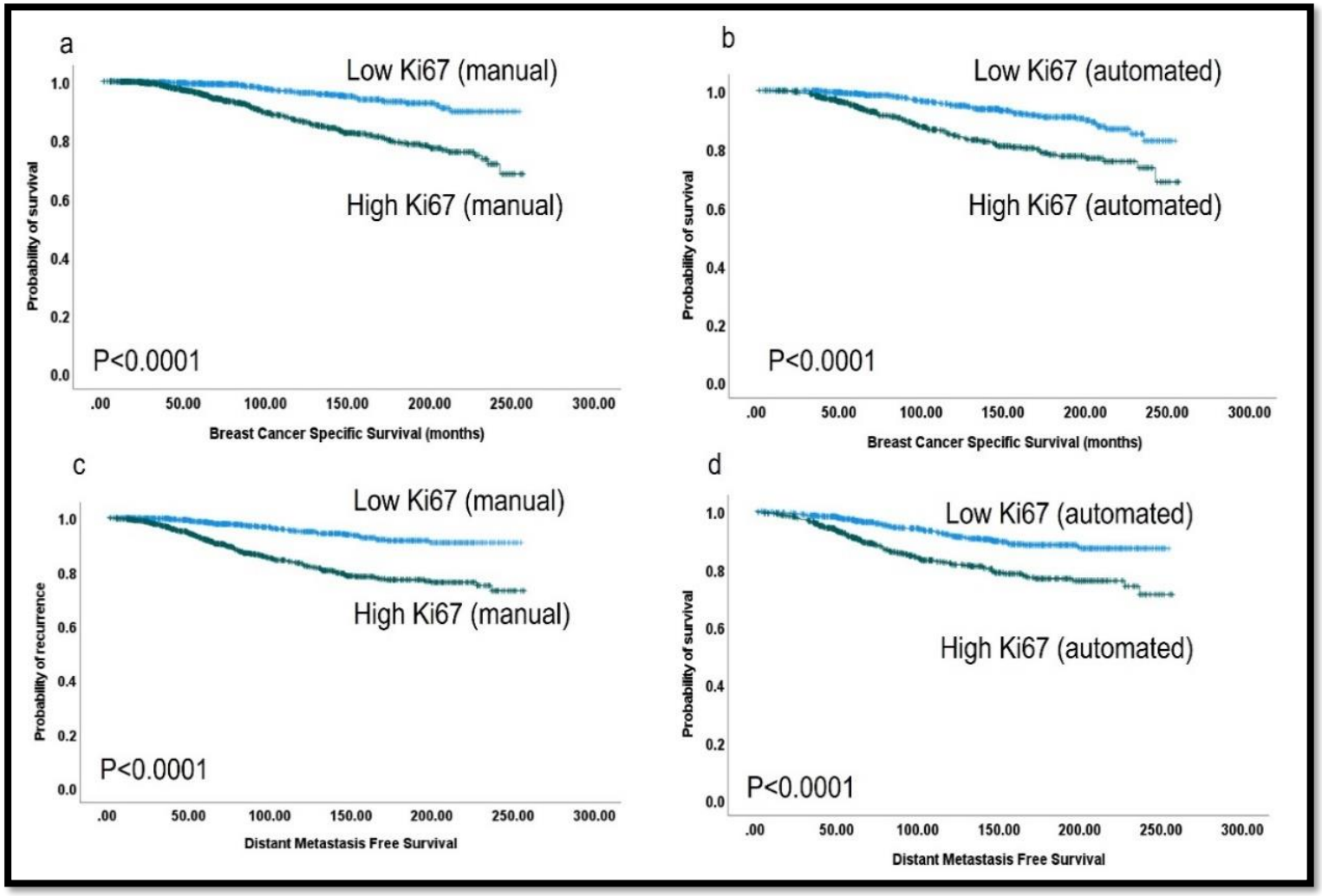


Figure 5-19: Kaplan Meier curves show the association of manual versus automated Ki67 score with both BCSS (a&b) and DMFS (c&d).

Table 34: Multivariate analysis of manual Ki67 scores in luminal BC

Parameters	Breast cancer specific survival		
	Hazard ratio	95% (CI)	P-value
Manual Ki67	2.1	1.3-3.2	**<0.0001**
Grade	1.8	1.4-2.4	**<0.0001**
Lymph node status	1.6	1.2-2.3	**0.001**
Lymphovascular invasion	1.6	1.1-2.2	**0.006**
Tumour size	1.3	1.0-1.9	**0.04**

95%CI, 95% confidence interval. Significant p values are in **bold.**

Table 35: Multivariate analysis of automated Ki67 scores in luminal BC

Parameters	Breast cancer specific survival		
	Hazard ratio	95% (CI)	P-value
Automated Ki67	2.0	1.3-3.2	**0.002**
Grade	1.9	1.5-2.5	**<0.0001**
Lymph node status	1.7	1.2-2.3	**0.001**
Lymphovascular invasion	1.5	1.1-2.1	**0.006**
Tumour size	1.4	1.0-1.9	**0.042**

95%CI, 95% confidence interval. Significant p values are in **bold**.

Table 36: Multivariate analysis of automated versus manual Ki67 scores in luminal BC

95%CI, 95% confidence interval. Significant p values are in bold.

Parameters	Breast cancer specific survival		
	Hazard ratio	95% (CI)	P-value
Colocalization score	1.5	1.1-2.2	**0.009**
Grade	1.9	1.1-2.2	**<0.0001**
Lymph node status	1.7	1.2-2.3	**0.001**
Lymphovascular invasion	1.6	1.2-2.3	**0.004**
Tumour size	1.5	1.1-2.2	0.064

5.4 Discussion

Ki67 is an essential indicator of cellular proliferation in malignancy and has a prognostic value in BC. Although, the Saint Gallen Guidelines

recommended the assessment of the Ki67 proliferation index [215], there is no standard method yet to evaluate its level of expression [204]. Several methods for Ki67 scoring have been proposed including the IKWG that recommended standardised scoring with online application to score 100 nuclei either negative or positive in each field directed then record the “Weighted global score”, which considered as the Ki67 index. This method seems to be tedious and requires routine digitalisation of the stained sections. In addition, the median required time to score a single case was 9 minutes as pathologists need first to review the Ki67-stained slide and then to estimate the percent area with 4 categories; negligible, low, medium, or high Ki67 expression then score 100 nuclei in each selected area in the whole slide, which would be challenging to achieve in the routine practice [164]. Others adopted a simple 5-grade scale to assess Ki67, depends on visual estimation of the ratio between positive and negative cells. Although, it was considered as a fast and easy task, it led to a discrepancy of 26% of cases compared to the manual counting in hotspot [463].

In this study, three main methods were compared for evaluating the percentage of Ki67 positivity in luminal (ER-positive, HER2 negative) BC where Ki67 has a specific prognostic and predictive role. These methods were chosen to achieve subjective versus objective Ki67 assessment. The clinical value of staining pattern and intensities, which could reflect the underlying biology was assessed. Several methods were considered similar to other published methods of Ki67 scoring however, some were

excluded due to the reported high level of discrepancy or for showing no or less significant association with patient outcome [463-466].

Although high Ki67 expression, regardless of assessment method, showed a significant association with parameters characteristics with aggressive tumour behaviour including higher grade, and LN metastasis, the best association was related to Ki67 expression using the quantification method. This indicates that the hotspot has the highest proliferative pool, which drives tumour behaviour and counting of those cells represents the optimal method for Ki67 assessment. Although the quantification method was the longest method compared to the other two methods, it was the most accurate method to predict patients` outcome in addition, it showed the highest degree in inter-observer concordance.

Although, each of the methods used had a significant association with disease outcome, calculation of Ki67 expression levels through counting of positive cells among 1000 invasive tumour cells within the hotspot showed the highest HR, the lowest p value and independent prognostic value compared to the other two methods. Using common cutoffs for categorisation of Ki67 expression into low and high groups, only Ki67 with quantification method showed consistent, significant association with poor outcome in a multivariate analysis while the other two methods lost the significance. It has been reported that a higher Ki67 value using the hotspot is strongly correlated with poor outcome [36, 467]. Also, it was showed that the determination of proliferation markers in BC should be standardised to hotspot counting as it showed greater risk of shorter tumour-free survival [465, 468]. In a previous study, using digital image

analysis for Ki67 scoring, quantification of Ki67 positive cells in hotspot was found to be superior to average Ki67 score within the whole slide and even outweigh mitotic count in terms of the association with BC outcome [439]. The IKWG recommends scoring the average Ki67 based on the lack of sufficient evidence to use the “hotspot” [200] and with the context of using core biopsy to assess Ki67 which may lack representative hotspot for the tumour proliferation index [164]. Others claimed that the prognostic powers of Ki67 in luminal BC, measured by either hotspot or average, were similar [449]. High Ki67 expression using all the previous scoring methods had strong association with poor outcome in patients who received endocrine therapy. However, high Ki67 expression using quantification method had a stronger association with poor outcome in those group of patients compared to the other two methods. This suggests the accuracy of quantification method to predict therapy response and confirm that high proliferative tumours showed worse outcomes and those patients should be considered for more aggressive therapy including chemotherapy. Furthermore, it suggests that Ki67 scoring using hotspot areas represents higher proliferation rates that could drive the biological behaviour of tumour [203].

In patients who received both endocrine therapy and chemotherapy, there was no association between high Ki67 and patient outcome. This suggests that patients with high Ki67 expression showed a better response to chemotherapy and the outcome of those patients became similar to patients with low Ki67. This also indicates that adding

chemotherapy to high Ki67 cases can improve the outcome consistent with previous studies [467].

Although, quantification of Ki67 in hotspot method was associated with highest hazard risk, this method is time consuming. To avoid this limitation and achieve the most reliable and clinically applicable method, the power of the quantification method was used in a simplified less time-consuming method. In this proposed method, the number of fields to count Ki67 depended on the cellularity of the tumour and the microscope field diameter. Consequently, pathologists scoring Ki67 in the clinical setting can calibrate their HPF and estimate the number of tumour cells per HPF in tumours with different cell density and assess the approximate number of HPFs that are required to count 1000 cells per case without the actual cell count in every case. The current results showed that this method reduced the scoring time and achieved a high concordance rate with the actual counting. Therefore, this simplified method can be implemented in routine practice. Although this method can be perceived by some as similar to the other two methods, it provides consistency and reproducibility and reduces the variation between cases with different cell density when scoring using fixed areas.

However, in cases with Ki67 score around the cutoff between high and low Ki67 categories, which may have therapeutic implications, detailed counting of 1000 cells within the hotspot may be considered to ensure reliability and accuracy.

In the current study, homogeneous Ki67 staining pattern was associated with longer survival while granular pattern showed strong association

with shorter survival. In addition, granular Ki67 expression was an independent prognostic factor in early-stage luminal BC. This indicates that the granular pattern could give information about patient prognosis.

Luminal BCs are heterogeneous in terms of Ki67 expression, which is evident by marked hotspot and this heterogeneity is one of the obstacles that makes the standardisation of Ki67 interpretation difficult and decreases its reproducibility [449]. In addition, Ki67 scores in hotspot versus average scores affect the categorisation of BC into the high versus low proliferation subgroups. In this study, high heterogeneity index showed a significant association parameters characteristic of aggressive tumour behaviour. Not only did high Ki67 heterogeneity index have a significant association with poor outcome in luminal BC patients, but also it predicted a poor response to endocrine therapy as there was a significant association between high heterogeneity index and poor outcome independent of tumour size and lymph node status in patients in who received endocrine therapy. Similarly, a previous study found a significant correlation between higher Ki67 heterogeneity index and recurrence in luminal BC [468]. Moreover, in this study, high heterogeneity index was significantly associated with tumours with high Ki67 expression, and this can explain the aggressive behaviour in luminal tumours showing high heterogeneity [468].

Both Ki67, which stains all the cells in the cell cycle, and mitotic scores which counts the number of cells in mitosis are used interchangeably to assess the proliferative activity of BC [436, 469, 470]. Although Ki67 has attracted a lot of attention as a measure of the proliferative activity of BC,

assessment of mitotic activity is a well-established approach, which is performed in H&E stained sections and is part of the Nottingham grading system [470]. Although not all cycling tumour cells will progress to the mitotic phase, there is a concern that the mitosis represents the smallest proportion of cycling cells, thus not precisely reflecting the proliferation rate of tumours [471, 472]. Therefore, this study aimed to assess the proportion of BC entering the mitotic phase.

Previous studies have demonstrated that mitotic counts and Ki67 expression are positively correlated [472, 473]; however, the dynamic relationship between them is not fully explored. In addition, the characterisation of BC in the cell cycle remains poorly defined. In the current study, proportion of BC cells in the cell cycle and in mitosis were quantified in the same areas using full face sections and a well-defined methodology [149, 200]. The cohort was categorised into four groups based on Ki67 positive cells and Ki67 positive mitotic counts using the same tissue area to evaluate the agreement between these two measures to better define the proliferative activity of BC cells.

In this study, strong association was observed between the proportion of cycling cells and cells in mitosis, and the mean percentage of cycling cells (Ki67 positive cells) in BC ranged from 24% to 30% based on the method of assessment, whether in the whole area or the hotspots, respectively. 52% of BC showed low proliferative activity, whereas 48% of cases showed high proliferation. The mean proportion of BC cells in mitosis was ~5%, and a high proportion of those cells was significantly associated

with features of aggressive tumour phenotypes including grade 3 tumours, triple negative subtype, the presence of LN metastasis and lymphovascular invasion. This suggests that quantification of cycling BC cells in mitosis may be useful to predict tumour aggressiveness. This is supported by a previous study showing that the entry of cancer cells into mitosis sensitises tumour cells to new anti-cancer strategies [474]. Similarly, Lee et al reported that the proportion of cells entering mitosis might be calculated from the mitotic counts compared to the Ki67 expression [475]. Furthermore, the correlation between mitosis and Ki67 hotspots in heterogeneous tumours has been explored and the majority of cases exhibited overlapping of Ki67 and mitosis within and adjacent to hotspots. This indicates that the same hotspots of both indicators may reflect the most biologically active part of the tumour, which drives the outcome of the disease.

To define the different proliferative activity of BC cells, the whole cohort was categorised into four proliferative groups according to Ki67 positive cells and Ki67- positive mitotic figures in 1000 cells within a hotspot. 42% of the cohort harboured both low mitosis and low Ki67, and this was significantly associated with parameters of good tumour behaviour. This suggests low Ki67, and low mitosis group may better distinguish tumours with better outcomes and waive those patients from more aggressive therapy.

Notably, 18% of this cohort showed discrepant Ki67 and mitosis, where low mitosis/high Ki67 group was significantly associated with the luminal

B subtype. The high Ki67 expression in this group may give a false proliferation index as this group did not show significant association with any of the other poor clinicopathologic parameters.

This group may be resulted from pre-analytic factors as mitosis seems to be affected with poor fixation more than Ki67 immunohistochemistry; however, this cohort is well fixed. Low mitosis/high Ki67 group represents the group of BCS with high proportion of cells in the cell cycle but only few of them entered mitosis. For example, some cells may stop at a certain point of the cycle, either stop their cycles or get into apoptosis [476]. Additionally, it may be related to the length of S-phase duration as it has been found that an increasing DNA content, as well as an increasing percentage of cells in S-phase, were both correlated with increasing Ki67 expression [475].

The last group (high mitosis/low Ki67) showed significant association with triple negative subtype, grade 3 tumours and higher NPI. This group was associated with poor prognosis and thus may require more aggressive therapy than the other proliferative groups. The explanation of this group seems to be difficult; it was hypothesised that mitotic blocks for DNA repair may lead to a larger number of mitotic figures being counted [476]. However, this could explain a small number of mitotic cells but not reach 8% of the cohort as found in the current study. This BC group may show the highest proliferative activity as most of the cycling cells were committed to cell division and hence captured in the mitosis phase. Similarly, a previous study of discrepant Ki67 and mitosis showed that

high mitosis with low Ki67 showed similar poor outcome to high mitosis with high Ki67 [475]. To further explore this group, the percentage of both Ki67 positive cells and Ki67 positive mitotic cells was compared in the same area, that percentage of Ki67 positive cells was found to outweigh percentage of Ki67 positive mitotic cells as Ki67 stains all malignant cells entering cell cycles including mitotic cells. The mean ratio between both groups (percentage of Ki67 positive cells to percentage of Ki67 positive mitotic cells was 16:1 (range from 5:1 to 60:1). This explains that Ki67 percentage was falsely interpreted when compared to the absolute mitotic counts. Previously, these groups were classified based on mitotic counts in H&E-stained slide and Ki67 expression [476]. However, in the current study, Ki67 expression and mitosis were assessed in the same area on hotspots of Ki67-stained sections for consistency.

The Ki67 cutoff for differentiating luminal A from luminal B BC has changed over time. The 2011 Saint Gallen Consensus Meeting defined as “low proliferation” tumours with a Ki67 index <14 % a cutoff established by comparison with PAM50 intrinsic multigene molecular test for classification of Luminal cancers [200]. However, during the 2013 Saint Gallen Conference, the majority voted that a threshold of ≥20 % was indicative of “high” Ki67 status [477]. At the same time, several studies reported low reproducibility of Ki67 results, mainly in the subset of cancers with intermediate proliferation activity (between 15 and 30 %) [390]. Again, in March 2015, during the Saint Gallen Conference, it was recommended that 20% is considered as the optimal cutoff of Ki67 in luminal BC [454]. The use of the median Ki67 value of local laboratory

was proposed as the optimal cutoff for Ki67 categorisation and this was accepted by a panel of experts. A previous meta-analysis of 64,196 patients showed that a cutoff of at least 25% has better discriminatory power compared to other cutoffs in luminal BC [216]. Another study included 8088 patients and used automated Ki67 scoring concluded that Ki67 >12 % had a worse 10-year BCSS in ER positive BC patients [217]. Moreover, Ki67 has been compared with EndoPredict risk scores and Ki67 above 25% was found to show a strong overlap with high-risk EndoPredict test [218, 478]. In the same way, Ki67 levels have been compared to Oncotype DX recurrence score (RS), and a positive correlation between Ki67 and RS was found with low percentage of discrepancy was detected between both [479-482].

It seems that cutoff values ranging from 10% to 30% have been the most common to dichotomise populations [219, 220]. In this study, the common cutoffs for Ki67 in luminal BC, used in previous literature, were compared according to patient outcome and the likelihood of metastasis. Univariate and multivariate analysis were carried out and these cutoffs were analysed in a large retrospective cohort of Luminal ER+/HER2− BC, with long term follow-up, to establish which is the optimal Ki67 cutoff to stratify cases based on patient survival. Although, Ki67 expression levels using the different cutoffs showed a significant association with outcome, 10% cutoff showed the best cutoff to stratify patients having “good” or “poor” prognosis and was independent prognostic index in presence of other confounders. Similarly, Al Eskandarany and his colleagues demonstrated

that 10% cutoff for Ki67 could stratify BC with grade I and II into statistically significant prognostic groups [221].

To resolve the problem of subjective Ki67 scoring, automated counting by a computer software device might be a breakthrough, however, not all institutes can afford this due to lack of the required infrastructure for pathology service digitalisation [53]. The majority of ER+/HER2- BC patients belong to an intermediate risk group in which some of the patients may benefit from adjuvant chemotherapy besides endocrine therapy [483]. Even though several risk calculators have been developed to estimate a patient's risk of BC recurrence and mortality, these methods are expensive and have long turnaround times [68, 484].

In this study, AI-based tool was developed to extract a Ki67-based risk score, named Ki67CL score, for risk stratification of intermediate risk BC patients who were treated with endocrine therapy. Ki67CL score can significantly categorise patients into different risk groups. Also, a significant association between high Ki67CL score and tumours characteristics associated with aggressive tumour behaviour was observed indicating that this score can be used as surrogate marker in luminal BC. In addition, automated Ki67 score showed a significant association with outcome in luminal BC patients in terms of both BCSS and DMFS. Ki67CL, automated Ki67 and manual Ki67 scores stratified endocrine treated patients into two distinct prognostic groups. However, there was no significant difference between any of these scores and patient outcome in the group of patients who received additional

chemotherapy suggesting that the Ki67 scores can help make the decision on whether to prescribe adjuvant chemotherapy or not.

Based on multivariate analysis Ki67CL score, automated Ki67 score and manual Ki67 were independent prognostic markers and added more prognostic information when each of those scores was plotted separately. This suggests the importance of the co-localisation of Ki67 positive and Ki67 negative tumour cells and automated Ki67 score besides the manual Ki67 scores in determining the risk of early-stage luminal BC patients. However, the lack of standardisation for manual Ki67 scoring encouraged the search for automated Ki67 scoring methods like digital image analysis. In clinical practice, both H&E morphological features and biomarker information are used as prognostic and predictive indicators. Combining the proposed Ki67-based prognostic marker with features extracted from H&E staining and clinical factors would have the potential to achieve better risk stratification performance.

Conclusion

Based on the previous findings, the quantification per 1000 cells is considered the optimal method of Ki67 scoring as it shows the highest hazard risk. Heterogeneity of Ki67 staining and patterns of its expression could give valuable information about the clinical outcome of BC. The proportion of BC cells in the cell cycle and mitosis is variable. Not only the number of cells in the cell cycle or mitosis but also the difference between them could provide valuable information on tumour

aggressiveness. 10% showed the optimal cutoff for categorisation of the proliferative activity in early stage-luminal BC.

Chapter 6 The clinical value of PR in BC

6.1 Introduction

Based on gene expression analysis, BC has been classified in five molecular subtypes linked to different prognosis and therapeutic responsiveness [485]. So far, high costs and technical issues have limited the use of genomic profile tools in routine clinical practice. Therefore, therapeutic decision-making process is still commonly based on clinical and immunohistochemical characteristics, including tumour size, grade, nodal status, ER expression, HER2 status and Ki67 values. In addition to ER, PR plays an important role in BC biology and its expression is dependent on oestrogen [287, 486, 487]. Previously, PR was only considered as an indicator of a functional ER in BC tumours [287]. PR is a member of the steroid nuclear receptor family of ligand-dependent transcription factors. PR consists of central portion of DNA binding domain, C-terminal ligand-binding domain and amino-terminal domain composed of intrinsically disordered protein [488].

There is an intimate interaction between PR and ER expression where ER regulates PR expression and, in return, PR modulates ER action which indicates that the ER pathway is intact [309, 489]. Additionally, PR plays a role in control of several normal cellular functions, including cellular integrity, growth and proliferation [490]. Although ER and PR are prognostic markers in invasive BC and they are the key markers for molecular subtype classification, only ER is used in the clinical settings as a well-established predictive marker of endocrine therapy [491]. Despite prognostic significance of PR in BC [309, 489, 492, 493], the

routine assessment of PR in BC remains controversial among laboratories. This is mainly related to its uncertain predictive value in the adjuvant setting [494-497]. This is mainly related to its uncertain predictive value in the adjuvant setting [299]. This lack of clinical significance compared to ER expression seems understandable because it has been shown that almost all PR positive tumours are ER positive, and ER negative PR positive tumours are very rare [299, 498].

In addition, different scoring cutoffs to define PR positivity have been published in literature [301, 499-501]. Also, the expression of PR in core biopsy and surgical excision specimens varies widely compared to ER [502, 503]. The United Kingdom Royal College of Pathologists (RCPath) breast dataset regards PR testing in BC as optional [297]. However, The ASCO/CAP guidelines and others recommend routine assessment of PR expression in ER-positive tumours [487]. In the current era of precision medicine, the routine use of PR should be reviewed, and evidence should be interpreted in the right context.

AI and machine learning technology are gaining ground in pathology because of their unmatched ability to make accurate predictions [504, 505]. Automated assessment of ER and PR has been widely investigated using AI-based tools to decrease the subjective interpretation and reduce inter- and intra-observer variability of these valuable biomarkers [506]. AI-based methods have the potential to achieve accuracy levels that are hard to obtain with conventional methods [507] and may allow for wide-spread usage of PR in clinical routine practice.

6.1.1 Hypothesis and aims:

This chapter hypothesised that PR has important predictive role in early-stage luminal BC and can help personalisation of patient therapy. Therefore, this chapter aims to utilise a large and well characterised cohort of luminal BC to 1) evaluate the optimal cutoff for categorisation PR expression levels, 2) assess the prognostic significance PR expression in the early-stage luminal BC 3) compare PR with the proliferation index, Ki67, 4) compare the expression of PR in core biopsy versus surgical excision specimens, 5) evaluate the automated expression of PR using AI- algorithm and 6) compare the automated PR score against the visual score.

6.2 Patients and Methods

6.2.1 Study cohort

This study was conducted on a large cohort (n=2304) of ER positive and HER2 negative BC with available clinicopathological, treatment and outcome data were available for this cohort (**Table 37**). Original clinical PR score which was carried out on the diagnostic core biopsies was available to a subset of cases (n=727), where 90% (651/727) were positive using 1% cutoff to define PR positivity. Ki67 was used for comparative analysis with PR. Adjuvant endocrine therapy was offered to 90% of patients included in the study while 12% of them received both endocrine therapy and chemotherapy.

Table 37: Clinicopathological characteristics of the study cohort

Variables	No (%)
Age at diagnosis (Year)	
<50	561 (24)
≥50	1716 (76)
Menopausal status	
Premenopausal	611 (27)
Post-menopausal	1693 (73)
Tumour size (cm)	
≤2	1682 (73)
>2	619 (27)
Histologic tumour grade	
1	543 (24)
2	1269 (53)
3	492 (23)
Histologic tumour types	
No Special Type (NST)	1316 (57)
Lobular	273 (12)
Other special types	150 (7)
NST mixed	565 (24)
Lymph node status	
Negative	1726 (75)
Positive	578 (25)
Lymhovascular invasion	
Absent	1931 (84)
Present	373 (16)
Nottingham prognostic index	
Good prognostic group	1274 (55%)
Moderate prognostic group	929 (40%)
Poor prognostic group	101 (5%)
Ki67 expression (10%)	
Low	897 (41)
High	1287 (59)

6.2.2 Immunohistochemical staining of PR.

The IHC staining was performed automatically on full face sections using DAKO Cytomation EnVision+ detection system using the same protocol used for PR staining in the clinical settings. Primary antibody (anti-PR (1E2) Rabbit Monoclonal Primary Antibody) was applied to tissue sections and incubated for 30 minutes. The correlation between PR status using full face sections and the PR status originated from core biopsies was tested to assess if there is any discrepancy in PR scores between the core biopsies and full-face sections.

6.2.3 Eyeballing evaluation of PR expression

The percentage of PR expression was assessed by visually estimation of the proportion of invasive tumour cells that showed nuclear staining for PR through scanning of the whole tumour at low power magnification using 10x objective lens. Areas of PR staining in normal ducts or in situ carcinoma were excluded from the final score. H. score was also calculated. Staining intensity is categorised into 0, 1, 2, or 3 (for negative, weak, moderate, and strong, respectively) (**Figure 6-1**). PR H. score was dichotomised into two groups based on a common cutoff used in literature for ER (H. score = 50).

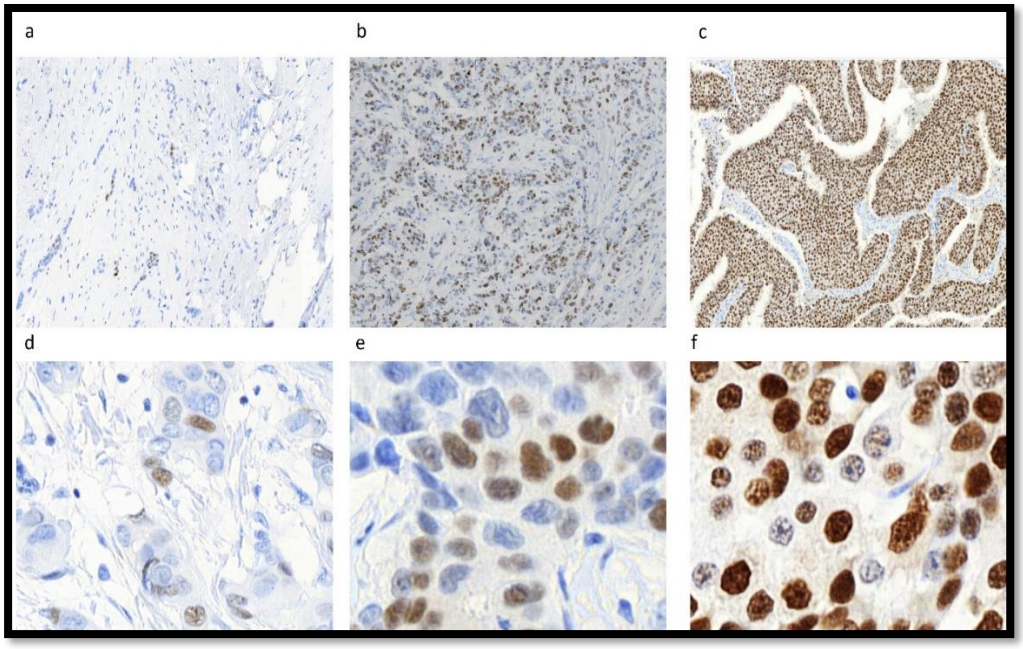


Figure 6-1: Expression of PR in luminal BC
a) low PR expression, b) moderate PR expression, c) high PR expression (20x), d, mild intensity of PR expression, e- moderate intensity of PR expression and f, moderate to strong intensity of PR expression (40x).

6.2.4 Exploring the optimal cutoff to define PR positivity.

In order to assess the optimal cutoff for PR for prognostic value, PR expression levels were firstly plotted against patient outcome using X-tile bioinformatics software [380]. Then, the previously recommended cutoffs in literature (1%, 10% and 20%) were compared in terms of patient outcome. These cutoffs are usually applied in either the clinical practice or in previous publications [487, 500, 501]. Then, the optimal cutoff was chosen based on the results of the univariate analysis using the most significant p-values associated with the highest HR after comparing the different cutoffs.

The sensitivity and specificity of the refined cutoff of PR expression in core biopsies was evaluated based on PR expression in full-face sections,

which was considered as the gold standard in this study. For defining the positivity of PR, a subset of cohort (n=1924) was utilised.

6.2.5 Automated scoring of PR

To assess the automated expression of PR, deep learning neural network model was trained to classify the neoplastic cells in the input images into 2 types of cells either negative or positive for PR staining. Training of the model was performed using large number of BC cases as described in detail in chapter 2.6.1.1.2. Automated PR score was calculated as the percentage of positive PR tumour cell in the whole WSI level (positive tumour cell/ all tumour cells). In addition, PR density score was evaluated by maximum PR density in the WSI level. PR density was calculated as the number of clusters inside areas showed high PR percentage. These scores were correlated with the clinicopathological parameters and patient outcome. Furthermore, automated PR scores were compared with eyeballing score. Automated PR expression was evaluated in 2304 specimens.

6.3 Statistical analysis

Statistical Package for the Social Sciences software v.27.0 (SPSS, Chicago, IL, USA) was used for statistical analysis. To assess the expression of PR either bimodal or linear, the cohort was categorised into 3 groups according to the percentage of PR expression in tumour cells, low ≤10, intermediate (21-70%), and high expression >70% and a distribution curve was carried out as shown previously in terms of ER expression [508]. Chi-square test was used for analysis of categorical

data. The percentage of Ki67 positivity was categorised into low versus high proliferative groups based on 10% cutoff. Similarly, 10% cutoff was used to classify automated PR expression levels into 2 groups. In addition, the median of PR density was used to classify it into low and high groups. The association between PR expression and clinicopathological parameters was assessed. In addition, univariate and multivariate analyses were used to assess the correlation of PR expression and patient outcome. The association of PR expression with the response of adjuvant therapy was also evaluated. Outcome analysis was assessed using Kaplan–Meier curves and the log-rank test. Cox regression models was used for the multivariate analysis. Estimated HR and their 95% confidence interval (95% CI) were calculated from univariate and multivariate regression models. For all tests, $P < 0.05$ (two-tailed) were statistically significant.

6.3.1 RESULTS

6.3.2 Eyeballing PR expression

The mean ± SD of eyeballing PR expression was 63± 39%, median was 75% (range 0-100). The mean ± SD of PR H. score was (145± 104), the median was 150 (range 0-300).

6.3.2.1 PR cutoff in BC

Using X tile software, a very high cutoff was identified as the best prognostic stratification of PR expression (75% and 25% in full face sections and in CNB, respectively). However, looking at the distribution of cases in each group and importantly at the discordance between full

face sections and CNB, other cutoffs were tested as well, these included 1%, 10%, and 20%. **Table 38, Table 39 and Table 40** show the prognostic significance of the various cutoffs, the specificity and sensitivity of each cutoff when assessed in CNB and the concordance between CNB and full-face sections. These analyses indicated that 10% cutoff can achieve best association with patient outcome including those who received endocrine therapy.

PR percentage showed bimodality of expression as 20% and 69% of PR were negative (≤10%) and highly expressed >70%, respectively (**Figure 6-2)** while only 11% showed moderate expression of PR (11-70%)

Table 38: Comparison between different cutoffs of PR expression in full face sections

PR cutoffs	Negative versus positive		BCSS in the whole cohort			DMFS in the whole cohort			BCSS in patients received endocrine therapy			DMFS in patients received endocrine therapy.		
	Negative (n %)	Positive (n %)	HR	95% CI	P-value	HR	95% CI	P-value	HR	95% CI	P-value	HR	95% CI	P-value
1%	226 (12)	1678 (88)	0.60	0.41-0.88	**0.010**	0.64	0.45- 0.91	**0.0140**	0.61	0.41-0.91	**0.015**	0.67	0.47-0.97	**0.030**
10%	381 (20)	1523 (80)	0.62	0.45-0.86	**0.004**	0.58	0.44-0.78	**<0.0001**	0.60	0.45-0.90	**0.011**	0.61	0.45-0.82	**0.001**
20%	452 (24)	1452 (76)	0.60	0.44-0.83	**0.001**	0.60	0.46-0.79	**<0.0001**	0.60	0.45-0.87	**0.005**	0.60	0.46-0.82	**0.001**
75%	865 (45)	1039 (55)	0.48	0.36-0.65	**<0.0001**	0.51	0.39-0.64	**<0.0001**	0.49	0.36-0.69	**<0.0001**	0.52	0.39-0.68	**<0.0001**

HR, hazard ratio, BCSS, breast cancer specific cancer, DMFS, distant metastasis free survival. 95%CI, 95% confidence interval. Significant p values are in **bold.**

Table 39: Comparison between different cutoffs of PR expression in needle core biopsies

PR cutoffs	Negative versus positive		BCSS			DMFS		
	Negative (n %)	Positive (n %)	HR	95% CI	P-value	HR	95% CI	P-value
1%	76 (10)	651 (90)	0.58	0.20-1.6	0.32	0.74	0.29- 1.9	0.52
10%	134 (18)	593 (82)	0.36	0.16-0.83	**0.01**	0.47	0.24-0.98	**0.04**
20%	155 (21)	572 (79)	0.32	0.14-0.70	**0.005**	0.4	0.23-0.88	**0.02**
75%	314 (43)	413 (57)	0.56	0.25-1.3	0.16	0.67	0.35-1.3	0.22

HR, hazard ratio, BCSS, breast cancer specific cancer, DMFS, distant metastasis free survival. 95%CI, 95% confidence interval, Significant p values are in **bold**

Table 40: Sensitivity and specificity of PR cutoff in needle core biopsies compared to full face sections.

PR cutoffs	Sensitivity	Specificity
1%	95%	60%
10%	**91%**	**86%**
20%	89%	74%
75%	74%	80%

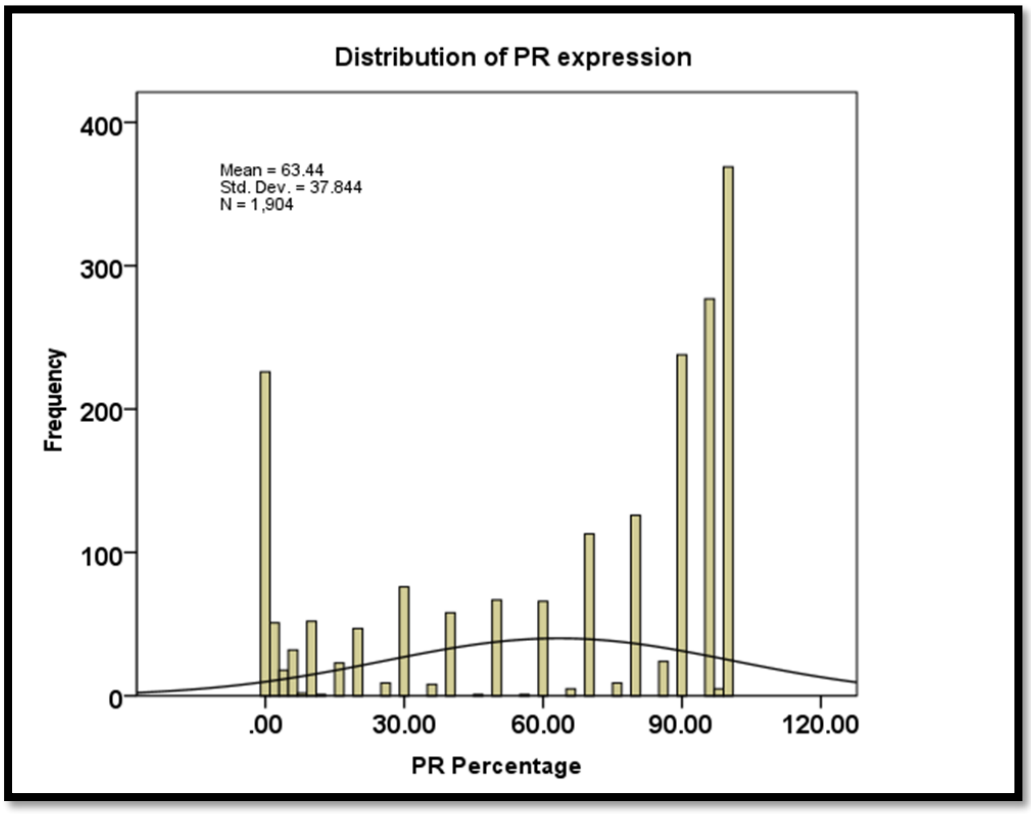


Figure 6-2: Normal distribution curve shows the bimodality of PR expression in BC.

6.3.2.2 PR expression and clinicopathological parameters in luminal BC

PR negativity (≤10%) showed strong association with older age patients (>50 years) ($p<0.0001$), post-menopausal status, grade 3 tumours and poor NPI scores ($p<0.0001$). However, no association was observed between PR status and the degree of pleomorphism, tumour size, LN status or LVI. The concordance rate was 80% between PR expression using CNB and full-face sections. However, 2% of negative PR cases in CNB were positive in full face sections. There was strong association between PR expression using full face sections and in core biopsies ($p<0.0001$) (**Table 41**). Moreover, there was a significant association between PR negativity and high Ki67 expression ($p=0.043$). Similarly, low PR H score showed significant associations with parameters characteristic of aggressive tumour behaviour including high tumour grade ($p<0.0001$) and NST tumour type ($p <0.0001$) (**Table 42**).

Table 41: Relationship between PR status and clinicopathological parameters

Variables	PR status		X^2
	Negative (n, %)	**Positive** (n, %)	**P-value**
Age at diagnosis			
<50	50 (12)	383 (88)	25.1
≥50	331(22)	1140 (78)	**<0.0001**
Menopausal state			
Premenopausal	58 (11)	462 (89)	35.1
Post-menopausal	323 (23)	1061 (77)	**<0.0001**
Tumour size (cm)			
≤2	258 (20)	1059 (80)	047
>2	123 (20)	464 (80)	0.49
Histologic tumour grade			
Grade 1	75 (17)	377 (83)	18.8
Grade 2	185 (18)	820 (82)	**<0.0001**
Grade 3	121 (27)	326 (73)	
Histologic tumour types			
No special type (NST)	209 (20)	809 (80)	15.4
Lobular	64 (27)	170 (73)	**0.002**
Other special types	28 (21)	102 (79)	
NST mixed	80 (15)	442 (85)	
Lymph node status			
Negative	282 (20)	1098 (80)	0.56
Positive	99 (19)	425 (81)	0.45
Lymhovascular invasion			
Absent	315 (20)	1236 (80)	0.47
Present	66 (19)	287 (81)	0.49
Nottingham prognostic index			
Good prognostic group	183 (18)	821 (82)	4.9
Moderate prognostic group	172 (22)	625 (78)	0.08
Poor prognostic group	26 (25)	77 (75)	
PR status (Original core biopsy)			
Negative	60 (88)	16 (12)	188.7
Positive	83 (13)	568 (87)	**<0.0001**

Table 42: Relationship between H. score of PR (low /high) and clinicopathological parameters

Variables	PR H. score		X^2
	Low H. score (n, %)	High H. score (n, %)	P-value
Age at diagnosis (years)			
<50	50 (12)	383 (88)	25.1
≥50	331 (22)	1140 (78)	**<0.0001**
Menopausal state			
Premenopausal	30 (11)	462 (89)	35.1
Post-menopausal	323 (23)	1061 (77)	**<0.0001**
Tumour size (cm)			
≤2	337 (26)	980 (74)	0.1
>2	155 (26)	432 (74)	0.7
Histologic tumour grade			
Grade 1	98 (22)	354 (78)	31.9
Grade 2	233 (23)	772 (77)	**<0.0001**
Grade 3	161 (36)	286 (64)	
Histologic tumour types			
No special type (NST)	276 (27)	742 (73)	21
Lobular	82 (35)	152 (65)	**<0.0001**
Other special types	29 (22)	101 (78)	
NST mixed	105 (20)	417 (80)	
Lymph node status			
Negative	360 (26)	1020 (74)	0.16
Positive	132 (25)	392 (75)	0.7
Lymhovascular invasion			
Absent	400 (26)	1151 (74)	0.01
Present	92 (26)	261 (74)	0.9
Nottingham prognostic index			
Good prognostic group	233 (23)	771 (77)	8.3
Moderate prognostic group	226 (28)	571 (72)	**0.01**
Poor prognostic group	33 (32)	70 (68)	

6.3.2.3 Outcome analysis

PR negativity showed a significant association with poor outcome in terms of shorter BCSS and poor DMFS, (p=0.004 and <0.0001, respectively) (**Figure 6-3).** Low PR H. score showed a significant association with shorter BCSS (p=0.002) and with DMFS (p=0.001) (**Figure 6-4**). When the cohort was stratified based on histologic tumour grade, PR status stratified only grade 2 tumours into two distinct prognostic subgroups (p=0.03) (**Figure 6-5).** In addition, when the cohort was classified based on the menopausal status, there was a statistically significant association between negativity of PR and poor outcome in postmenopausal patients in terms of both BCSS (p=0.002) and DMFS

(p<0.0001) (

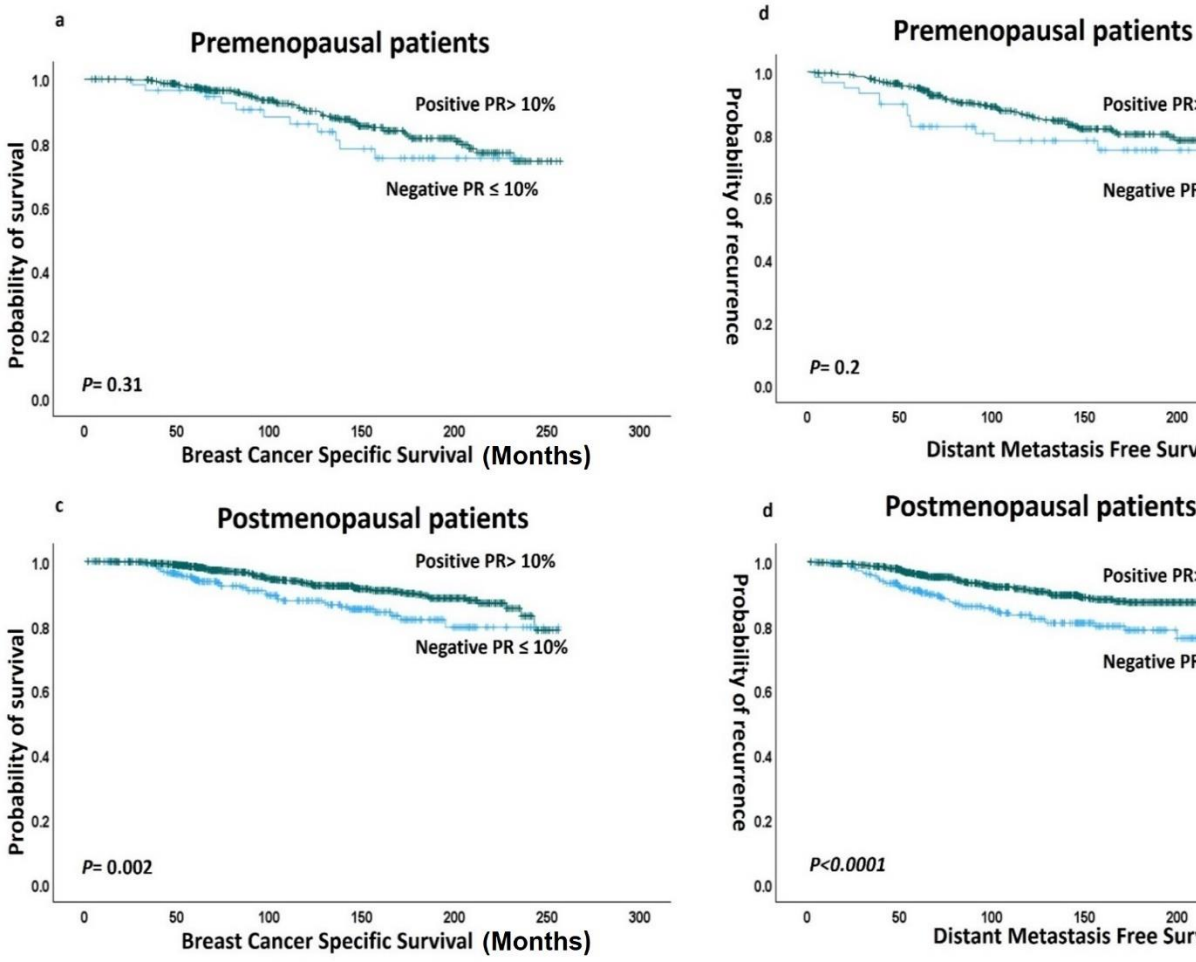


Figure 6-6).

When adjuvant system therapy was considered, PR negativity showed strong association with poor outcome in patients who received endocrine therapy in terms of BCSS (p=0.009) and DMFS (p=0.001) (**Figure 6-7**). Similarly, low PR H. score showed a significant association with poor outcome in patients who received endocrine therapy (**Figure** 6-8**).** However, in patients who received endocrine therapy and chemotherapy, there was no association between either PR status or PR H. score and patient survival.

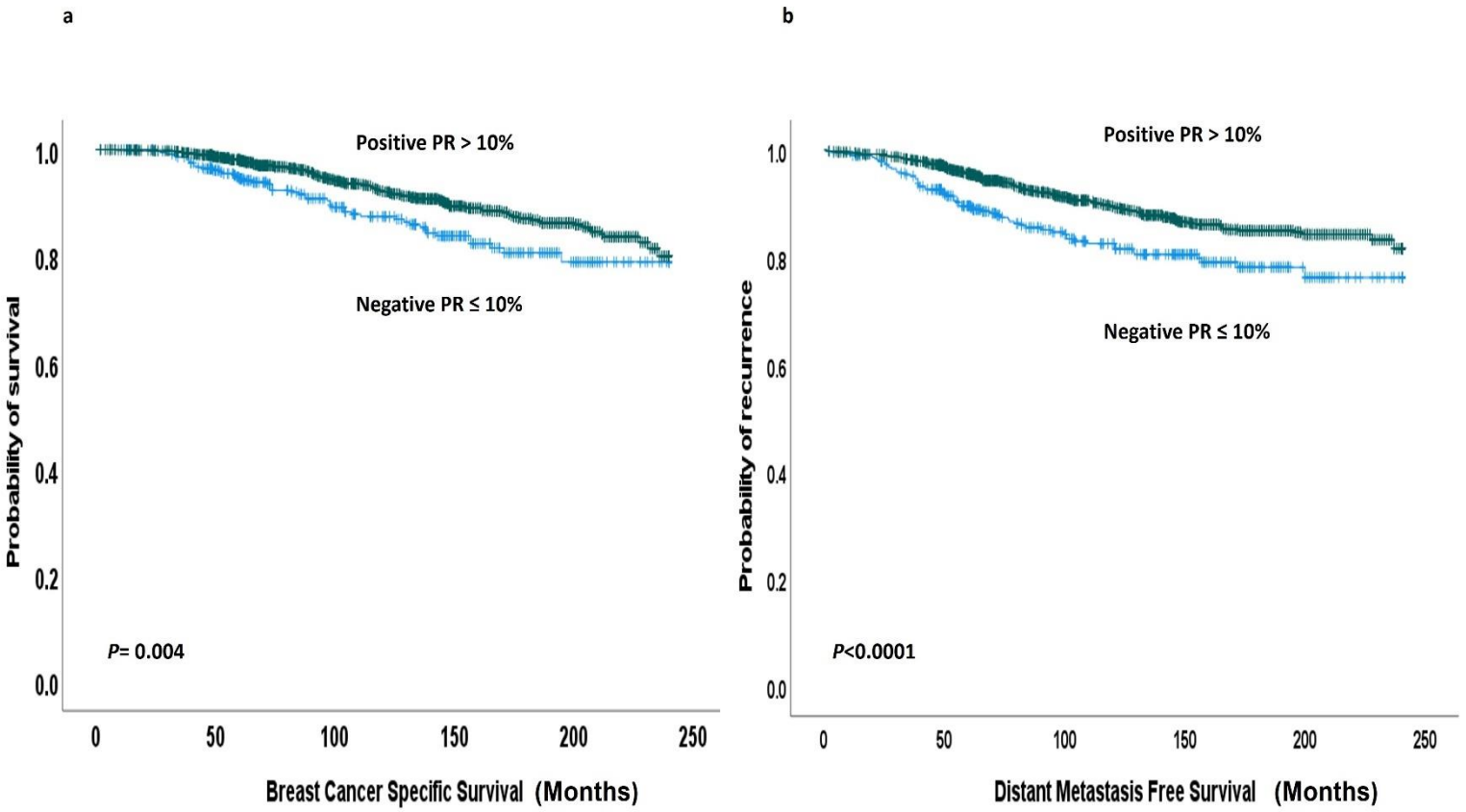


Figure 6-3: Kaplan Meier curves show the association of PR status with a) BCSS and b) DMFS.

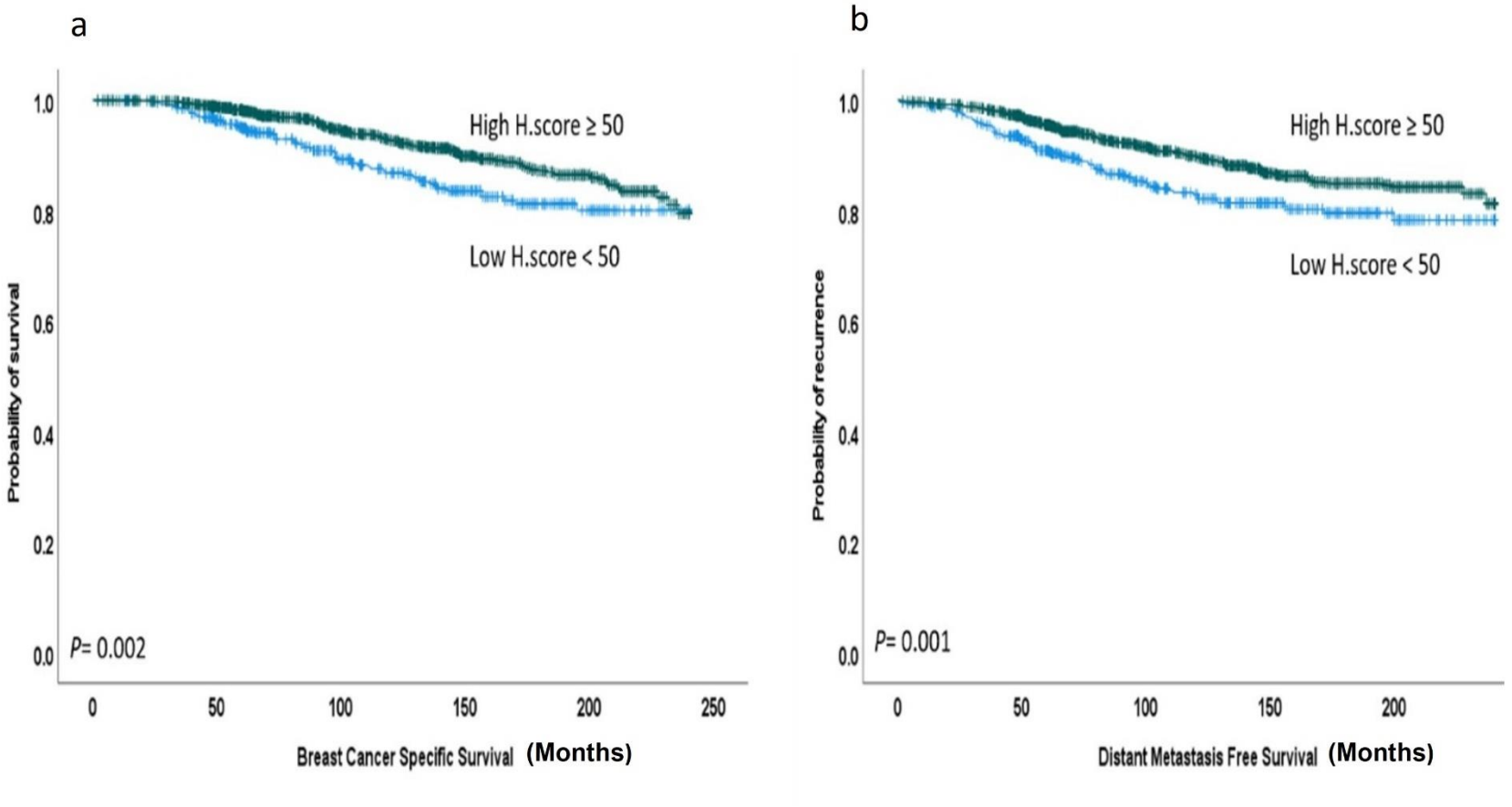


Figure 6-4: Kaplan Meier curves show a significant association between low PR H. score and a) BCSS and b) DMFS**.**

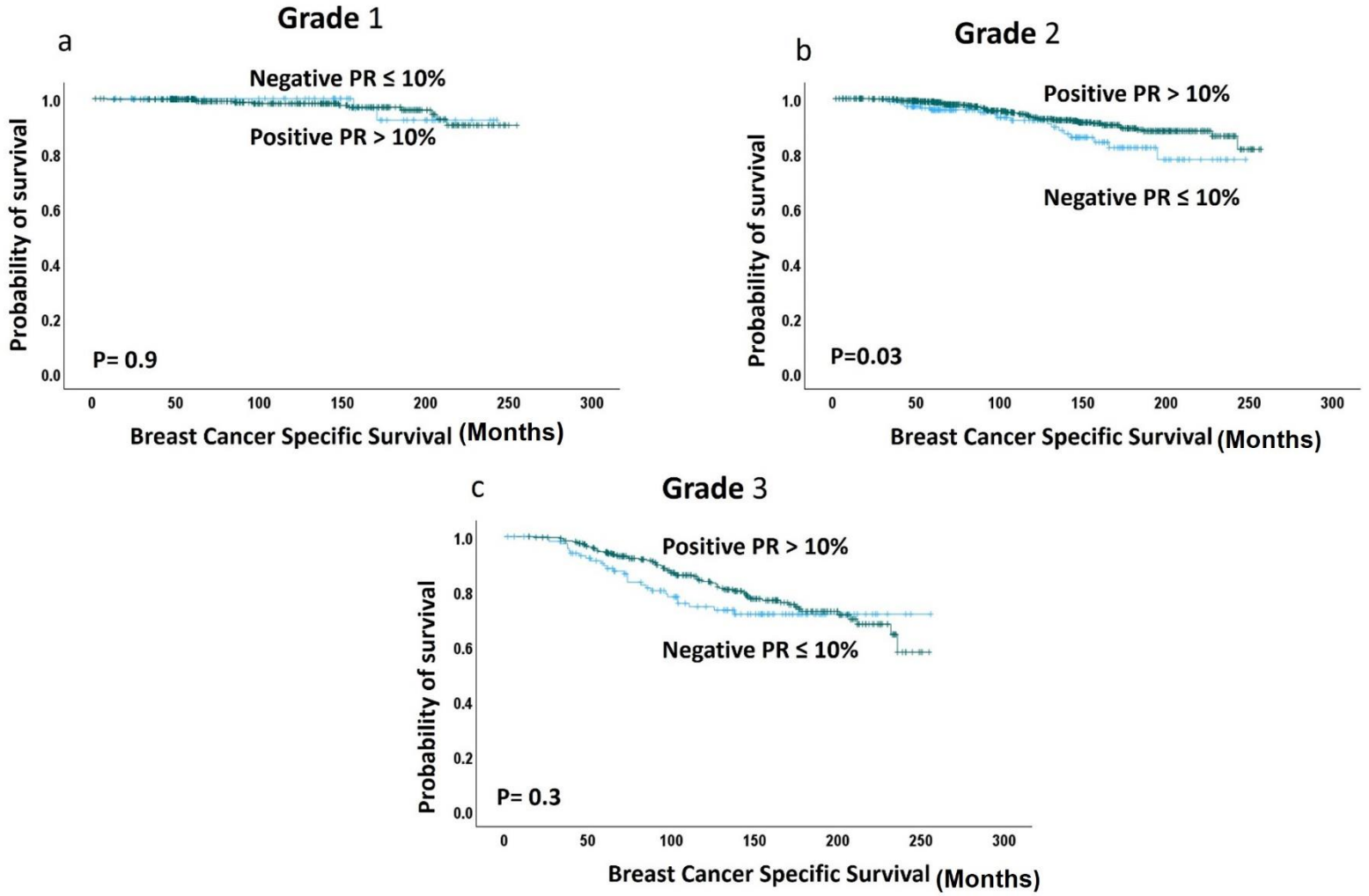


Figure 6-5: Kaplan Meier curves show a significant association of PR negativity with short BCSS in grade 2 tumours (b), but no association was found in either grade 1 or grade 3 tumours (a&c).

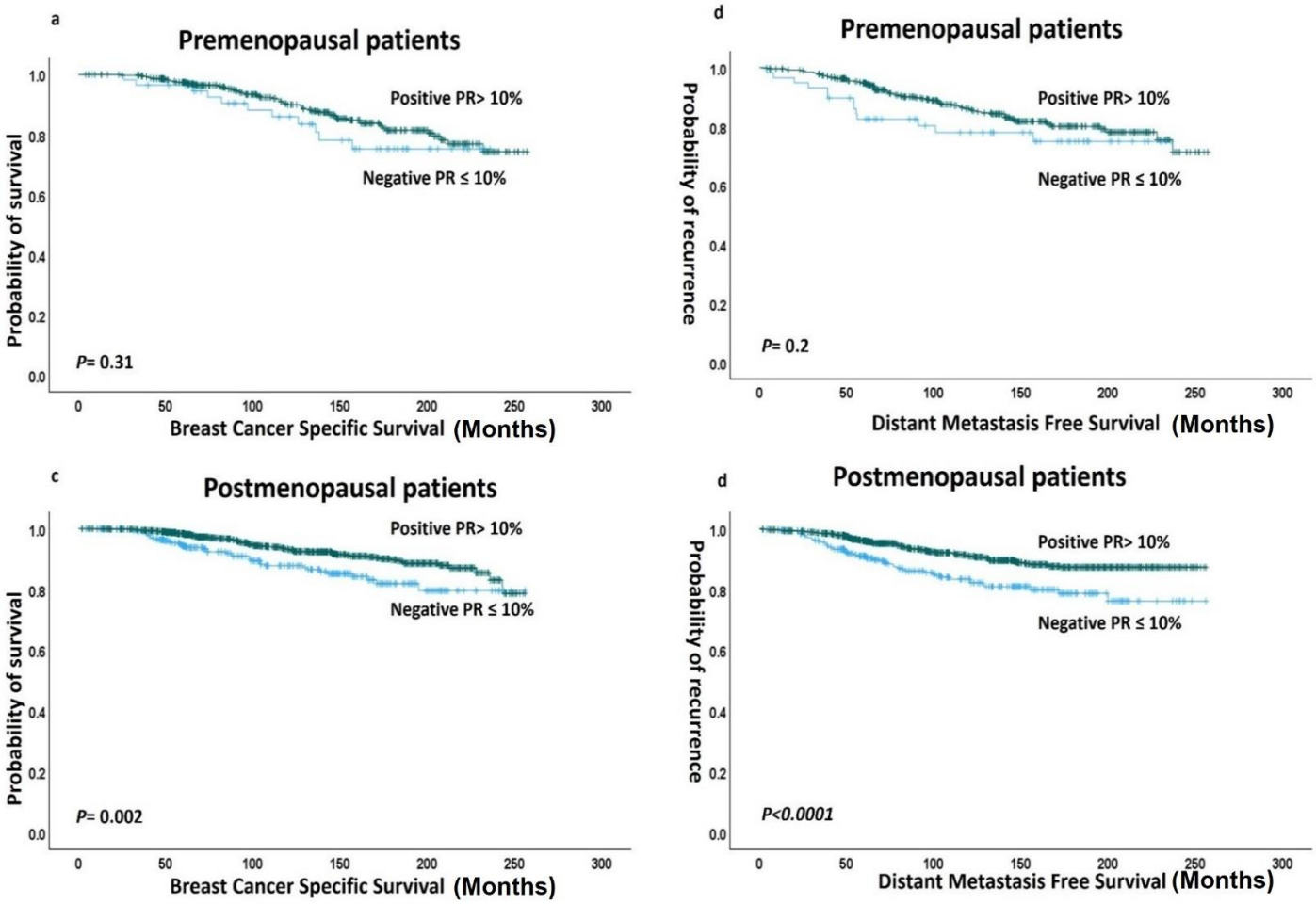


Figure 6-6: Kaplan Meier curves show significant associations between PR negativity and poor BCSS and DMFS in post-menopausal patients (c&d) while no association was observed in terms of pre-menopausal patients (a&b).

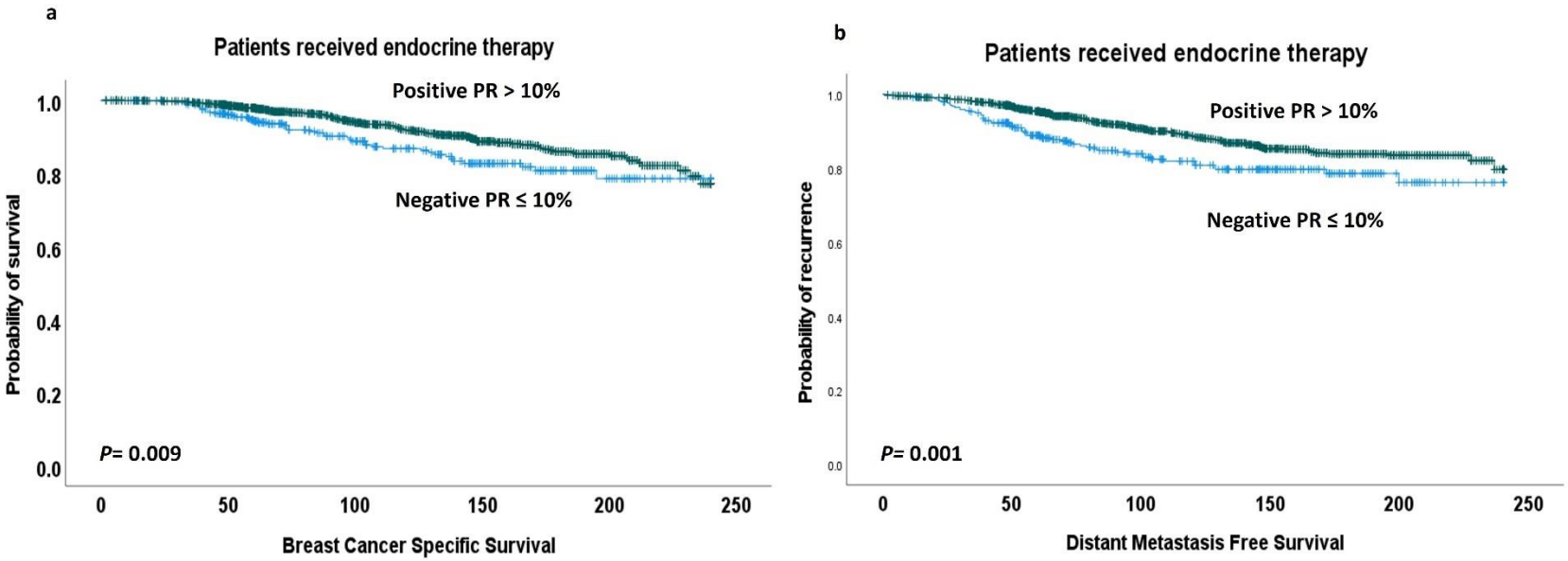


Figure 6-7: Kaplan Meier curves show significant association between PR negativity and poor BCSS (a) and DMFS (b) in patients who received endocrine therapy.

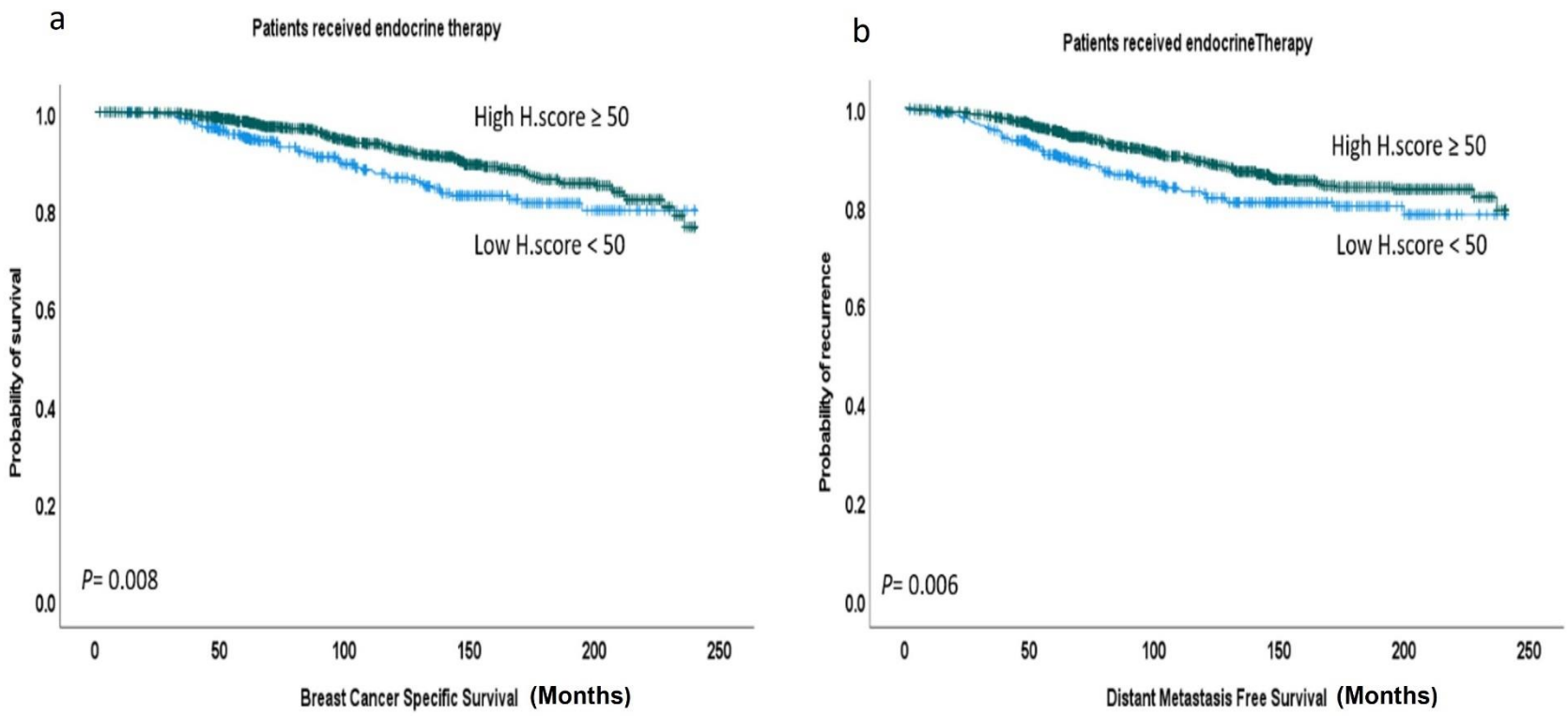


Figure 6-8: Kaplan Meier curves show significant association between low PR H. score and poor BCSS (a) and DMFS (b) in patients who received endocrine therapy.

6.3.2.4 Comparative analysis of eyeballing PR and Ki67 expression

In this study, Ki67 showed strong association with worse outcome in terms of shorter BCSS and DMFS in ER+, $HER2^-$ patients who received endocrine therapy (both $p<0.0001$). In the group of patients who received both endocrine and chemotherapy, there was no association between PR status and patient survival ($p=0.38$). Similar results were obtained with Ki67 ($p=0.8$) (**Table 43**).

In multivariate analysis, PR was a prognostic factor independent of histologic grade, tumour size, nodal status and Ki67 in the whole cohort. However, in the group of patients who received endocrine therapy, multivariate analysis included tumour size, grade and LN status, PR remained an independent prognostic factor ($p=0.009$) while Ki67 lost its significance ($p=0.093$) (

Table 44).

Moreover, when PR and Ki67 expression was compared in CNB, PR status showed independent prognostic significance in the group of patients who received endocrine therapy (p=0.039) while Ki67 lost its significance in the multivariate analysis (p=0.44) (**Table 45**).

When the cohort was stratified based on Ki67 expression, negative PR expression showed association with poor outcome in both the high and low proliferative groups (Ki67 positive and negative classes) in terms of both BCSS and DMFS (**Figure 6-9**).

Table 43: Comparison between the prognostic and predictive values of PR and Ki67 in univariate analysis in luminal BC.

Survival data	Number of patients	PR (negative/positive)		Ki67 (low/high)	
		Log rank	P-value	Log rank	P-value
BCSS in the whole cohort	1904	10.0	**0.004**	43.29	**<0.0001**
DMFS in the whole cohort		12.9	**<0.0001**	36.7	**<0.0001**
BCSS in patients received only endocrine therapy	1477	5.1	**0.020**	30.6	**<0.0001**
DMFS in patients received only endocrine therapy		8.1	**0.005**	29.86	**<0.0001**
BCSS in patients received both endocrine and chemotherapy	238	0.8	0.38	0.05	0.8
DMFS in patients received both endocrine and chemotherapy		1.6	0.20	0.1	0.75

BCSS, breast cancer specific cancer, DMFS, distant metastasis free survival. Significant p values are in **bold**.

Table 44: Multivariate analysis of Ki67 and PR in luminal BC using full face sections.

Parameters	BCSS in the whole cohort (n=1904)			BCSS in endocrine treated patients (n=1715)		
	Hazard ratio	95% (CI)	P-value	Hazard ratio	95% (CI)	P-value
PR	0.62	0.45-0.86	**0.005**	0.64	0.44-0.90	**0.009**
Ki67	1.50	1.10-2.10	**0.027**	1.40	0.90-2.01	0.093
Grade	1.90	1.01-2.00	**<0.0001**	2.10	1.60-2.92	**<0.0001**
Lymph node status	2.00	1.50-2.70	**<0.0001**	2.09	1.60-2.83	**<0.0001**
Tumour size	1.50	1.10-2.10	**0.006**	1.50	1.12-2.10	**0.01**

BCSS, breast cancer specific cancer, DMFS, distant metastasis free survival. 95% (CI): confident interval. Significant p values are in **bold**.

Table 45 : Multivariate analysis of Ki67 and PR in BR using core needle biopsies.

Parameters	BCSS in endocrine treated patients (n=727)		
	Hazard ratio	95% (CI)	P-value
PR	0.38	0.15-0.95	**0.039**
Ki67	0.67	0.24-1.96	0.44
Grade	2.4	1.1-5.60	**0.038**
Lymph node status	2.1	0.81-5.34	0.13
Tumour size	2.2	0.79-6.34	0.12

BCSS, breast cancer specific cancer, DMFS, distant metastasis free survival. 95%CI, 95% confidence interval. Significant p values are in **bold**.

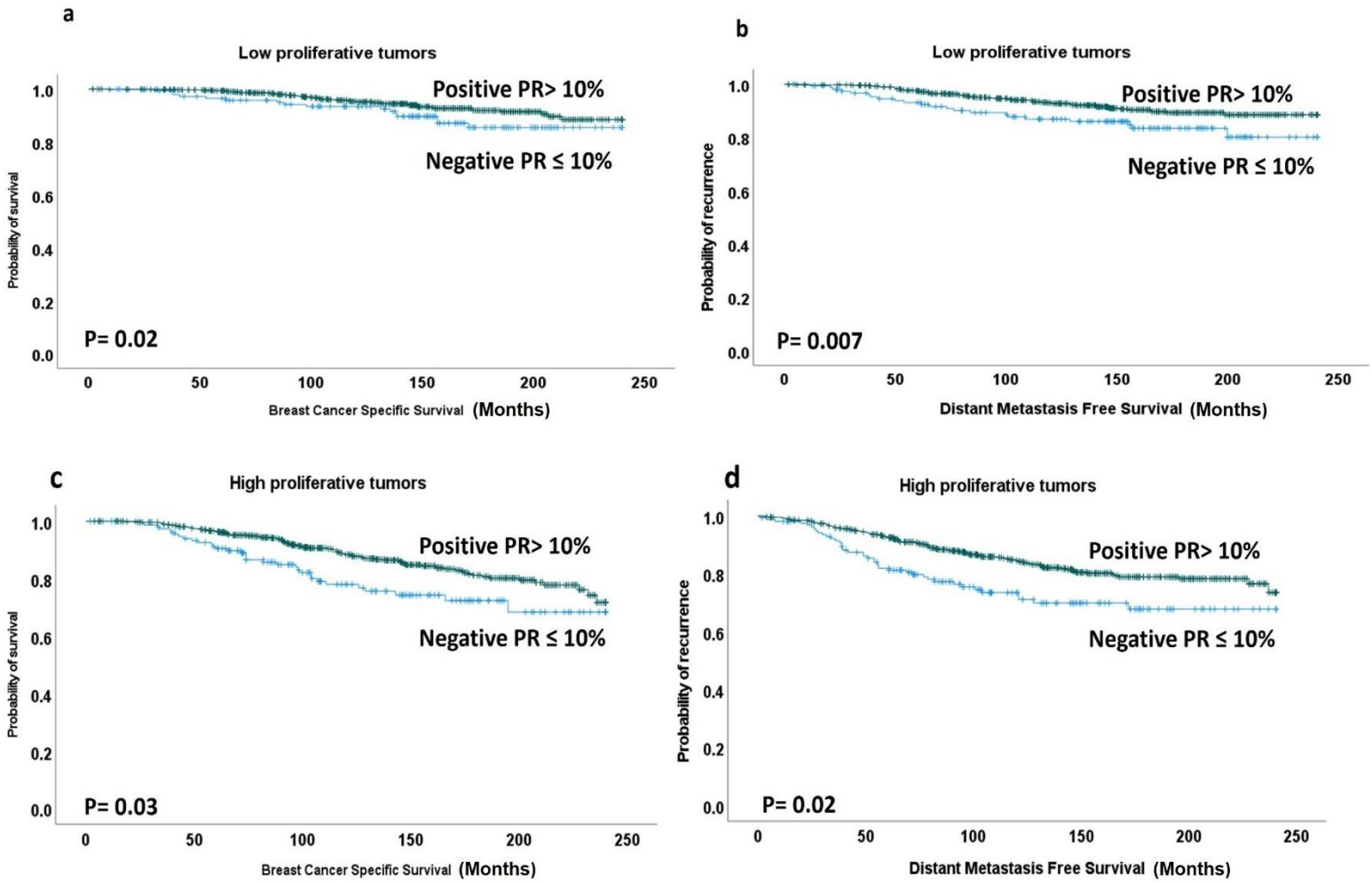


Figure 6-9: Kaplan Meier curves show a significant association between negative PR status and both BCSS and DMFS in low Ki67 expression tumours (a&b) as well as high Ki67 tumours (c&d).

6.3.3 Automated assessment of PR expression and clinicopathological parameters

The mean ± SD of automated PR expression was 48± 26%, median was 52% (range 0-100). Using 10% cutoff for the automated PR expression categorisation, negative PR status showed a significant association with older age patients >50 years (p=0.015), post-menopausal status (p=0.002), larger tumour size (p<0.0001), high tumour grade (p<0.0001), LVI (p=0.009) and poor NPI group (p<0.0001) (**Table 46**).

PR density was also calculated based on the number of clusters inside the areas which have high PR expression levels. 1393 cases were used to train the algorithm while 884 cases were used a validation set. The mean ± SD of PR density score was 6± 8, median was 3 (range 0-100). In contrast to PR percentage, high PR density showed a strong association with parameters characteristics of aggressive tumour behaviour including larger tumour size (p<0.0001), high tumour grade (p<0.0001), NST tumour type (p <0.0001), LVI (p<0.0001), LN metastasis (p<0.0001) and poor NPI group (p<0.0001) (**Table 47**).

Table 46: Relationship between automated PR status and clinicopathological parameters in luminal BC

Variables	Automated PR status	

	Negative (n, %)	**Positive** (n, %)	**X^2** **P-value**
Age at diagnosis			
<50	47 (8)	514 (92)	6.1
≥50	208 (12)	1508 (88)	**0.015**
Menopausal state			
Premenopausal	47 (8)	564 (92)	9.8
Post-menopausal	209 (12)	1484 (88)	**0.002**
Tumour size (cm)			
≤2	147 (9)	1535 (91)	35.9
>2	109 (18)	510 (82)	**<0.0001**
Histologic tumour grade			
Grade 1	34 (6)	509 (94)	88.2
Grade 2	110 (9)	1159 (91)	**<0.0001**
Grade 3	112 (23)	380 (77)	
Histologic tumour types			
No special type (NST)	176 (13)	114	18.8
Lobular	29 (11)		**<0.0001**
Other special types	13 (9)		
NST mixed	38 (7)	0 (87)	
		244 (89)	
		137 (91)	
		527 (93)	
Lymph node status			
Negative	183 (11)	1543 (89)	1.8
Positive	73 (13)	505 (87)	0.18
Lymhovascular invasion			
Absent	200 (10)	1731 (90)	6.86
Present	56 (15)	317 (85)	**0.009**
Nottingham prognostic index			
Good prognostic group	87 (7)	1187 (93)	55.1
Moderate prognostic group	148 (16)	781 (84)	**<0.0001**
Poor prognostic group	21 (21)	80 (79)	
Ki67 expression			
Low	64 (7)	833 (93)	24.1
High	178 (14)	1109 (86)	**<0.0001**

Table 47: Relationship between PR density and clinicopathological parameters in luminal BC

Variables	PR density		X^2
	Low (n, %)	**High** (n, %)	**P-value**
Age at diagnosis			
<50	203 (46)	243 (54)	4.4
≥50	764 (51)	728 (49)	**0.035**
Menopausal state			
Premenopausal	228 (46)	244 (54)	3.4
Post-menopausal	753 (51)	718 (49)	0.063
Tumour size (cm)			
≤2	862 (59)	609 (41)	173.88
>2	119 (24)	372 (76)	**<0.0001**
Histologic tumour grade			
Grade 1	321 (70)	136 (30)	280.1
Grade 2	599 (54)	503 (46)	**<0.0001**
Grade 3	61 (15)	343 (85)	
Histologic tumour types			
No special type (NST)	507 (45)	627 (55)	56.4
Lobular	137 (58)	98 (42)	**<0.0001**
Other special types	90 (78)	25 (22)	
NST mixed	247 (52)	232 (48)	
Lymph node status			
Negative	826 (54)	718 (46)	35.9
Positive	155 (37)	264 (63)	**<0.0001**
Lymhovascular invasion			
Absent	904 (53)	789 (47)	57.6
Present	77 (28)	193 (72)	**<0.0001**
Nottingham prognostic index			
Good prognostic group	758 (68)	364 (32)	331.03
Moderate prognostic group	218 (28)	568 (72)	**<0.0001**
Poor prognostic group	5 (9)	50 (91)	
Ki67 expression			
Low	483 (65)	262 (35)	116.2
High	438 (39)	676 (61)	**<0.0001**

6.3.4 Automated PR scoring and patient survival.

In terms of patient outcome, there was a significant association between negative PR expression and poor BCSS in the whole cohort as well as endocrine therapy treated patients (p<0.0001) (**Figure 6-10)**. However, high PR density had a significant association with short survival (p<0.0001). When the analysis was reported to the group of patients who received both endocrine therapy and chemotherapy, no association was detected between either automated PR score or PR density and patient survival, (p= 0.12 and 0.33, respectively).

In multivariate analysis, low automated PR score showed independent association with poor patient survival (p=0.004) and this association was maintained in patients who received endocrine therapy (p=0.018) (**Table 48**). High PR density showed independent association with shorter BCSS in endocrine therapy treated patients (**Table 49**). Multivariate analysis including eyeballing PR score, automated score and PR density showed that only PR density maintains the association with patient survival (p<0.0001) (**Table 50**).

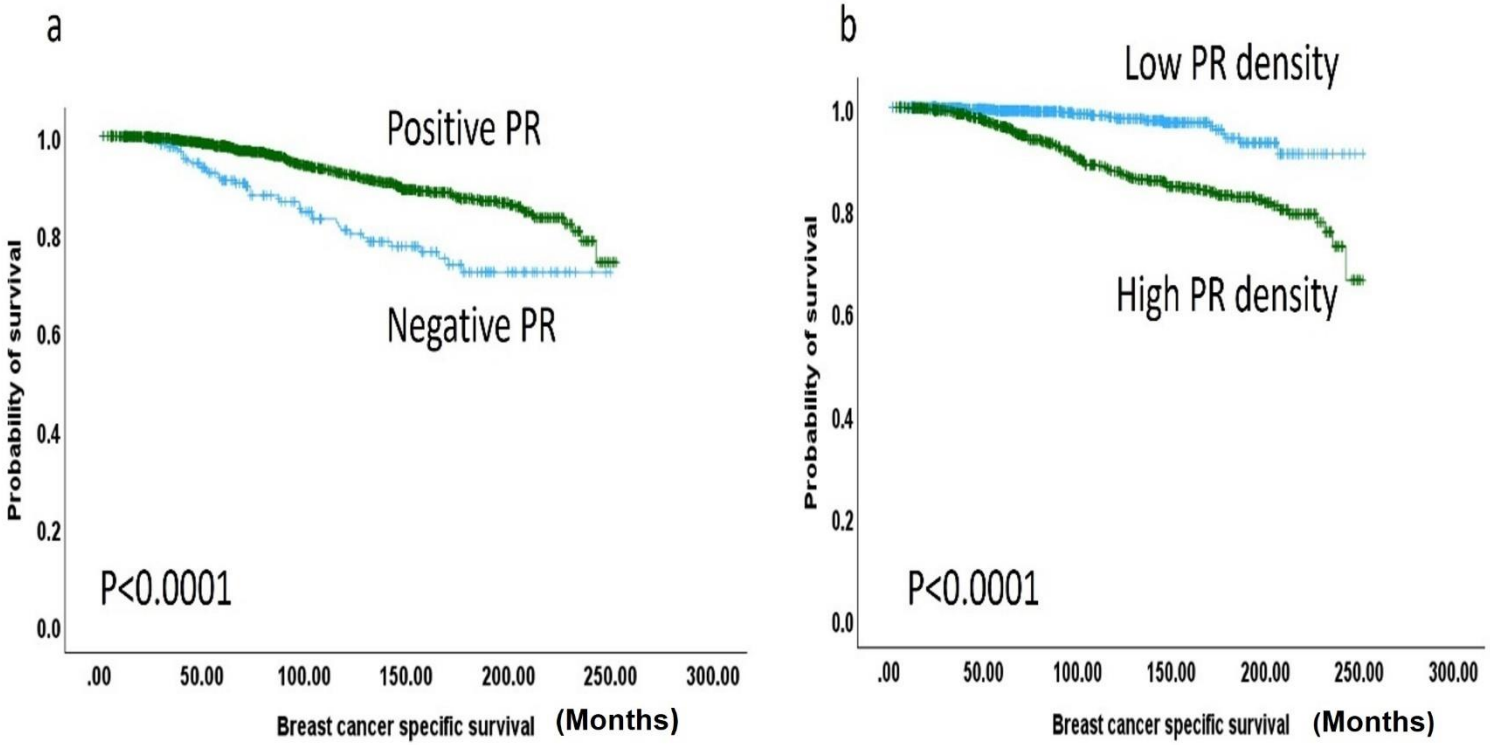


Figure 6-10: Kaplan Meier curves show a significant association between negative PR status using automated scores and short survival (a) while high PR density was associated with worse survival (b).

Table 48: Multivariate analysis of automated PR expression in endocrine therapy treated patients.

Parameters	BCSS		
	Hazard ratio	95% (CI)	P-value
Automated PR	0.63	0.43-0.93	**0.018**
Tumour size	1.69	1.1-2.2	0.44
Grade	1.4	1.8-3.1	**<0.0001**
Lymph node status	1.8	1.3-2.6	**<0.0001**

BCSS, breast cancer specific cancer. 95%CI, 95% confidence interval, Significant p values are in bold

Table 49: Multivariate analysis of PR density in endocrine therapy treated patients.

Parameters	BCSS		
	Hazard ratio	95% (CI)	P-value
PR density	2.6	1.6-4.4	**<0.0001**
Tumour size	1.2	0.82-1.7	0.39
Grade	2.5	1.8-3.4	**<0.0001**
Lymph node status	1.7	1.2-2.4	**0.003**

BCSS, breast cancer specific cancer. 95%CI, 95% confidence interval, Significant p values are in bold

Table 50: Multivariate analysis of eyeballing PR scores, automated PR scores and PR density in endocrine therapy treated patients.

Parameters	BCSS		
	Hazard ratio	95% (CI)	P-value
Eyeballing PR	0.66	0.36-1.2	0.17
Automated PR	0.74	0.38-1.4	0.36
PR density	4.6	2.6-8.1	**<0.0001**

HR, hazard ratio, BCSS, breast cancer specific cancer. 95%CI, 95% confidence interval, Significant p values are in bold

6.4 DISCUSSION

Approximately 75% of BC are positive for ER and/or PR. The ER-positive tumours express ER, PR, ER responsive genes and other genes that encode typical proteins of luminal epithelial cells, so they are termed the luminal group [187]. Adjuvant hormonal therapy is the main systemic treatment for the majority of hormone receptor-positive early-stage BC thus. In luminal BC, a comprehensive individual assessment should include many factors including ER, PR, HER2, and Ki67 [509, 510]. PR positivity is seen in 65-75% of BC, and in approximately 80-90% of ER positive cases. However, this proportion varies depending on the cutoff used for defining PR positivity [511]. In addition, the predictive significance of PR expression in BC has showed controversy for decades, and there has been equivocal data about the role of PR expression in early-stage BC [305, 512]. Currently, many physicians prefer to discriminate low-risk from high-risk BC tumours by using multigene assays, including Oncotype DX test rather than using surrogate histopathological-based classification [513]. Although, these tests provide prognostic information and some of them are approved to

guide treatment decisions, these tests are still too expensive to be afforded by the institutions in many countries. This project shown that PR can categorise luminal BC tumours into low risk and high-risk groups to help the decision of patient's treatment. This study included a large cohort of luminal BC and full-face sections were used to assess PR. PR expression is usually assessed using immunohistochemistry following the same standard and quality assurance measurements applied for ER staining. Therefore, using PR can provide an additional quality assurance method for the accuracy of ER status [514]. In addition, PR is one of the main component genes in many well-established multigene prognostic signatures [257].

As several cutoffs are reported in the literature, this chapter aimed to evaluate and compare different cutoffs to reach an optimal cutoff that can be reliably used for dichotomisation of luminal BC patients into 2 prognostic groups.

At the beginning of this study, testing the optimal cutoff of PR using X tile was carried out where 75% positivity was the best cutoff for prognostic purposes using full face sections while in CNB, 25% cutoff was the best one. Since this cutoff was considered very high in the context of routine practice for consistency with ER, three commonly used cutoffs (1%, 10% and 20%) were also tested. 10% cutoff achieved the best association with patient outcome in both full-face sections and in CNB. Similarly, a previous study showed that the differences in patients` outcome were most significant using a cutoff of 10% [500]. These findings suggest that

in order to incorporate PR in decision of patient's therapy, 1% cutoff is not sufficient but 10% could be optimal in ER^+/HER2 BC. Furthermore 10% cutoff would be consistent with recent ASCO/CAP guidelines of ER expression categorisation into low and high groups [487]. In this study, PR showed a bimodality of expression in BC, which suggest that PR positivity in BC is an all-or-none phenomenon. Similarly, many studies reported bimodal expression of PR as well as ER [508, 515, 516].

Low PR expression was significantly associated with post-menopausal status and two studies also showed similar finding [501, 517], suggesting that PR expression affected by the menopausal status. Therefore, PR was analysed separately in subgroups of premenopausal and postmenopausal women. It was found that low PR had association with poor outcome in postmenopausal patients. Additionally, negative PR expression showed strong association with high-grade tumours which can be an indication of aggressive tumour behaviour. Previous studies have shown PR loss in ER+ tumours resulted from the low circulating levels of oestrogen in those patients in addition to the methylation the in PR promoter gene [518] or due to aberrant growth factor signalling [501, 519]. While others hypothesise that the absence of PR expression means non-functional ER as PR is found to be the end product of oestrogen action [305, 520]. These mechanisms of negative PR expression could explain the highly aggressive characteristics of tumours lacking PR expression. Consequently, ER-positive/PR-negative breast tumours might be treated by completely blocking ER action using Selective estrogen receptor modulators (SERMs) via oestrogen withdrawal using

aromatase inhibitors, by targeted ER degradation using Selective Estrogen Receptor Degraders (SERDs), or by combined therapy targeting both ER and growth factor signalling pathways [521].

In the current study, negative PR expression showed strong association with poor outcome. Similarly, in previous studies, low PR expression showed association with poorer prognosis in ER positive, HER2-negative BC, independently of other variables [302, 305, 497, 522, 523]. Also, ER-positive/PR-negative BC has a similar prognosis to hormone receptor-negative BC and ER+/PR- tumours, as defined by gene expression profiles, share gene expression patterns with ER-/PR tumours [299].

Ki67 expression has important prognostic role in BC and with combination of PR, they could stratify luminal BC into distinct prognostic groups [524]. In this study, high Ki67 shows a significant association with poor outcome in BC patients, however, negative PR expression showed independent poor prognostic factor and inclusion, or exclusion of grade had an impact on the association of Ki67 with outcome but not affects the prognostic significance of PR. This indicates that Ki67 works as complementary of grade and PR works independently in BC. Also, the PR status of proliferating cancer cells reflect the biology of luminal-type BC.

Additionally, loss of PR predicts endocrine resistance, and this indicates that PR expression seems to be a key marker of prediction of endocrine response. Previously, it has been reported that PR is a powerful,

independent prognostic variable in patients receiving endocrine therapy in early BC and absence of PR expression should be re-evaluated as a biomarker for poor prognosis and considered as a key for additional systemic therapy [525]. In this study, negative PR expression in patients who revived chemotherapy, showed improved outcome similar to patients who had positive PR in the same group of patients. Similar findings were shown regarding high Ki67 expression. A clinical trial proved that the addition of chemotherapy added benefit in patients with low or absent PR expression [516]. Also, a previous report showed that although there was a significant association between high Ki67 and poor outcome, no association with outcome was observed in endocrine therapy and chemotherapy treated patients [526]. In addition, it has been shown that high Ki67 is indicative of a greater benefit from chemotherapy when added to endocrine therapy in ER+/HER2- BC [527].

Therefore, these results suggest that adding chemotherapy to PR negative and or high Ki67 cases can improve the outcome in luminal BC patients.

In multivariate analysis, the lack of PR showed poor outcome in patients who received endocrine therapy, and this suggests that PR expression could predict the response of endocrine therapy. A previous paper also showed that PR negativity can influence the therapeutic decision to offer adjuvant chemotherapy in addition to adjuvant endocrine therapy in selected patients [528]. However, others claim that PR expression cannot predict the benefit of endocrine efficacy in ER-positive BC in clinical trials

and they showed that benefit from tamoxifen is similar in both PR-positive and PR-negative patients [300, 394]. Similar to the current study, a previous study concluded that low proliferative tumours lacking PR should be classified as a high-risk group for developing recurrence [492].

Not only can the status of PR predict the response of adjuvant therapy but also, the intensity of expression has an important role; low PR H. score showed a significant association with worse outcome in patients who received endocrine therapy. This suggests that the addition of intensity to percentage of PR expression can be used to provide confirmatory information about the level of clinical response to endocrine therapy. Similarly, ER H. score levels are associated with different patient outcome [529]. However, calculating H. score of PR for each patient may be challenging in the routine practice.

Intra-tumoural heterogeneity of ER and PR expression is commonly reported in BC [495, 530]. 80% concordance per tumour was found that between a full section of resected tumour and core biopsy for PR expression [516, 531]. In this study, there was a significant association between both biopsies and the concordance rate was 88%, however, about 2 % of negative PR tumours in CNB were PR positive >10% in full face sections. This might have resulted from substantial heterogeneity of PR staining in BC or due to variation in pre-analytic variables, but in the view of these low percentage of discrepant cases, PR assessment in CNB is sufficient for the prognostic and predictive value.

Subjectivity of PR scoring may also hinder its value and automated scoring methods could eliminate this problem and improve diagnostic accuracy and patient care [532]. In this study, PR expression levels were assessed using DL models. It was found that negative expression of PR using automated methods had significant associations with aggressive tumour parameters and poor outcome. Notably, PR density was a novel prognostic parameter in this study and its value outweighed the value of both manual and automated PR expression.

Conclusion

These findings highlighted that PR optimal cut off should be 10% instead of 1%. PR provides prognostic and predictive value independent on the proliferation activity in ER positive tumours. PR has potential benefit to identify candidate patients who might benefit from more aggressive adjuvant therapy. PR density score assessed using AI tools can be a novel predictor of patient outcome in luminal BC.

Chapter 7 Identification of potential prognostic markers in BC

7.1 Introduction

As BC is the second leading cause of cancer-related death in women worldwide, the identification of new clinical biomarkers in BC is crucial to further predict prognosis, metastases, and therapeutic response [533-535]. BC associated biomarkers comprise macromolecules, such as nucleic acid (DNA/RNA), proteins, and intact cells [534, 536]. Identifying biomarkers may solve the problem of drug resistance which is a challenging obstacle in BC treatment [534, 536]. Moreover, an increasing percentage of patients are demanding personalised therapy, that requires the development of novel biomarkers for diagnostic and prognostic procedures [537, 538]. The markers that have been studied in this chapter were selected from the panel of gens associated with atypical mitosis. Linker histone H1 is one of the important members of histone family which displays both transcriptional and genetic alterations in a variety of cancers [314, 539, 540]. Linker histone, also known as histone H1 (H.H1), interacts with DNA and involved in the formation and maintenance of higher-order chromatin structure [316, 541, 542]. Decreased H.H1 expression is a common feature of neoplastic cells in lymphoma, where it acts as a tumour suppressor [321, 322]. However, high H.H1 expression is associated with malignant potential of prostate cancer and considered as a marker of cellular proliferation [323].

In addition, *Polo-like kinases 1* (*PLK1*) is a key components of PLK family [326]. *PLK1* plays a role in DNA replication during S phase, and it acts as a modulator of the DNA damage response and maintenance

of genome stability during replication [340, 543]. PLK1 is reported to be overexpressed (at both mRNA and protein level) in many tumours compared to the normal tissue counterpart [544, 545]. Although, *PLK1* is considered a potential oncogene and *PLK1* inhibitors have recently been developed as potential anticancer agents [329, 330]. Recent data showed that down regulation of *PLK1* can induce aneuploidy and tumorigenesis [546]. Elevated PLK1 expression is associated with vascular invasion and high proliferation in basal-like BC [339, 547].

Enabling replicative immortality and uncontrolled cell cycle are hallmarks of cancer cells and minichromosome maintenance proteins (MCMs) play vital roles in controlling replication times within a cell cycle [548]. MCM complex presents a critical first step in formation of pre-replication complex at the initiation stage of DNA replication [343]. The MCM complex is composed of six highly conserved MCM proteins, namely MCM2-7 [549, 550]. Several MCM proteins are highly associated with tumorigenesis, and they can also predict tumour progression [344, 345]. In addition, MCMs are candidate markers for cellular proliferation and increased levels of MCMs indicate proliferation of malignant cells [551]. Although, MCM2 and MCM3 were significantly upregulated in BC [552, 553], the expression levels, functions, and prognostic values of MCM7 in BC need to be explored. *MCM7* plays an important role in initiating DNA replication and proliferation of cancer cells [346-348]. High *MCM7* expression was associated with a negative outcome in patients with non-small cell lung carcinoma [554]. Therefore,

inactivation of this *MCM7* was hypothesised to be a promising therapeutic target in various types of cancer [344].

7.1.1 Hypothesis and aims.

This chapter hypothesised that H.H1, PLK1 and MCM7 have important clinical value in BC and their role will differ among the different molecular subtypes of BC. Therefore, the aim of this chapter is to assess the expression of H.H1 and its prognostic and predictive values in BC, b) evaluate the clinical significance of PLK1 in BC and its molecular classes, and c) assess the expression of MCM7 and explore its clinical value in BC. All markers were assessed at both proteomics and transcriptomic levels.

7.2 Patients and Methods:

7.2.1.1 Proteomic analysis

For the assessment of H.H1, PLK1 and MCM7 protein expression, a large cohort of primary BC (n = 1311) from patients who presented at Nottingham City Hospital, Nottingham, UK was utilised. Tissue samples were prepared as a TMA as previously described in chapter 2.4 [555]. The clinical and tumour characteristics, including histologic tumour type, tumour grade, and LVI, were available **(Table 51).**

Table 51: Clinicopathological characteristics of the study cohort

Variables	No (%)
Tumour size (cm)	
≤2cm	794 (61)
>2cm	517 (39)
Tumour grade	
Grade 1	204 (16)
Grade 2	505 (39)
Grade 3	602 (45)
Tumour types	
No Special Type (NST)	855 (65)
Lobular	118 (9)
Other special types	54 (4)
NST mixed	284 (22)
Molecular subtypes	
Luminal	1041 (79)
HER2 enriched.	64 (5)
Triple negative	194 (16)
Lymph node status	
Negative	809 (62)
Positive	502 (38)
Lymhovascular invasion	
Absent	934 (71)
Present	377(29)
Nottingham prognostic index	
Good prognostic group	437 (33)
Moderate prognostic group	659 (50)
Poor prognostic group	215 (17)

7.2.1.2 Transcriptomic analysis

H.H1, PLK1 and MCM7 mRNA expression were analysed at the transcriptomic level using publicly available cohorts (n=11359) from the bc-GenExMiner database (all DNA microarray data) and the Kaplan Meier (KM) Plotter (n=3951) online dataset (https://kmplot.com/analysis/) [556].

7.2.2 H.H1, PLK1 and MCM7 immunostaining

The specificity of anti-H.H1 mouse monoclonal antibody (Clone r1415-1, Thermo Fisher Scientific, UK), anti-PLK1 rabbit polyclonal antibody (ab109777, Abcam, UK) and anti-MCM7 rabbit polyclonal antibody (# PA5-79651, Invitrogen, Thermo Fisher Scientific, UK) were validated by WB using cell lysates of MCF7, MCF10A, MCF10 DCIS, and SKBR3 human BC cell lines obtained from American Type Culture Collection, Rockville, MD, USA. The H.H1 primary antibody was used at 1:500 dilution. While PLK1 and MCM7 primary antibodies were used at 1:1000 dilution. Anti-β-actin primary antibody (Sigma-Aldrich) was used as a loading control.

Three series of full-face sections of BC cases (n=14) representative of several molecular subtypes were used to evaluate the distribution of H.H1, PLK1 and MCM7 expression. Tissue sections were stained using the Novocastra Novolink™ Polymer as described in detail in chapter 2.5.3. H.H1 antibody was diluted at 1:100 and incubated for 30 minutes at room temperature. PLK1 was diluted at 1:250 and incubated for 30 min at room temperature. MCM7 was diluted at 1:500 and incubated for 30 min at room temperature. The IHC-stained slides were digitally scanned using a high-throughput scanner (Panoramic 250 Flash III; 3D-Histech, Budapest, Hungary), and the WSIs were viewed with CaseViewer software (version 2.2.0.85; 3D-Histech).

7.2.3 H.H1, PLK1 and MCM7 protein expression evaluation

H.H1 showed nuclear expression and the percentages of positive tumour cells were calculated. A semiquantitative evaluation of nuclear H.H1 expression was used to calculate a modified histochemical score (H-score), and the final H. score was obtained by giving a range of 0 to 300 [390, 450, 557, 558]. To assess the inter-observer concordance, a second observer (Nabeelah Almalki) scored 20% of the cohort, and the ICC was calculated. H. scores of 170 was set as the optimal cutoff of H.H1 expression using X tile.

PLK1 immunoreactivity was observed predominantly in the cytoplasm and the percentage of positive tumour cells, and H. score were calculated. For assessing the inter-observer concordance, a second observer (Dr Michael Toss) scored 20% of the cohort. Using X-tile software, H. score of 170 was the optimal cutoff to categorise PLK1 expression into low and high groups and then PLK1 was correlated with clinicopathological parameters and patient outcome.

MCM7 showed both nuclear and cytoplasmic expressions. The percentage of positive tumour cells and H. score were evaluated. However, H. score did not show any association with clinicopathological data or patient outcome. Therefore, the percentage of MCM7 expression was adopted in the final analysis. For assessing the inter-observer concordance, a second observer (Dr Michael Toss) scored 20% of the cohort. MCM7 expression levels were categorised into low and high expression based on X tile software. The generated cutoffs were 5% and

95% for the nuclear and cytoplasmic expression of MCM7, respectively. MCM7 expression was correlated with the clinicopathological parameters and patient outcome.

7.2.4 Statistical analysis

Statistical Package for the Social Sciences software v.27.0 (SPSS, Chicago, IL, USA) was used for statistical analysis. X-tile bioinformatics software version 3.6.1 (School of Medicine, Yale University, New Haven, CT, USA) was used to define the optimum cutoff for each marker [380]. While the mRNA levels were categorised using the optimal splitting criterion in the online website. Chi-square test was used for analysis of categorical data. Outcome analysis was assessed using Kaplan Meier curves and the log-rank test. The survival analysis for MCM7 was based on 10-year survival. Cox regression models was used for the multivariate analysis. Estimated HR and their 95% confidence interval (95% CI) were calculated. For all tests, $P < 0.05$ (two-tailed) were statistically significant.

7.3 Results

7.3.1 H.H1 expression in BC

7.3.1.1 *H.H1* mRNA expression

There was a significant association between high *H.H1* mRNA levels and NST tumour type, and basel-like BC (both p<0.0001) (**Figure 7-1**). Additionally, high *H.H1* mRNA levels was significantly associated with poor outcome in the whole cohort of bc-GenExMiner data (p=0.04) and basel-like BC (p=0.001) (**Figure 7-2**).

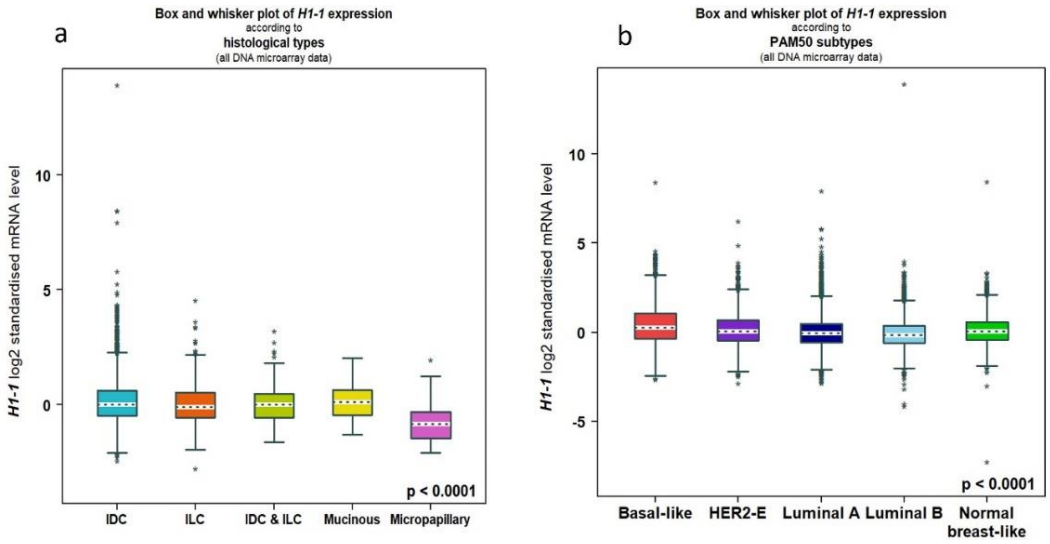


Figure 7-1: Box and whisker plots show a significant association between high *H.H1* mRNA expression and NST tumour type (a) while (b) shows that *H.H1* mRNA levels were significantly higher in basal-like BC.

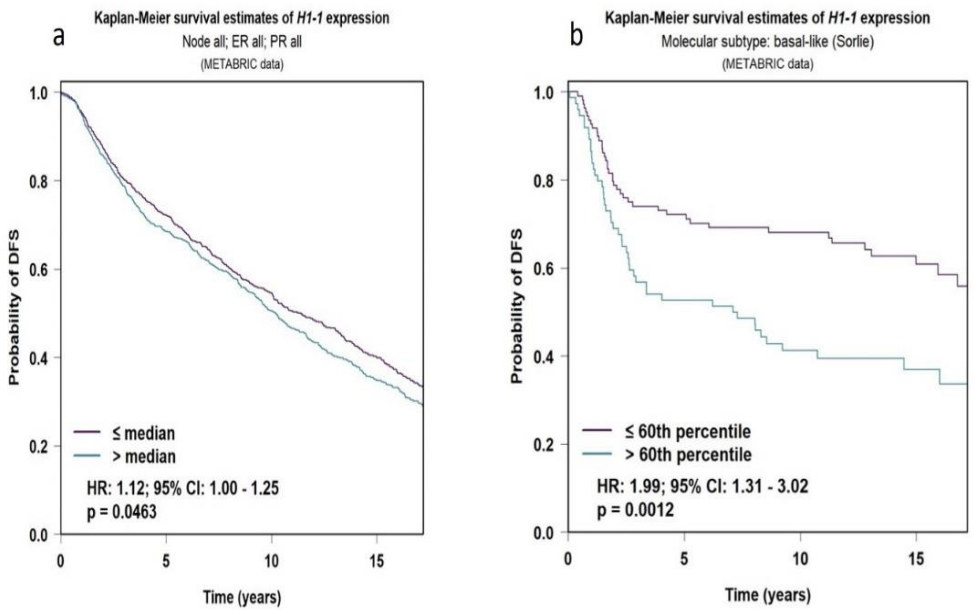


Figure 7-2: Kaplan Meier curves show significant associations of high *H.H1* mRNA expression and poor outcome in the whole METABRIC cohort (a) and basal-like BC (b).

7.3.1.2 H.H1 protein expression

Specific bands for H.H1 were observed in western blot at the predicted molecular weight, 30 kilodaltons (KDa) confirming the specificity of the H.H1 antibody (**Figure 7-3**). H.H1 expression showed both nuclear and nucleolar subcellular localisation (**Figure 7-4**). There was almost perfect interobserver agreement in the H.H1 scoring (K=0.86). The mean ± SD of H.H1 H. score was 134± 64, median was 120 and ranged from 0-300. High H.H1 expression (H-score >170) was observed in 28% (372/1311) of the patients. High H.H1 expression showed a significant association with features characteristic of aggressive cancer behaviour including high tumour grade, NST tumour type, TNBC and poor NPI group (all p<0.0001) (**Table 52**).

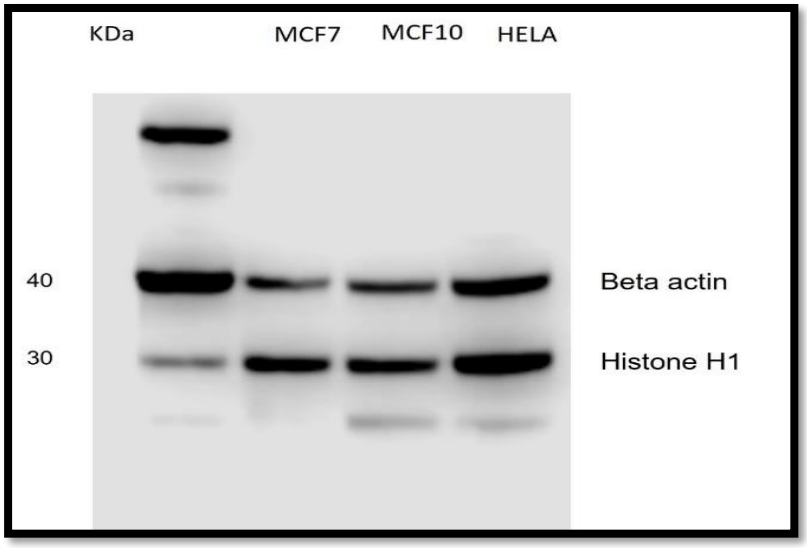


Figure 7-3: Western blot of H.H1 antibody shows specific bands at 30 KDa and B actin at 42 KDa.

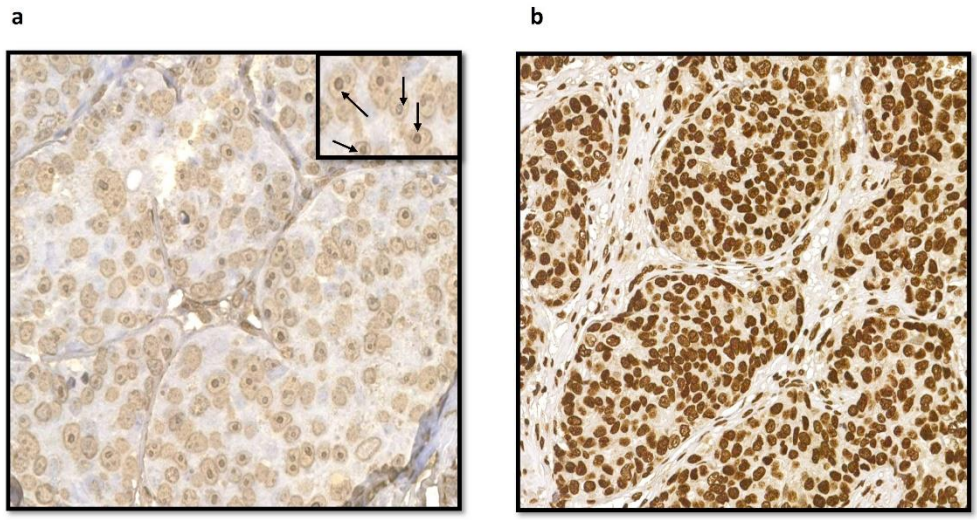


Figure 7-4*:* H.H1 protein expression in BC
a- mild nuclear and nucleolar expression (black arrows) of H.H1, b- strong expression of H.H1 (20x)

Table 52: Relationship between H.H1 expression and clinicopathological parameters in BC

Categories	Low H.H1 expression	High H.H1 expression	X^2 *P*-value
Age (years)			
<50	272 (67)	136 (33)	7.15
≥50	667 (74)	236 (26)	**0.007**
Tumour size			
≤ 2cm	583 (73)	211 (27)	3.231
> 2cm	356 (69)	161 (31)	0.073
Tumour grade			
Grade 1	187 (92)	17 (8)	76.339
Grade 2	383 (76)	122 (24)	**<0.0001**
Grade 3	369 (61)	233 (39)	
Histologic types			
No special type (NST)	554 (65)	301 (35)	57.723
Lobular	100 (85)	18 (15)	**<0.0001**
Other special types	49 (91)	5 (3)	
Mixed NST	236 (83)	48 (17)	
Molecular subtypes			
Luminal	792 (76)	249 (24)	43.954
Triple negative	109 (56)	85 (44)	**<0.0001**
HER2 enriched	34 (56)	30 (34)	
Lymph node status			
Negative	582 (72)	227 (28)	0.104
Positive	357 (71)	145 (29)	0.747
Lymphovascular invasion			2.227
Absent	680 (73)	254 (27)	0.136
Present	259 (69)	118 (31)	
Nottingham Prognostic index			
Good prognostic group	356 (82)	81 (18)	32.118
Moderate prognostic group	445 (68)	214 (32)	**<0.0001**
Poor prognostic group	138 (63)	77 (37)	

Significant p values are shown in **bold**.

7.3.1.3 H.H1 protein expression and patient outcome

In terms of patient outcome, high H.H1 expression was significantly associated with shorter BCSS and DMFS (p=0.006 and p=0.01), respectively (**Figure 7-5).** When the whole cohort was classified according to the molecular classes, high H.H1 expression was strongly associated with a shorter BCSS in TNBC (p=0.007). Such association was not observed in luminal or HER2 enriched subgroups (**Figure 7-6)**.

When the association between high H.H1 expression and endocrine therapy in luminal BC was investigated, there was a significant association between high H.H1 expression and poor prognosis in patient who did not receive endocrine therapy (p=0.001). In contrast, no association was observed in the endocrine therapy treated patients (p=0.603) (**Figure 7-7**). Looking at the TNBC molecular subtype, a significant association was found between high H.H1 expression and shorter survival in patients who did not receive chemotherapy (p=0.006); however, there was no significant difference was observed in patient outcome between low and high H.H1 expression in those group patients who received chemotherapy (p=0.43) (**Figure 7-8**).

In a multivariate Cox regression analysis including tumour size, tumour grade, and LN status, high H.H1 expression was an independent predictor of poor outcome (p=0.029) (**Table 53**).

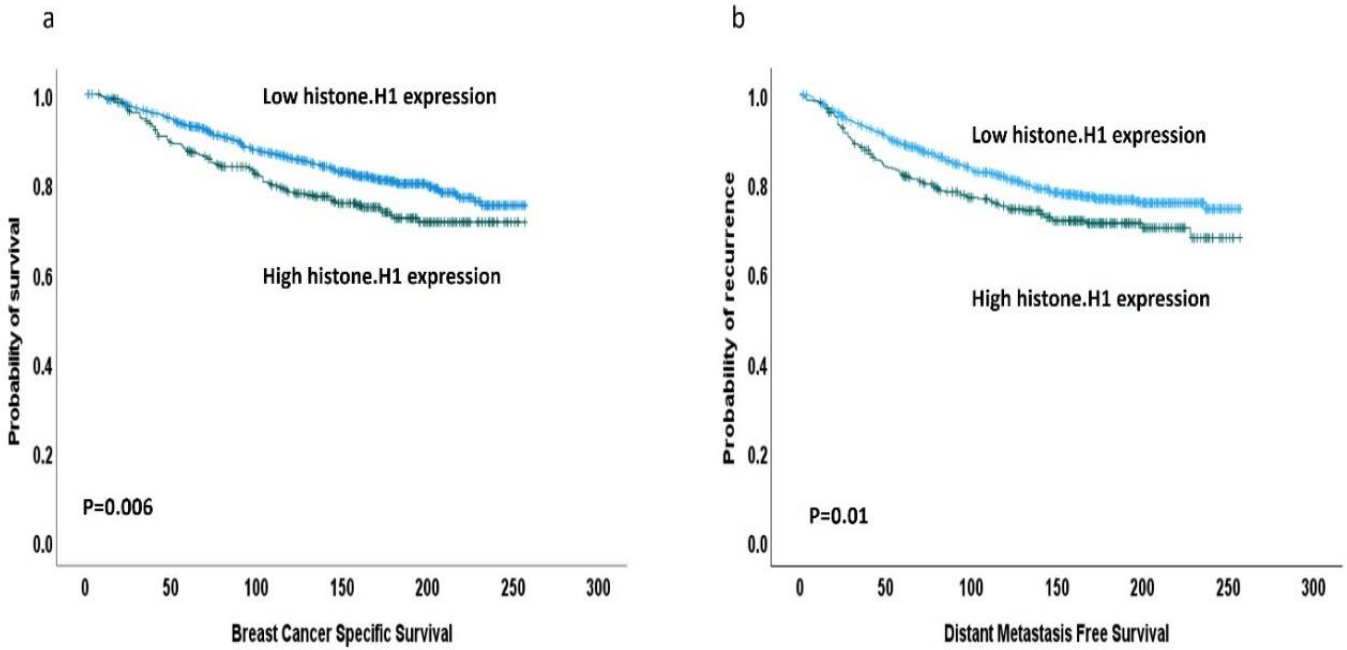


Figure 7-5 : Kaplan Meier curves show the association of high H.H1 expression and poor BCSS (a) and shorter DMFS (b).

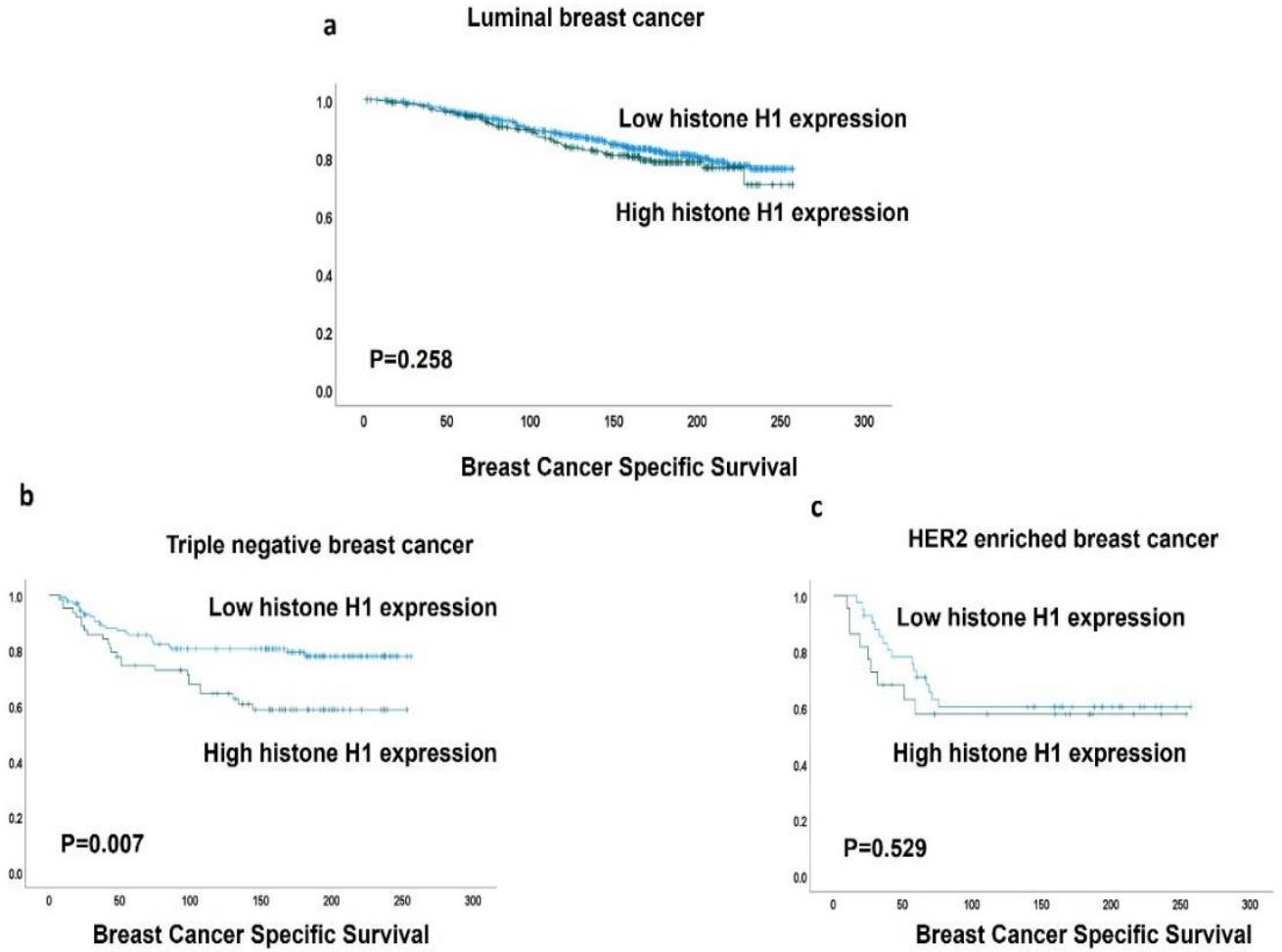


Figure 7-6: Kaplan Meier curves show the significant association of H.H1 and BCSS in TNBC (b) while no association was observed in luminal or HER2 enriched BC (a&c).

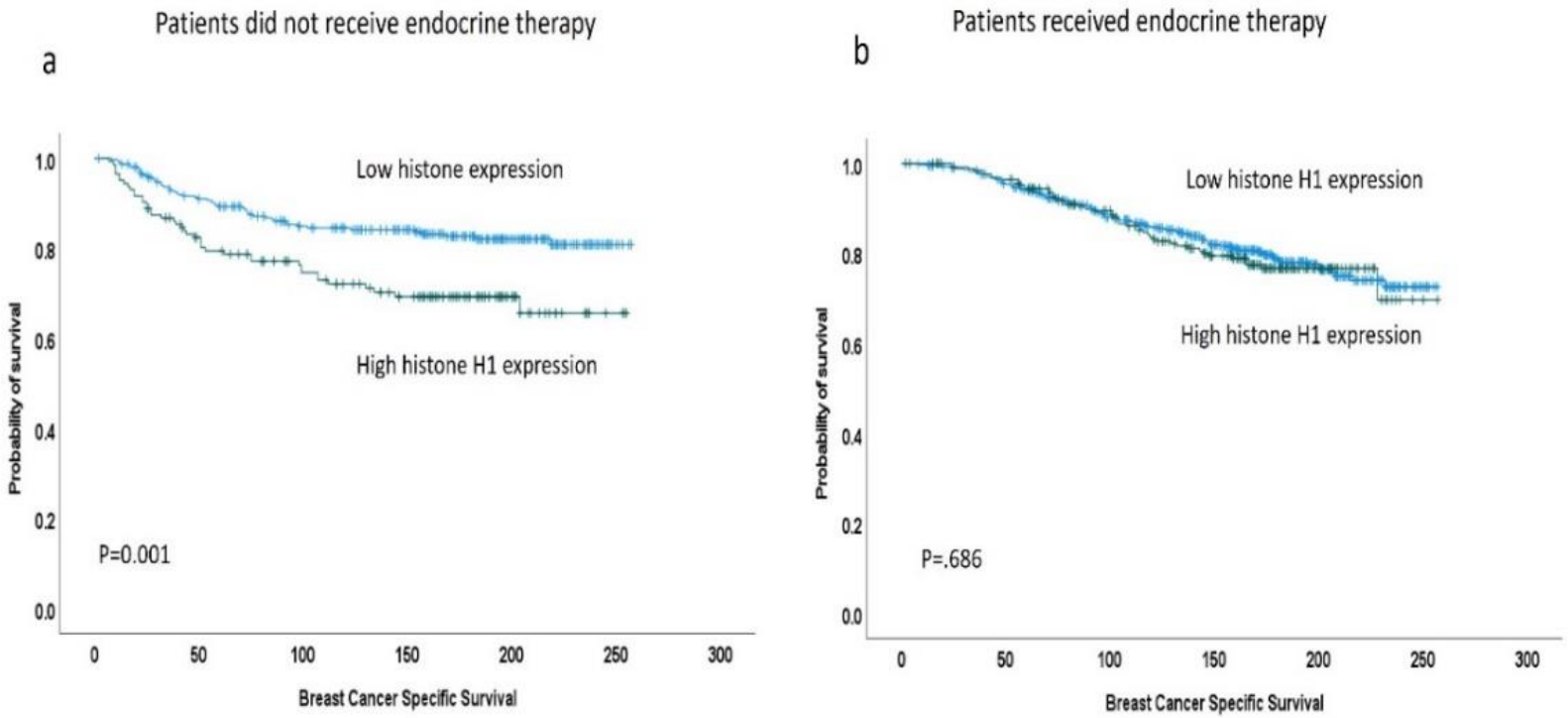


Figure 7-7: Kaplan Meier curves show the association of H.H1 and short BCSS in endocrine therapy naïve luminal BC patients (a) but, no association was observed in endocrine therapy treated patients (b).

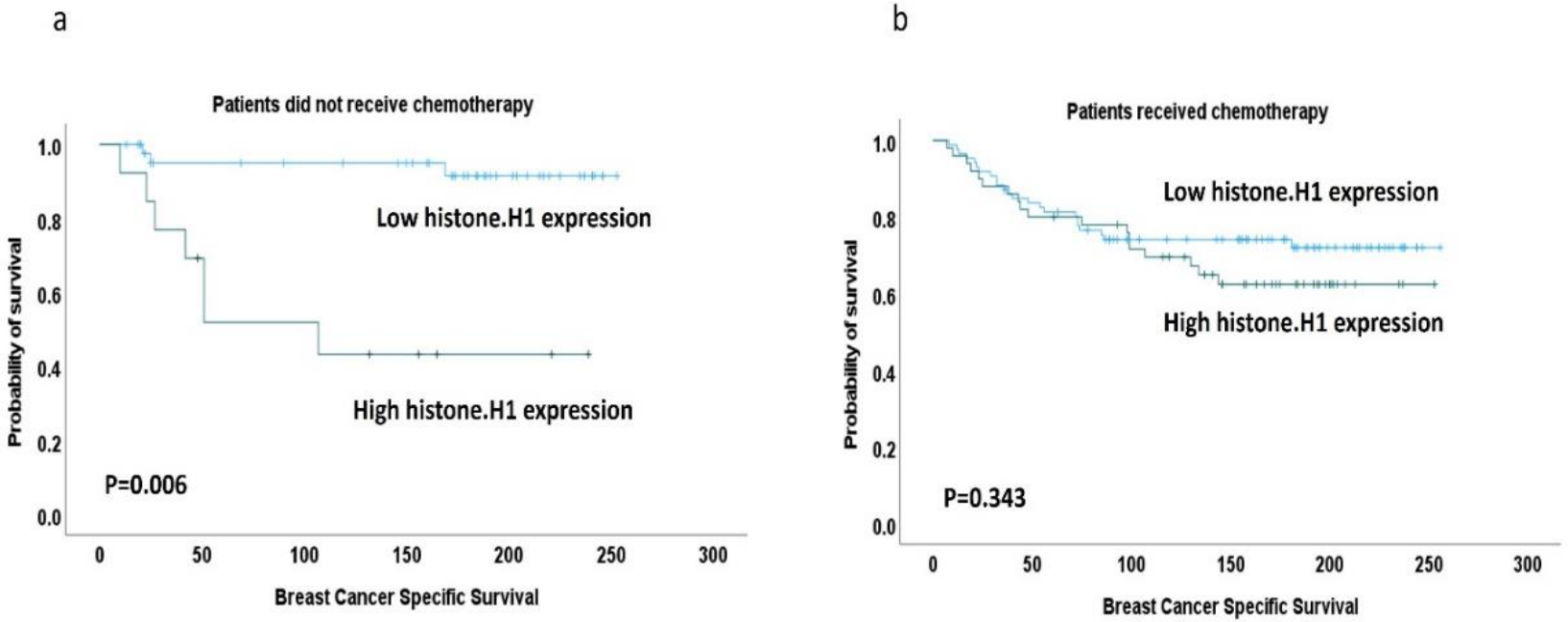


Figure 7-8: Kaplan Meier curves show the association of H.H1 and shorter BCSS in TNBC patients who didn't receive chemotherapy (a) but, no association was observed in chemotherapy treated patients (b).

Table 53: Multivariate analysis of H.H1 expression in BC

Parameters	BC specific survival		P value
	Hazard ratio	95% (CI)	
H.H1 expression	1.3	1.1-1.7	**0.029**
Grade	2.0	1.6-2.6	**<0.0001**
Lymph node status	1.6	1.4-2.2	**<0.0001**
LVI	1.8	1.4-2.3	**<0.0001**
Tumour size	1.5	1.1-2.0	**0.002**

95%CI, 95% confidence intervals. Significant p values are in bold.

7.3.2 PLK1 in BC

7.3.2.1 *PLK1* mRNA expression

PLK1 mRNA expression levels showed a significant association with NST tumour type compared to other tumour types (p<0.0001). It was also associated with basel-like BC (p<0.0001) (**Figure 7-9)**. Low *PLK1* mRNA expression levels showed a significant association with poor outcome in the whole cohort (*p*=0.0004) while in TNBC, high *PLK1* expression had a significant association with poor BC outcome (*p*= 0.01) using the Kaplan Meier Plotter dataset, (

Figure 7-10).

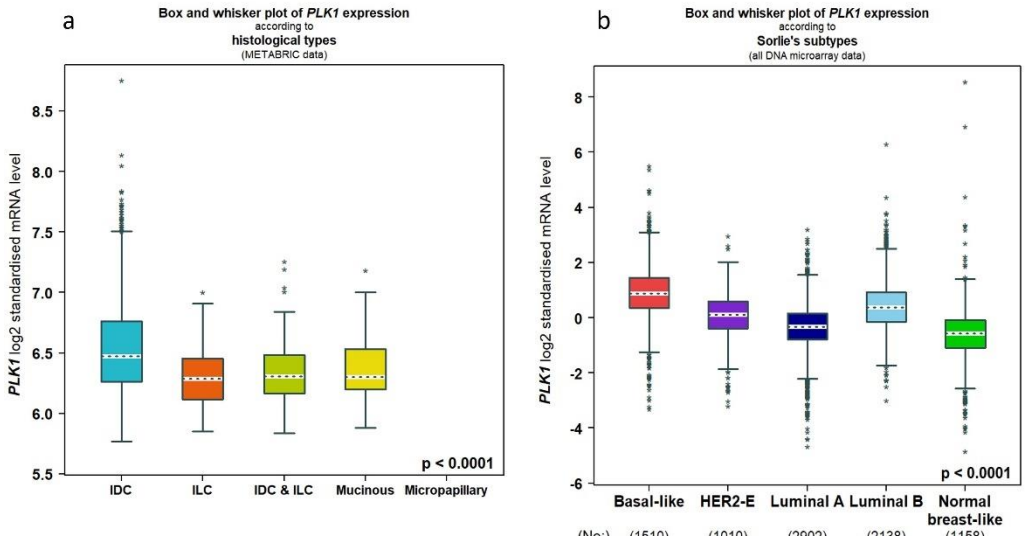


Figure 7-9: Box and whisker plots showing significantly higher *PLK1* mRNA levels in NST tumour (a) and in basal-like BC (b).

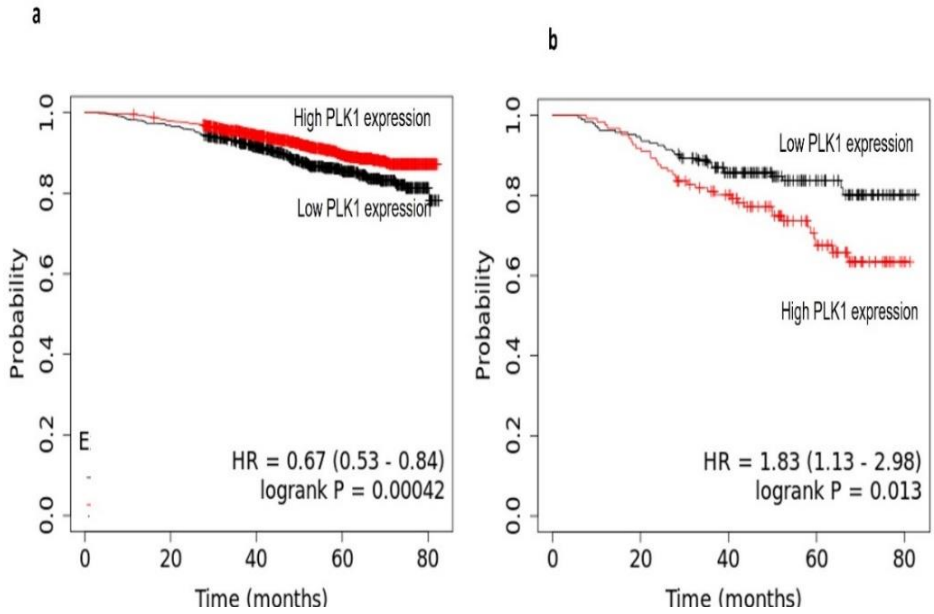


Figure 7-10: Kaplan Meier curves show association between *PLK1* mRNA levels and patient survival using Kaplan Meier Plotter cohort.

High *PLK1* shows a significant association with longer survival in the whole cohort (a) while (b) shows the association between high *PLK*1 expression and shorter survival in basal-like BC.

7.3.2.2 PLK1 protein expression

Specific band for PLK1 was observed in western blot at the predicted molecular weight, 60 KDa confirming the specificity of PLK1 antibody (**Figure 7-11)**. PLK1 showed cytoplasmic sub-cellular localisation and its expression was detected in 73% (886/1208) of tumours (**Figure 7-12)**. The mean ± SD of PLK1 H. score was 151± 80, median was 100, (range 0-300). There was excellent inter-observer concordance in scoring of PLK1 H. scores (K =0.85). Low PLK1 was observed in 86% (1046/1208) of patients while 14% showed high PLK1.

There was a significant association between high PLK1 expression and TNBC while low expression was associated with the HER2 enriched

subtype (p=0.020). Although, there was no association between PLK1 and the overall grade, a significant association was observed between high PLK1 expression and high mitosis scores (p=0.045). There was no association between PLK1 expression and tumour size, LN invasion or LVI (**Table 54**).

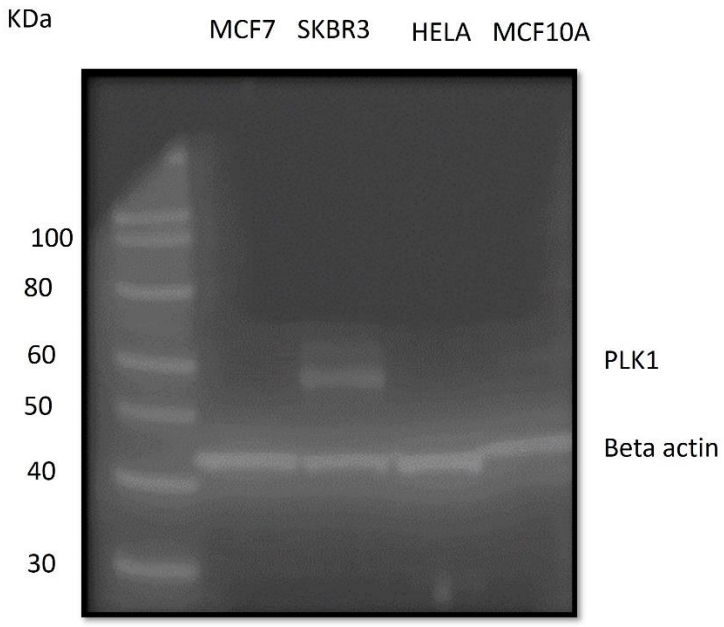


Figure 7-11: Western blot of PLK1 antibody shows a specific band at 60 KDa while B-actin shows specific bands at 42 KDa.

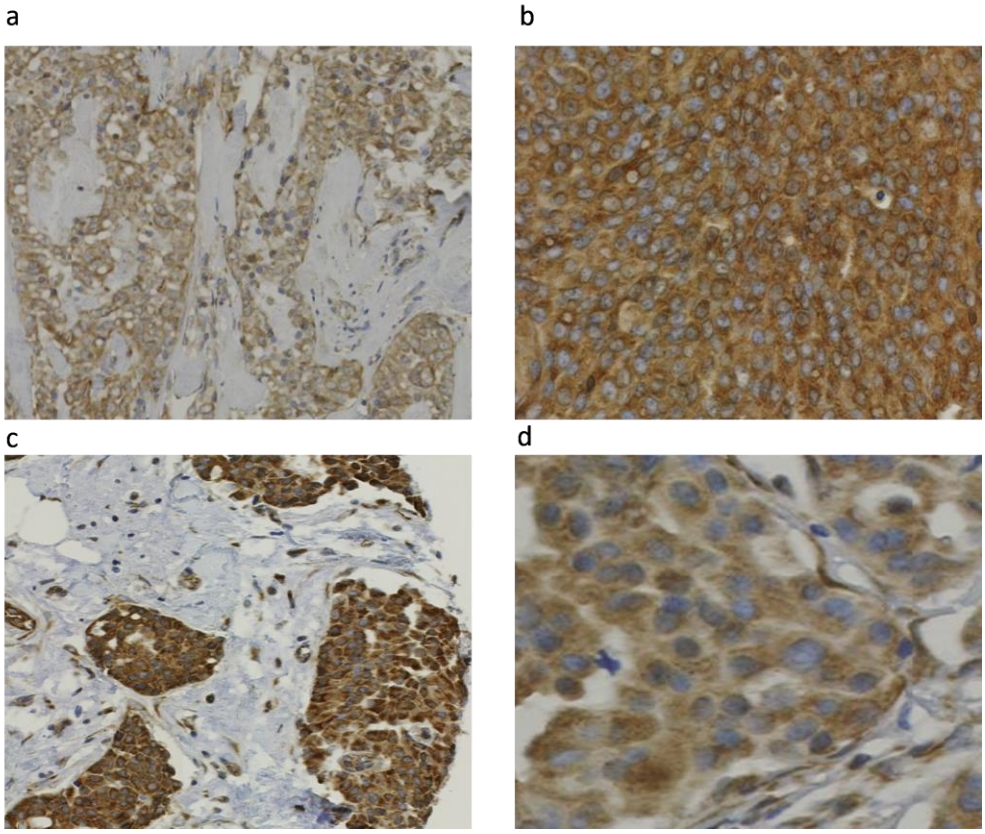


Figure 7-12: PLK1 expression in BC
a- Mild expression, b- moderate expression, c- strong expression (20x), d- high power view shows cytoplasmic PLK1 expression (40x).

Table 54: Relationship between PLK1 expression and clinicopathological parameters

Variables	PLK1 expression		X^2
	Low expression	High expression	*P*-value
Age at diagnosis (years)			
<50	300 (78)	82 (22)	1.2
≥50	671 (81)	155 (19)	0.3
Menopausal state			
Premenopausal	327 (78)	94 (22)	3.1
Post-menopausal	644 (82)	143 (18)	0.08
Tumour size (cm)			
≤2 cm	576 (79)	154 (21)	2.5
>2 cm	395 (83)	83 (17)	0.1
Histologic tumour grade			
Grade 1	144 (80)	35 (20)	0.01
Grade 2	384 (81)	93 (19)	0.9
Grade 3	443 (80)	109 (20)	
Histologic tumour types			
No special type (NST)	635 (81)	147 (19)	
Lobular	77 (80)	19 (20)	2.1
Other special types	32 (73)	12 (27)	0.5
NST mixed	277 (80)	59 (20)	
Molecular subtypes			
Luminal	791 (81)	188 (19)	10.5
HER2 enriched.	52 (93)	4 (7)	**0.005**
Triple negative	119 (73)	43 (27)	
Lymph node status			
Negative	580 (80)	143 (20)	0.03
Positive	391 (81)	94 (19)	0.8
Lymhovascular invasion			
Absent	687 (80)	173 (20)	0.47
Present	284 (82)	64 (18)	0.49
Nottingham prognostic index			
Good prognostic group	314 (79)	83 (21)	4.7
Moderate prognostic group	490 (81)	117 (19)	0.6
Poor prognostic group	167 (82)	37 (18)	

7.3.2.3 PLK1 protein expression and patient outcome

Low PLK1 expression showed a significant association with worse BCSS and DMFS in the whole cohort (*p*=0.032 and *p*=0.049, accordingly) (**Figure 7-13**). With classification the whole cohort into different molecular subtypes, low PLK1 expression was associated with shorter BCSS in both luminal and HER2 enriched subtypes (*p*=0.030 and 0.045, respectively). While in TNBC, high PLK1 expression showed a significant association with poor outcome (*p*=0.026) (**Figure 7-14**).

In endocrine therapy treated patients, low PLK1 expression showed a significant association with poor outcome (*p*=0.02) while high PLK1 expression showed association with worse BCSS in TNBC patients who received chemotherapy (*p*=0.036) (**Figure 7-15**). Multivariate Cox regression analysis showed that low PLK1 expression was an independent risk factor of poor prognosis in luminal BC (**Table 55**). On the other hand, high PLK1 expression was an independent poor prognostic risk factor in TNBC, with loss of the prognostic power of tumour grade and tumour size in those group of patients (**Table 56**).

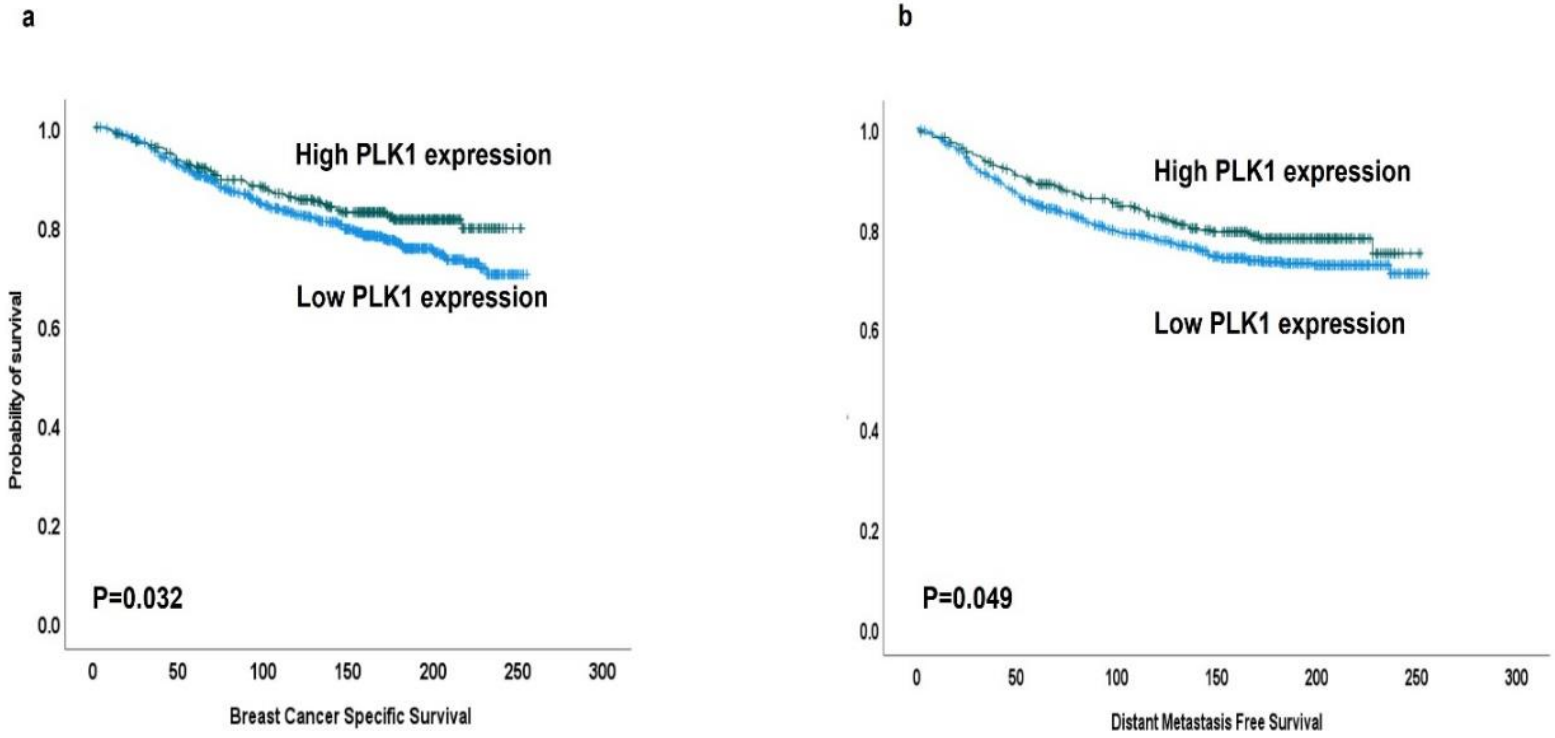


Figure 7-13: a- Kaplan Meier curves show the association of PLK1 expression with BCSS (a) and DMFS (b).

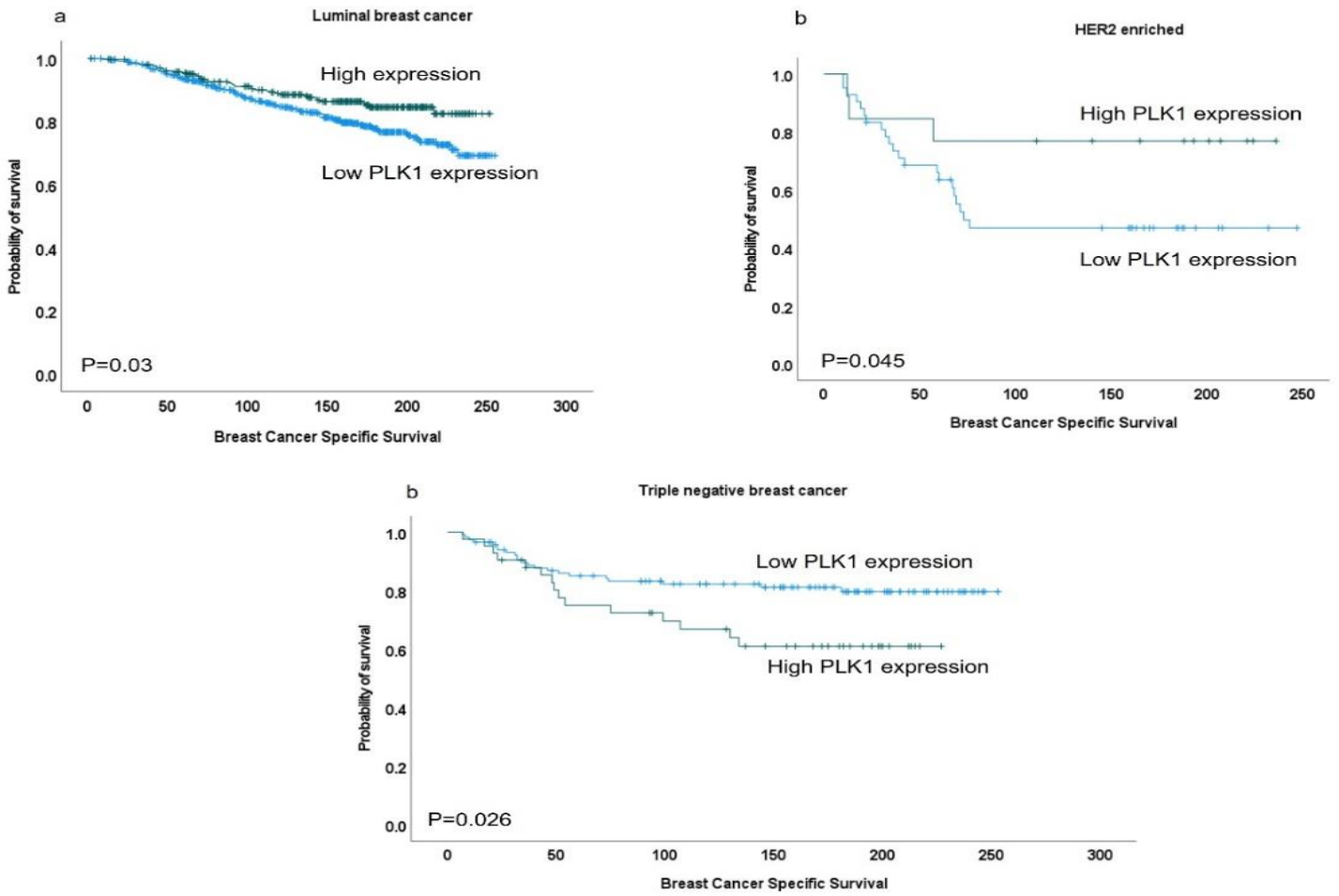


Figure 7-14: Kaplan Meier curves show association of low PLK1 with shorter survival in luminal and HER2+ BC (a&b) while high PLK1 expression shows association with poor outcome in TNBC (c).

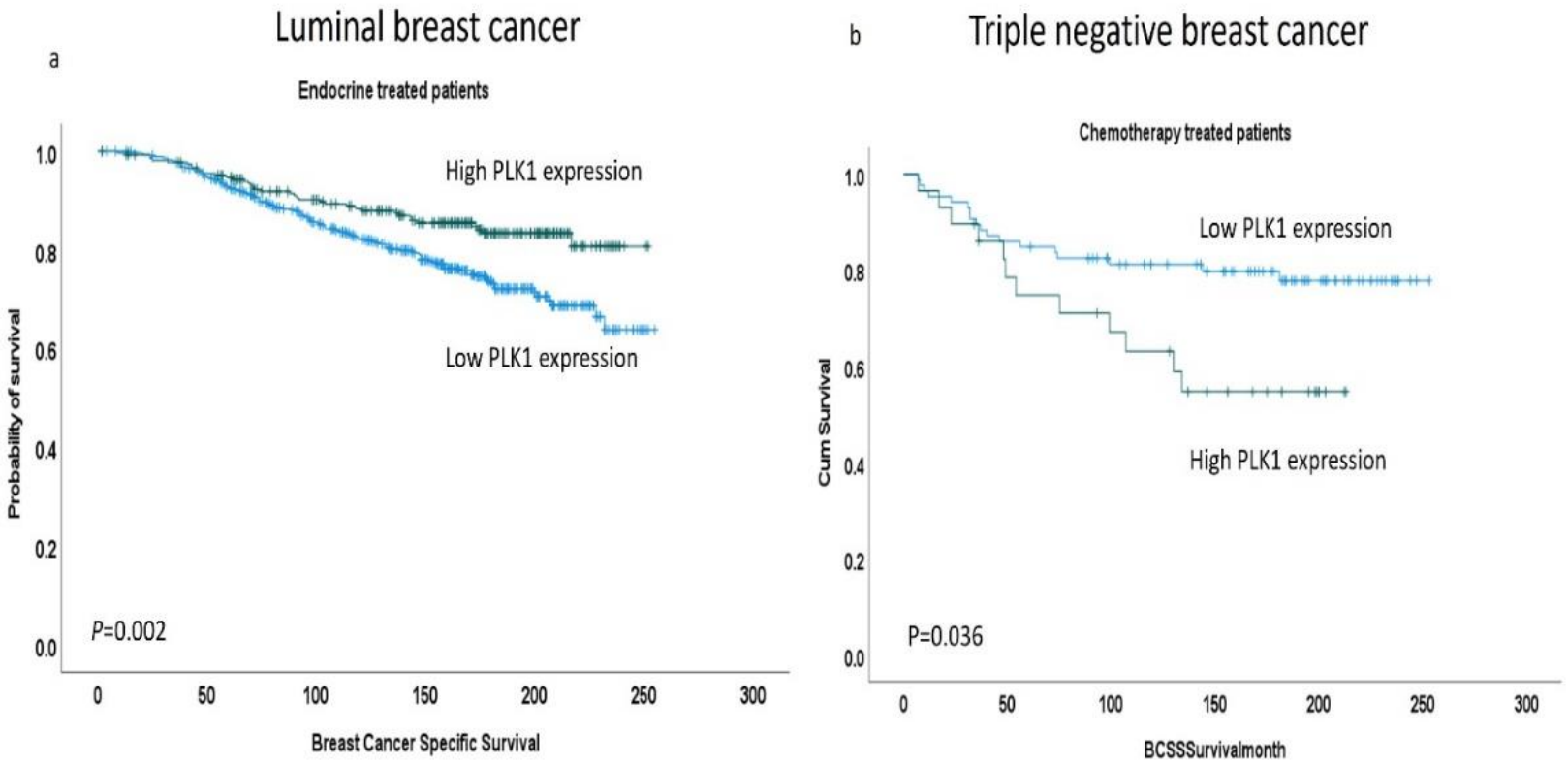


Figure 7-15: Kaplan Meier curves show the association between PLK1 and adjuvant therapy
Low PLK1 had a significant association with poor BCSS in luminal BC patients who received endocrine therapy (a) while b showed a significant association between high PLK1 and poor outcome in TNBC patients who received chemotherapy

Table 55: Multivariate analysis PLK1 expression in luminal BC

Parameters	BCSS			DMFS		
	Hazard ratio	95% (CI)	P-value	Hazard ratio	95% (CI)	P-value
PLK1	0.57	0.40-0.81	**0.002**	0.6	0.48-0.8	**0.006**
Grade	1.6	1.1-2.2	**<0.0001**	2.0	1.6-2.6	**<0.0001**
Lymph node status	1.8	1.4-2.1	**<0.0001**	1.8	1.5-2.2	**<0.0001**
Tumour size	1.6	0.4-0.8	**0.002**	1.7	1.3-2.3	**<0.0001**

BCSS, breast cancer specific cancer, DMFS, distant metastasis free survival. 95% confidence interval. Significant p values are in **bold.**

Table 56: Multivariate analysis PLK1 expression in TNBC

Parameters	BCSS			DMFS		
	Hazard ratio	95% (CI)	P-value	Hazard ratio	95% (CI)	P-value
PLK1	1.7	0.8-3.3	0.10	1.9	1.0-3.7	**0.04**
Grade	1.2	0.5-3.2	0.61	1.4	0.5-3.9	0.49
Lymph node status	2.4	1.6-3.7	**<0.0001**	1.8	1.2-2.8	**0.006**
Tumour size	1.1	0.6-2.2	0.65	1.3	0.7-2.4	0.06

BCSS, breast cancer specific cancer, DMFS, distant metastasis free survival. 95%CI, 95% confidence interval. Significant p values are in **bold.**

7.3.3 MCM7 in BC

7.3.3.1 MCM7 mRNA expression

There was a significant association between high *MCM7* mRNA levels expression and NST tumour type and basal-like BC (both p<0.0001) (**Figure 7-16)**. In addition, there was a significant association between high *MCM7* mRNA levels and poor survival in the whole cohort and luminal BC (both p<0.0001) using BC Gene Ex-miner (**Figure 7-17**). Similar findings were observed in the whole cohort and luminal BC using the Kaplan Meier plotter cohort. However, no association was found between *MCM7* mRNA levels and patient outcome in basal-like BC in either the Gene ex miner data or the Kaplan Meir plotter cohort (**Figure 7-18**).

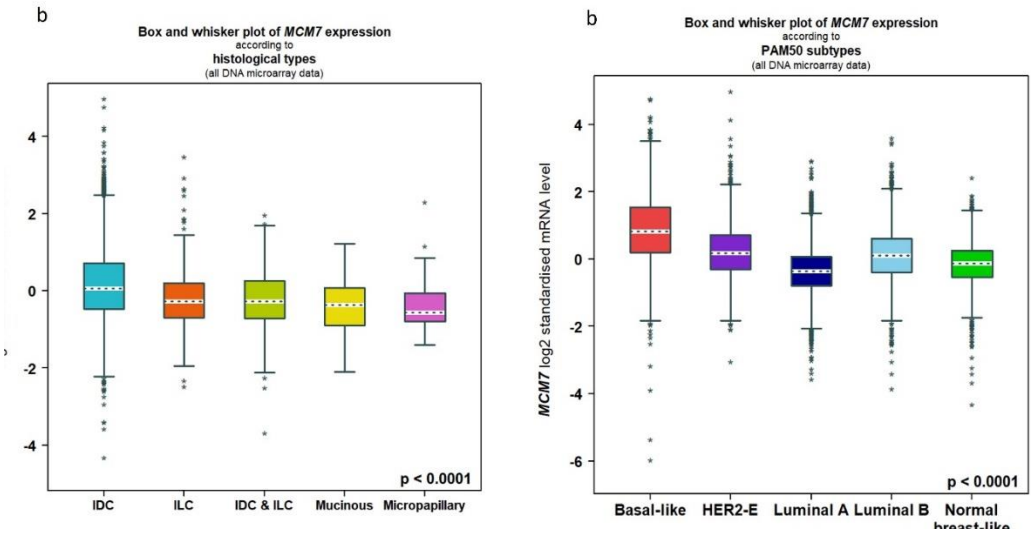


Figure 7-16: Box and whisker plots showing significantly higher *MCM7* mRNA expression in NST tumour type (a) and basal-like BC in gene ex miner data.

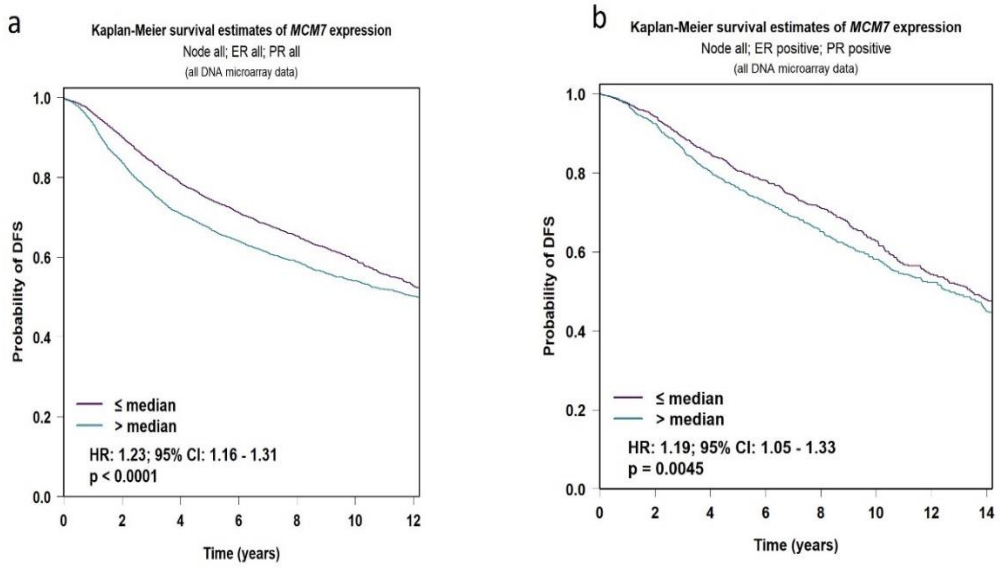


Figure 7-17: Kaplan Meier curves showing a significant association between high *MCM7* mRNA levels and poor outcome in the whole cohort (a) and luminal BC in gene ex miner data. (b).

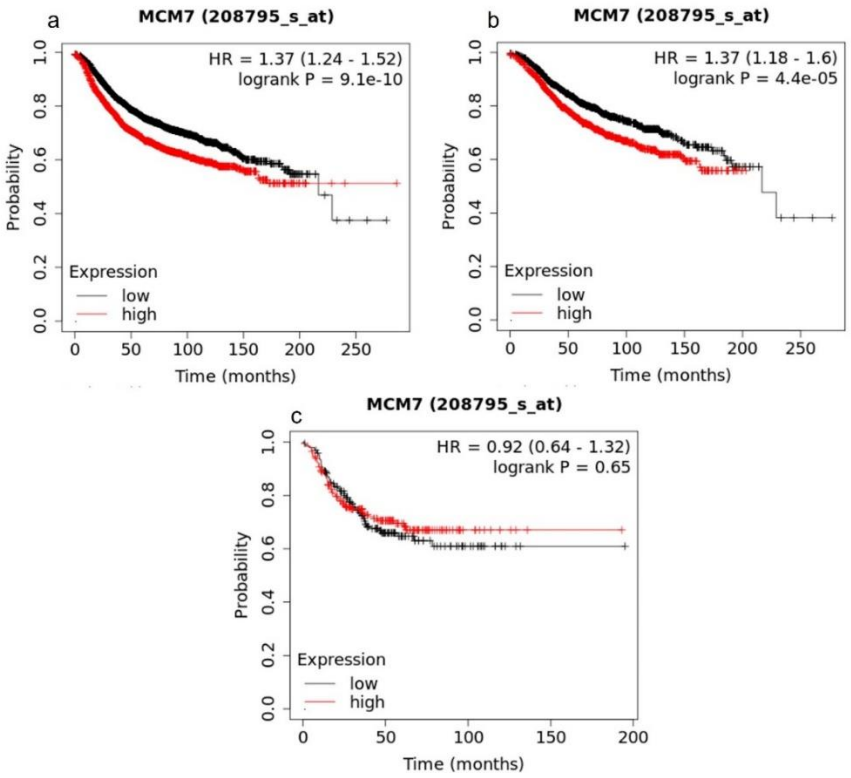


Figure 7-18: Kaplan Meier curves show association of *MCM7* mRNA levels and molecular subtypes
A significant association was found between high *MCM7* mRNA expression and poor survival in a) whole cohort and b) luminal BC, while no association was observed in basal-like BC (c) using the Kaplan Meier Plotter dataset.

7.3.3.2 MCM7 protein expression

Specific bands for MCM7 were observed in western blot at the predicted molecular weight, 80 KDa confirming the specificity of the antibody (

Figure 7-19). MCM7 showed both nuclear and cytoplasmic expression. MCM7 nuclear expression was detected in 43% (500/1156) of patients while cytoplasmic expression was found in 17% of patients (**Figure 7-20**). High nuclear MCM7 expression was significantly associated with younger patient (<50 years) (p=0.02), premenopausal status (p=0.03), high tumour grade (p<0.0001), no special type (NST) tumours (p=0.01), (p=0.01), TN phenotype and poor NPI (p=0.002) (**Table 57**). However, cytoplasmic expression did not show association with any of clinicopathological parameters or patient outcome. Therefore, all the analysis was confined to nuclear expression of MCM7.

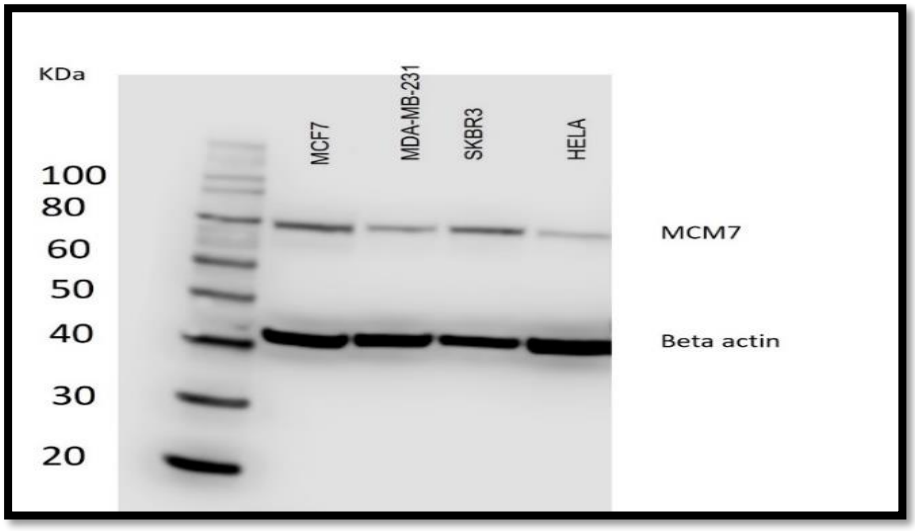


Figure 7-19: Western blot of MCM7 antibody shows specific bands at 80 kDa while B-actin shows specific bands at 42 kDa.

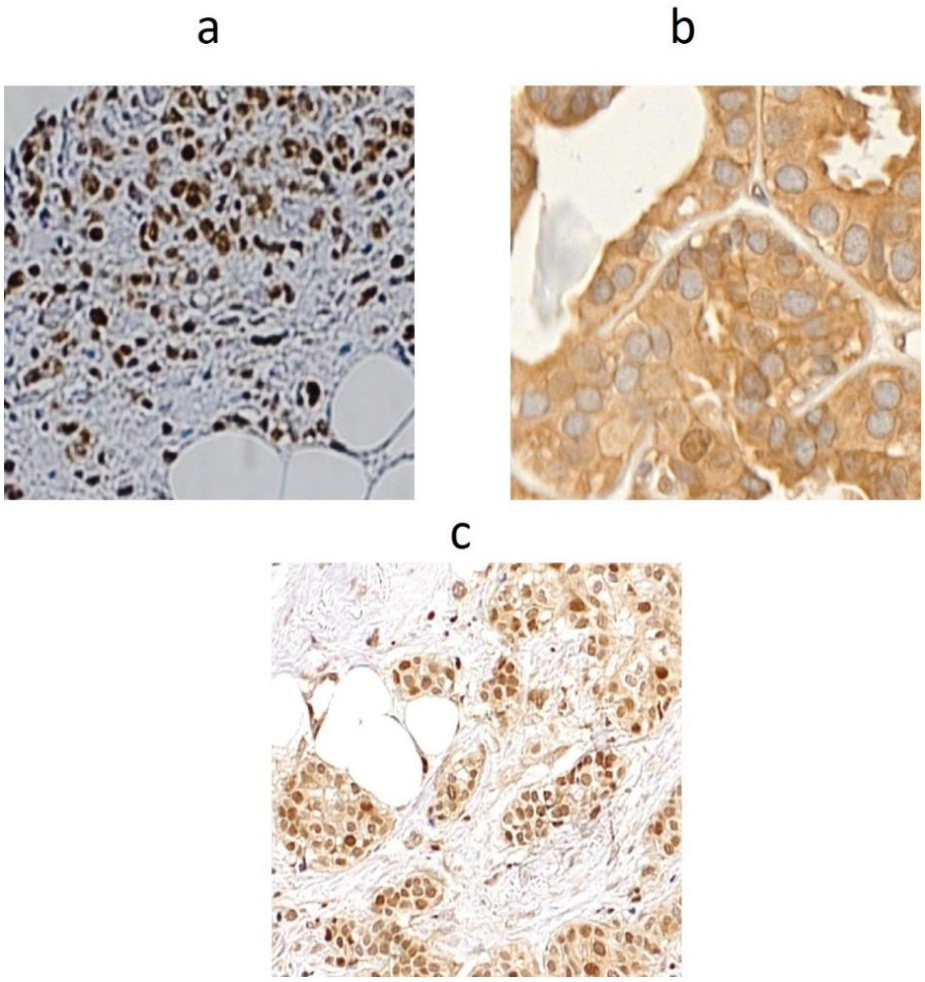


Figure 7-20: BC cases exhibiting nuclear (a), cytoplasmic (b), and both nuclear and cytoplasmic (c) MCM7 expression (x20).

Table 57: Relationship between nuclear MCM7 expression and clinicopathological parameters

Variables	MCM7 expression		X^2
	Low ≤5 %	High >5%	***P*-value**
Age at diagnosis (years)			
<50	188 (52)	174 (48)	4.9
≥50	468 (59)	326 (41)	**0.02**
Menopausal state			
Premenopausal	212 (53)	191 (47)	4.3
Post-menopausal	44 (59)	309 (41)	**0.03**
Tumour size (cm)			
≤2 cm	411 (58)	294 (42)	1.7
>2 cm	245 (54)	206 (46)	0.1
Histologic tumour grade			
Grade 1	110 (66)	58 (34)	19.1
Grade 2	285 (61)	180 (39)	**<0.0001**
Grade 3	261 (50)	262 (50)	
Histologic tumour types			
No special type (NST)	428 (56)	333 (44)	
Lobular	46 (45)	57 (55)	10.9
Other special types	28 (56)	22 (44)	**0.01**
NST mixed	154 (64)	88 (36)	
Molecular subtypes			
Luminal	543 (59)	375 (41)	9.2
HER2 enriched.	27 (49)	28 (51)	**0.01**
Triple negative	82 (48)	90 (52)	
Lymph node status			
Negative	422 (58)	309 (42)	0.7
Positive	234 (55)	191 (45)	0.37
Lymhovascular invasion			
Absent	479 (57)	360 (43)	0.1
Present	177 (56)	140 (44)	0.7
Nottingham prognostic index			12.78
Good prognostic group	248 (64)	139 (36)	**0.002**
Moderate prognostic group	311 (53)	277 (47)	
Poor prognostic group	97 (54)	84 (46)	

7.3.3.3 MCM7 protein expression and patient outcome

In the whole cohort, high MCM7 expression showed a significant association with short BCSS (p=0.004), but no association was found in terms of DMFS (p=0.1) (**Figure 7-22**). When the whole cohort was classified according to the molecular classes, there was a significant association between high MCM7 expression and shorter survival in luminal BC (p<0.0001), but no such association was observed in patients with TN or HER2+ molecular subtypes, (p=0.12 and p=0.61, respectively) (**Figure 7-23).**

In the luminal BC samples, there was an association between high MCM7 expression and poor outcome in endocrine therapy treated patients (p=0.002) (**Figure 7-24).** In contrast, there was no association between MCM7 expression and outcome in the subgroup of luminal low risk BC patients who did not receive endocrine therapy (p=0.12). Therefore, further analysis for luminal BC patients who received endocrine therapy was carried out. This analysis showed that high MCM7 was significantly associated with poor outcome in more aggressive tumours, including tumours with (a) LN metastasis (p=0.004), (b) higher-grade (grade ≥2; p<0.0001) and (c) high NPI (NPI>3.4; p=0.003). However, there was no association between MCM7 expression and patient outcome in the more indolent tumours such as tumours without LN metastasis (p=0.08), with low grade tumours (grade 1; p=0.29) and in those with a low NPI (p=0.39) (**Figure 7-21**). In the subgroup of high-risk luminal BC patients who received both endocrine therapy and chemotherapy, the association

between MCM7 expression and poor outcome was lost (p=0.07) (**Figure 7-26).**

In the whole cohort, multivariate analysis showed that high MCM7 was independently associated with shorter survival in presence of other prognostic variables including tumour size, LN status and LVI (**Table 58**). In addition, high MCM7 showed an independent association with poor survival in luminal BC patients who were treated with endocrine therapy (**Table 59**).

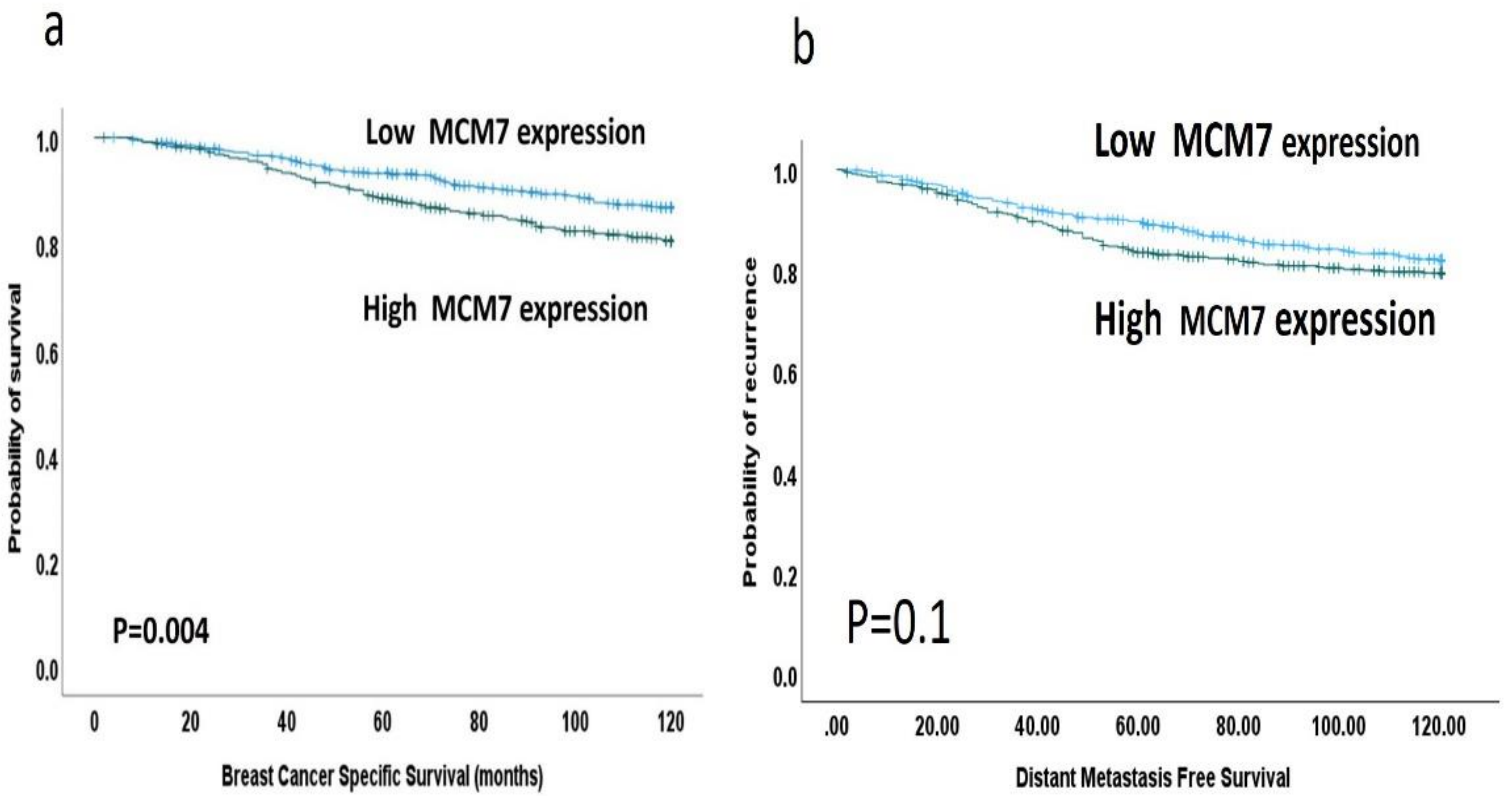


Figure 7-22: Kaplan Meier curves showing a significant association between high MCM7 expression and shorter BCSS (a) in the whole cohort while no association was observed with DMFS (b).

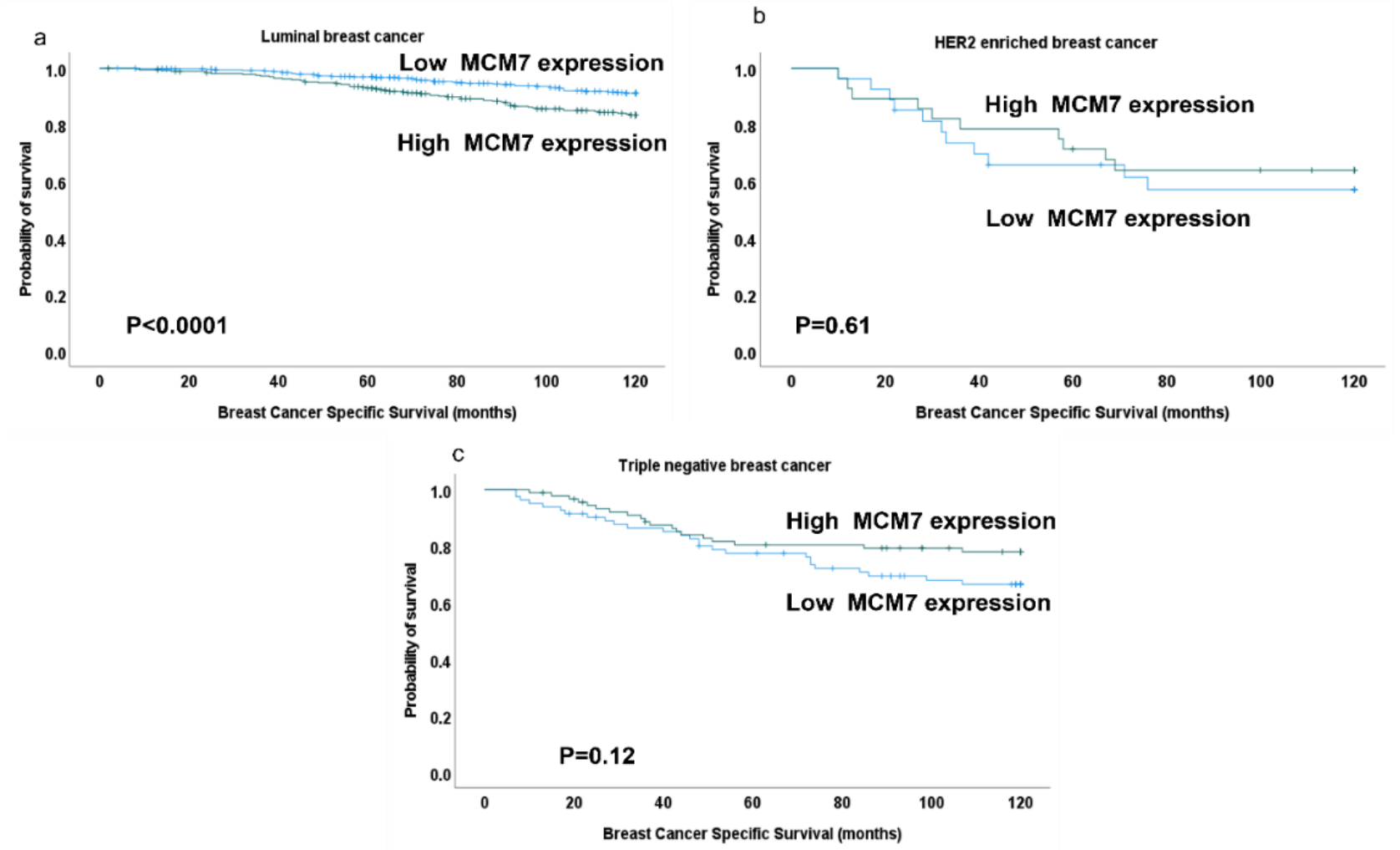


Figure 7-23: Kaplan Meir curves showing the association of high MCM7 with shorter BCSS in luminal BC (a) while no association was observed in both HER2 enriched and TNBC (b&c).

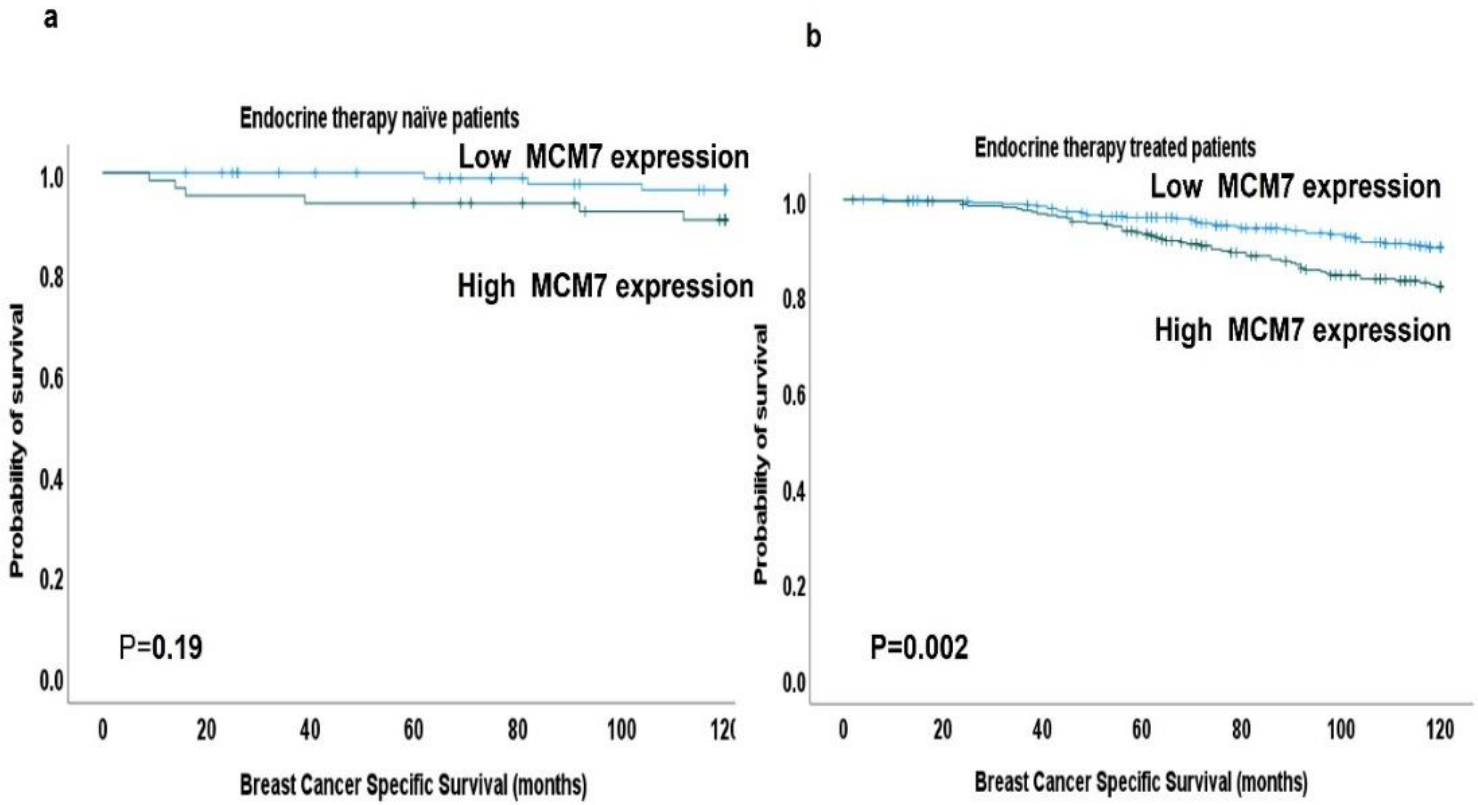
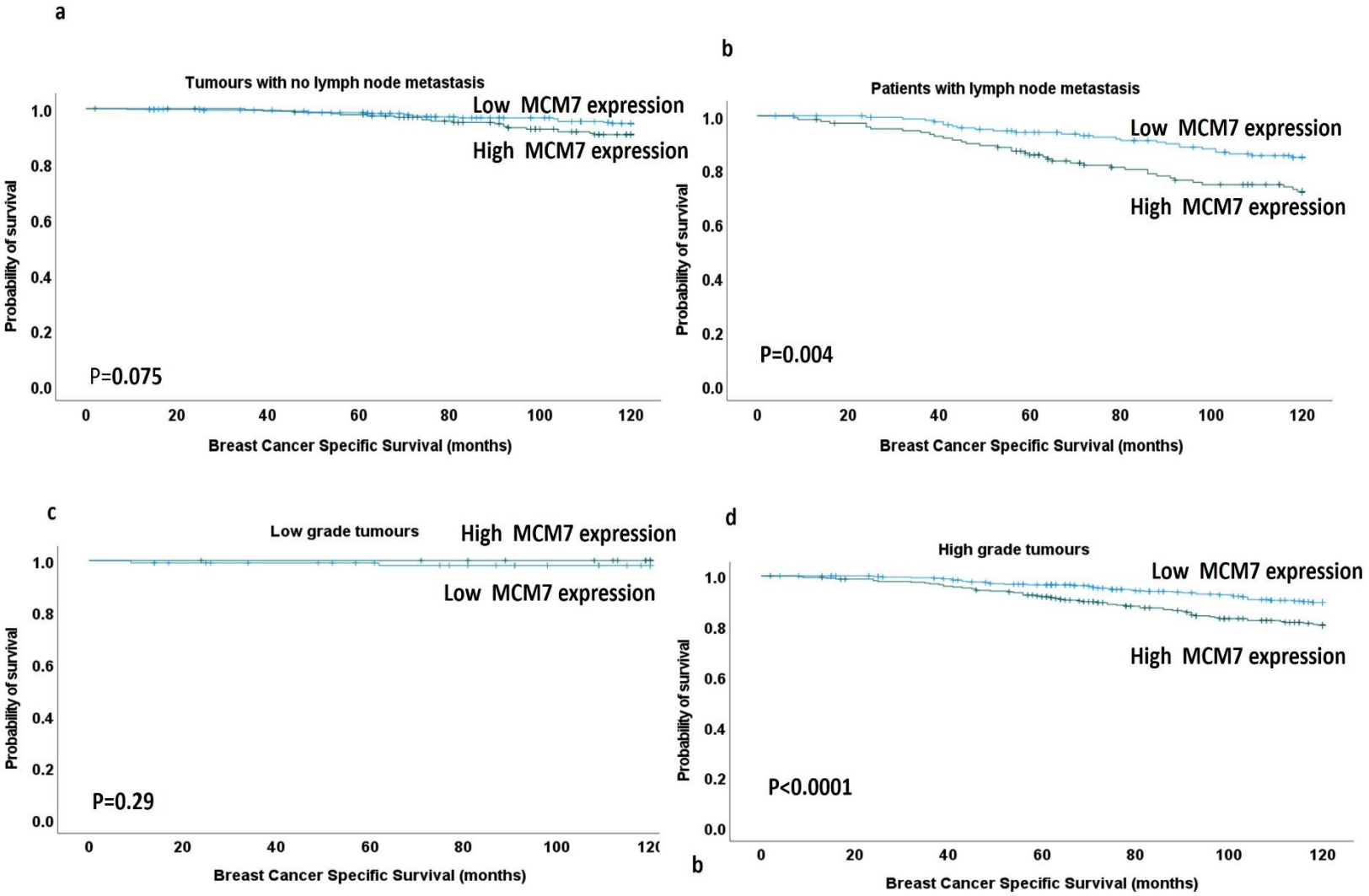


Figure 7-24: Kaplan Meir curves showing the association between high MCM7 with short BCSS in endocrine therapy naïve patients (a) where no association was observed in those who received endocrine therapy (b).


a
Tumours with no lymph node metastasis
Low MCM7 expression
High MCM7 expression
Probability of survival
P=0.075
Breast Cancer Specific Survival (months)
b
Patients with lymph node metastasis
Low MCM7 expression
High MCM7 expression
Probability of survival
P=0.004
Breast Cancer Specific Survival (months)
c
Low grade tumours
High MCM7 expression
Low MCM7 expression
Probability of survival
P=0.29
Breast Cancer Specific Survival (months)
d
High grade tumours
Low MCM7 expression
High MCM7 expression
Probability of survival
P<0.0001
Breast Cancer Specific Survival (months)

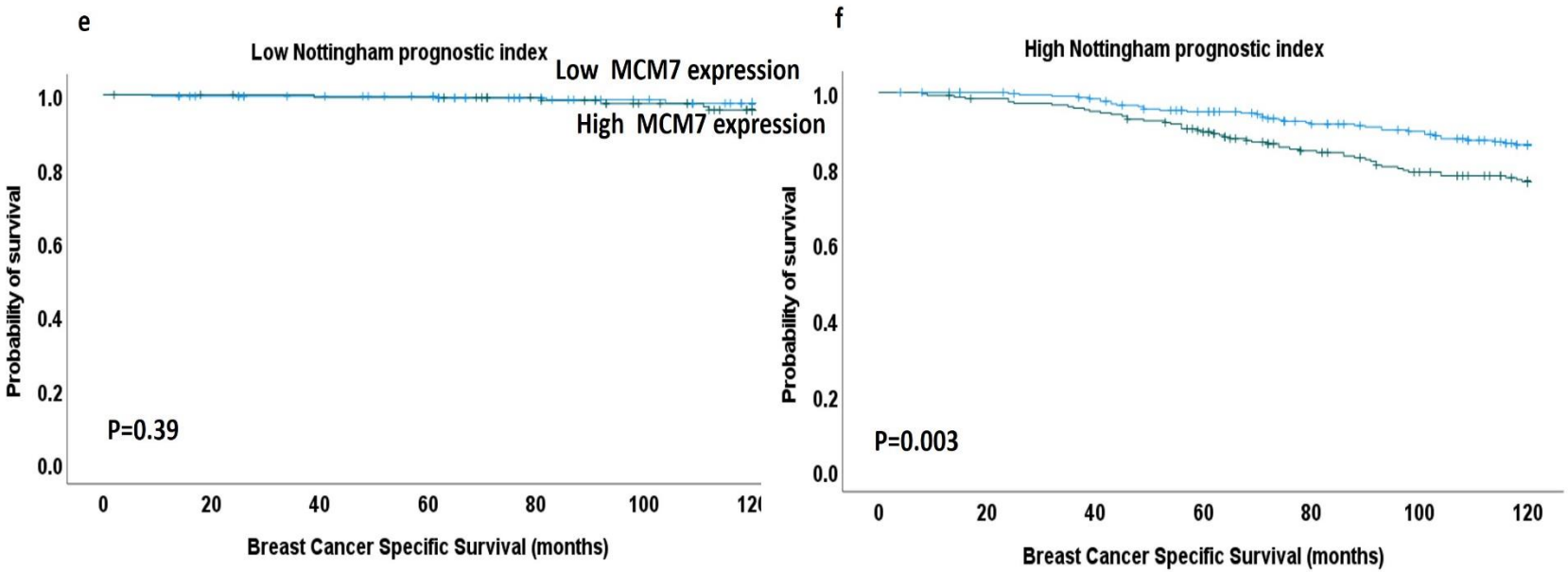


Figure 7-25: Kaplan Meier curves show a significant association between MCM7 and tumour behaviour

High MCM7 and shorter survival in tumours with LN metastasis, higher grade tumours and tumours with high NPI (b,d, f) while no association between high MCM7 and patient survival in the more indolent tumours (a,c,d).

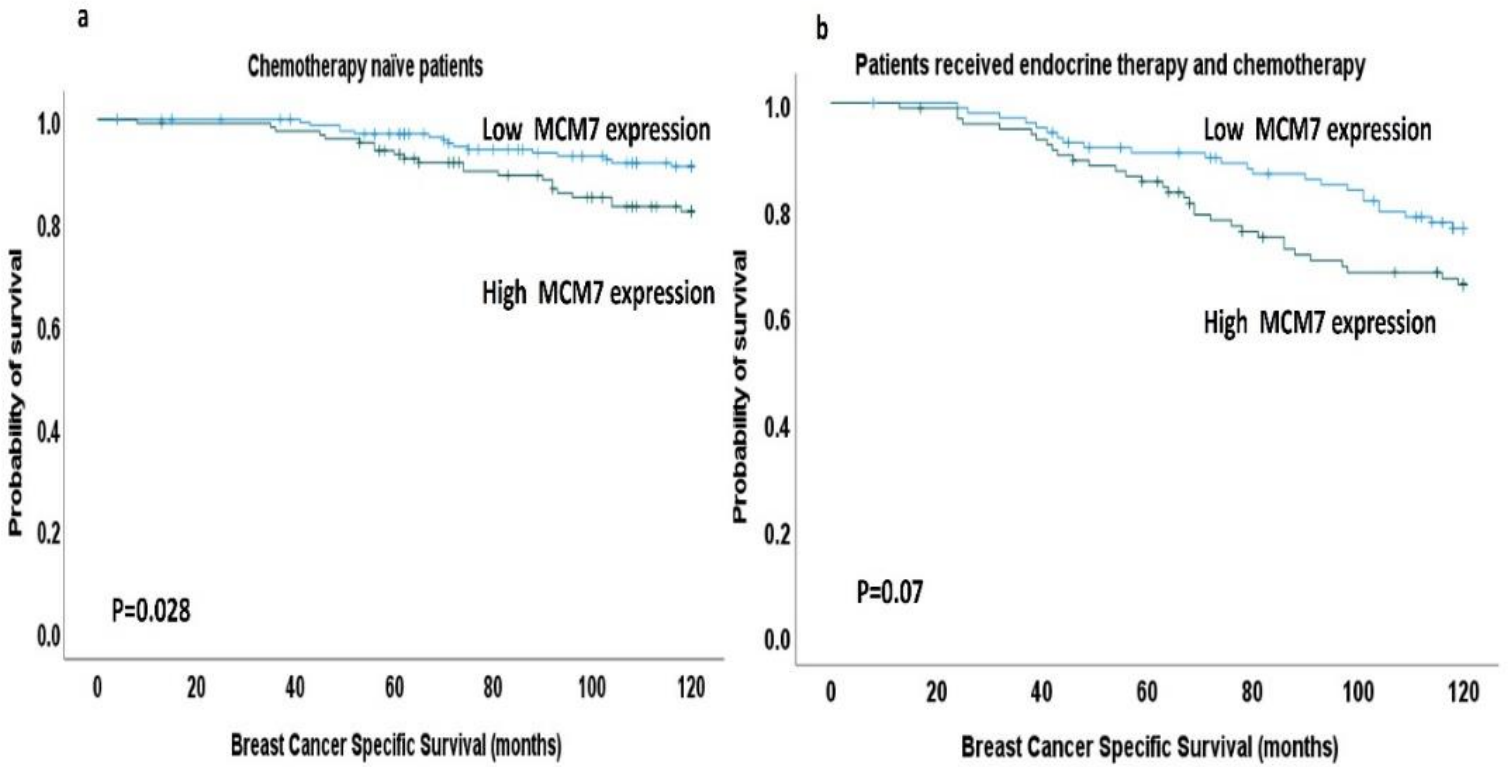


Figure 7-26: Kaplan Meier curves showing a significant association between MCM7 expression and adjuvant therapy high MCM7 expression and shorter survival in chemotherapy naïve luminal BC patients (a) while no association was observed in both endocrine therapy and chemotherapy treated patients (b).

Table 58: Multivariate analysis of MCM7 expression in BC

Parameters	BCSS		
	Hazard ratio	95% (CI)	*P-value*
MCM7	1.4	1.1-2.0	***0.04***
Tumour size	1.5	1.1-2.1	***0.009***
Tumour grade	2.7	2.1-3.7	***<0.0001***
Lymph node status	2.3	1.6-2.4	***<0.0001***

95%CI, 95% confidence interval. Significant p values are shown in **bold.**

Table 59 Multivariate analysis of MCM7 expression in luminal BC patients who were treated with endocrine therapy.

Parameters	BCSS		
	Hazard ratio	95% (CI)	*P-value*
MCM7	1.7	1.1-2.5	***0.013***
Tumour size	1.3	0.85-2.0	*0.227*
Tumour grade	2.8	1.8-4.2	***<0.0001***
Lymph node status	2.5	1.6-3.9	***<0.0001***

95%CI, 95% confidence interval. Significant p values are shown in **bold.**

7.3.4 The correlation between H.H1, PLK1 and MCM7

There was a week positive correlation between H.H1 and PLK1 (r=0.2, p<0.0001), but no association was observed between PLK1 and MCM7. In addition, no association was detected between H.H1 and MCM7.

7.4 DISCUSSION

The use of traditional biomarkers for prognosis and prediction in BC has led to improvements in overall survival [187, 559, 560]. However, to deliver additional benefits of precision medicine, there is a continuous need for identification of novel biomarkers that can predict disease prognosis and treatment response especially in the era of targeted therapies [561].

H.H1 contributes to BC aggressiveness and poor prognosis by taking part in regulating both DNA methylation and histone H3 methylation [562-565]. H.H1 binds to nucleosomal and linker DNA to assist in the stabilisation of higher-order chromatin structures [563]. It is also involved in DNA repair, cell proliferation, and in the regulation of a variety of cellular processes via interactions with cytosolic and nuclear proteins [566, 567]. Generally, H.H1 plays a role in protein-DNA interaction to regulate transcription and promote chromatin condensation [568]. Although, increased H.H1 protein expression is associated with aggressive behaviour and metastatic potential in prostate and bladder cancers [569, 570]. However, these features have not been investigated in BC.

In this study, the role of H.H1 in BC was explored based on both transcriptomic and proteomic data. At the mRNA level, high *H.H1* mRNA levels were associated with poor survival of BC patients, especially in TNBC. Similarly, XIE et al showed that *H.H1* gene expression is overexpressed in BC compared to normal breast. In addition, *H.H1*

overexpression had a significant association with short OS for BC patients [571] consistent with the current study.

At the proteomic level, H.H1 high expression was associated with parameters characteristic with aggressive tumour behaviour and poor outcome in BC patients. Similar findings were reported in patients with pancreatic adenocarcinoma and high H.H1 protein expression was also positively correlated with the Gleason grade and Ki67 positivity in prostate cancer [323]. These results confirm the previous theory that H.H1 is associated with tumour aggressiveness and poor prognosis [562] whereas H.H1 overexpression may contribute to avoidance of G1 arrest or cellular senescence in late-stage cancer [323]. In addition, H.H1 promotes the activation of tumour genes and mediates downstream signalling pathways in BC [572]. H.H1 overexpression promotes colony stimulating factor 2 (CSF2), which is a pivotal orchestrator of BC growth through its role in suppressing the immune response and regulating stem cell phenotype [573]. In view of the biological properties of H.H1, it was believed that its accumulation is related to the transition of the cells to the state of terminal differentiation rather than to proliferation itself [563]. However, depletion of H.H1 in fibroblasts results in alteration of the cell cycle, leading to decreased cell growth [574]. Additionally, depletion of H.H1 causes G1 arrest and slows proliferation in *invitro* studies [314].

Distinct H.H1 variants may have different properties and implications for gene expression [563]. For instance, H.H0 is expressed mainly in differentiated and nondividing cells, while somatic H1 variants (H.H1-

H.H5) are present in dividing cells with differing abundances depending on cell type. Of special interest for therapeutic purposes is that re-expression of the H.H0 isoform during tumour growth contributes to cell differentiation and impairs the long-term proliferative potential of cells [575]. In contrast, H.H4 knockdown cells show a cell death phenotype due to necrosis and apoptosis. Additionally, depletion of H.H2 in BC cell lines caused G1 arrest and, consequently, slow proliferation [574].

This chapter investigated both mRNA and protein expression of H1.1 in BC and evaluated patient outcomes in comparison to mRNA and protein expression, and furthermore investigated outcomes in different molecular subtypes.

When the cohort was split into molecular subclasses, high H.H1 expression showed a significant association with poor outcome in TNBC. Similar associations with aggressive behaviour in basal-like BC have been reported using RNA extraction whereas knockdown of H.H1 suppresses tumourigenicity of basal-like BC cell lines [576]. Therefore, H.H1 could function as a potential epigenetic target for those patients [577].

One of the promising prognostic surrogates in BC is PLK1, which is a key regulator of mitotic cell division and DNA repair [578]. Overexpression of *PLK1* is associated with poor prognosis in BC, and *PLK1* is considered an oncogene and a potential cancer target [579]. However, *in vivo* studies using a mouse mammary gland tumour model, *PLK1* overexpression prevented tumour development in 85% of the mouse colony and

significantly delayed tumour latency in the rest of the colony [580]. In addition, the number of tumours per animal was significantly reduced after *PLK1* overexpression in these models due to defective mitosis and enhancing apoptotic cell death [581]. Therefore, despite the therapeutic benefits of inhibiting PLK1 due to its essential role in tumour cell cycles, PLK1 overexpression may have tumour-suppressive properties [580]. Notably, evidence showed that the FDA approved proteasome inhibitor bortezomib (Velcade), which is an experimental small molecule that can reduce PLK1 protein levels and has been used to treat multiple myeloma and mantle cell lymphoma [582].

PLK1 protein expression was localised in cytoplasm only and it was suggested that PLK1 has different subcellular localisations depending on the function to exert, it locates from the cytoplasm and centrosomes during interphase, when the cells enter in mitosis, PLK1 concentrates at the spindle poles [583, 584].

Strong association was observed between high PLK1 expression and TNBC consistent with previous reports [585-587]. In addition, PLK1 has been shown to promote proliferation and angiogenesis in TNBC cell lines by inducing specific signalling pathway [588, 589].

In the whole cohort, there was a significant association between low PLK1 expression and poor outcome, however, with classification into three molecular classes, low PLK1 expression had a significant association with shorter BCSS and DMFS in luminal BC while high PLK1 was associated with poor prognosis in TNBC.

A previous study in colorectal cancer showed that reduced levels of PLK1 were correlated to aggressiveness and poor survival [590]. However, it was reported that high expression of PLK1 is associated with DNA repair deficiency, and worse prognosis in luminal BC [341].

The discrepancy between the prognostic value of PLK1 in luminal and TNBC can be explained by the different actions that PLK1 plays in those molecular subtypes. For example, in luminal BC, PLK1 mediates ER-regulated gene transcription participating in the expression of genes involved in tumour-suppressive functions and high expression of these genes correlates positively with a clinical benefit in BC patients [591, 592]. Here, low PLK1 expression was associated with poor response to endocrine therapy indicating the need for addition of chemotherapy for luminal BC patients with low PLK1 expression.

However, in TNBC, PLK1 acts as an oncogene, with its role of regulation of cell cycle and mitosis. Inhibition of PLK1 decreases the viability of various TNBC cell lines and this suggests that PLK1 is essential for survival of TNBC cells [585]. Therefore, it is implied that inhibition of PLK1 may be especially effective for the aggressive type of TNBC [588, 593, 594]. In chemotherapy treated patients, high PLK1 expression was associated with poor prognosis which confirm the aggressiveness of TNBC with high PLK1 expression. Similarly, it was reported that the resistance to some chemotherapeutic drugs has been linked to PLK1 overexpression [354]. In addition, PLK1-mediated mitotic events like, microtubule rearrangement was found to reduce the efficacy of

chemotherapeutic agents [329]. Notably, low PLK1 expression was associated with worse outcome in HER2 enriched BC, which can explain that most of HER2 enriched BC in this cohort are ER positive.

MCM7, which is an important subunit of the MCM complex, is involved in tumour formation and progression and plays an essential role in initiating DNA replication [347]. Overexpression of MCM7 has been found in several types of cancers [350, 595]. Knockdown of MCM7 significantly inhibited cellular proliferation in hepatocellular carcinoma [350] and its upregulation has been associated with cisplatin resistance in bladder cancer [596]. To investigate the potential oncogenic properties and prognostic value of MCM7 in BC, MCM7 expression was assessed at the transcriptomic and proteomic levels. MCM7 showed both nuclear and cytoplasmic subcellular localisation. It is hypothesised that cytoplasmic localisation of MCM7 may be linked to senescence rather than to the G2 phase of the cell cycle [597]. However, in this study, no specific role of the cytoplasmic subcellular localisation of MCM7 in BC was found.

High nuclear MCM7 expression showed a significant association with parameters characteristics of aggressive cancer behaviour including high grade tumours and poor NPI. Similarity, high MCM7 expression has been associated with high grade ovarian carcinoma [598] and oral squamous cell carcinoma [599]. The present research in BC, and previous research in other cancers, therefore, indicates the strong correlation between MCM7 and tumour grade, which represents the

degree of tumour proliferation. The function of MCM7 is related to cell proliferation and the proteins found only in dividing cells [600]. High nuclear MCM7 expression showed a significant association with short survival in BC. Similarly, high MCM7 expression has previously been associated with worse OS in several types of cancer including HCC [347], lung [601], bladder [602] and prostate [346]. In BC, *MCM7* transcripts were upregulated and significantly related to shorter survival [603], this was consistent with the current findings at the mRNA level. The poor prognostic role of MCM7 could be explained by the knockdown of MCM7 leading to significantly inhibited cellular proliferation *in vitro* and *in vivo* via suppression of cyclin D1 expression [595]. This suggests that MCM7 might be a potential therapeutic target for cancer treatment. In this chapter, high MCM7 expression was associated with endocrine resistance in luminal BC patients and interestingly, *MCM7* and its family are top candidate genes responsible for endocrine-resistant development. Recent evidence suggests that blocking the expression of these genes can lead to the inhibition of the growth of tamoxifen-resistant cancer cells [604]. It has been hypothesised that simvastatin, (which reduces the expression of MCM7) combined with tamoxifen induce DNA damage in tamoxifen-resistant BC cells and apoptosis of tamoxifen-resistant BC cells [605].

Conclusion

Based on the previous findings, H.H1 expression is associated with poor prognosis in BC. Given the poor prognostic role of H.H1 in TNBC, it may represent a potential therapeutic target for patients with this aggressive

disease. This chapter confirmed the role of PLK1 in TBNC as an oncogene, whereas treatment of TNBC patients with PLK1 inhibitors is recommended, caution should be taken when PLK1 inhibitors that could be used in the clinical trial in luminal BC patients.

The current findings indicate that MCM7 expression has prognostic value in BC and can be used to identify luminal BC patients who can benefit from adjuvant chemotherapy. Anti MCM7 drugs could be a valuable therapeutic strategy for MCM7-positive BC patients.

Chapter 8 Assessment of intra-tumour heterogeneity in BC using AI

8.1 Introduction

BC morphological features assessed by visual examination of the routinely stained sections are considered the standard method for BC grading and prognostication [606]. Although such assessment provides core pathological data that can guide therapy, it is still limited by the capacity of the human eye to capture all features and cellular details or to integrate multiple features in a meaningful index or score that can be applied in clinical practice. The emerging era of DP and AI could provide a better tool for assessment of multiple morphological features extracted from the scanned histological images in a more comprehensive manner [607]. BCs are highly heterogenous tumours, characterised by inter and intra-tumour heterogeneity (ITH) which largely affect patient prognosis and treatment options [608]. ITH represents the diversity exists in tumour cell subpopulations, including cancer and stromal cells across different regions of a tumour that differ in their genetic, phenotypic, or behavioural characteristics (spatial heterogeneity), and between a given primary tumour and its metastasis (temporal heterogeneity) [3]. ITH can be classified into histological and functional ITH. Histological ITH includes morphological subtype, cellular pleomorphism, differentiation, mitosis, stromal reaction, tissue composition, vascularity, infiltrative borders, and tissue necrosis. On the other hand, different genetic makeup of malignant cells within the same tumour constitutes the functional ITH [609, 610]. ITH is a consequence of differences in several cancer-cell-intrinsic parameters, such as genetic profile, interplay between the genome, transcriptome and proteome, migration and invasion capabilities,

proliferation, stemness and intrinsic cell plasticity [3, 611]. In addition, cell-autonomous traits and microenvironmental factors can also contribute to ITH.

Most of the previously used techniques to assess ITH depended on genetic analysis of micro-dissected multi-regions from the same tumour which is challenging to be replicated in routine practice [366]. AI algorithms enables the quantification of complex features of a tumour phenotype in the spatial context of the tissue, which is an imperative task for ITH assessment [365].

8.1.1 Hypothesis and aims.

This chapter hypothesised that AI-based tools can robustly assess the spatial ITH in BC, highlighting its prognostic implication. A large cohort was used and composed entirely of luminal BC, eliminating any morphological heterogeneity owing to different molecular classes.

The aim of this chapter is using the power of AI to assess the different morphological features that reflects ITH within primary tumour and its effect on patients' outcome and therapy effect.

8.2 Patients and Methods

8.2.1 Study cohort

A large cohort (n=2561) of ER positive and HER2 negative (luminal) BC was utilised in this study. The clinicopathological data, including tumour size, histological grade, histological tumour type, axillary LN status and NPI, were available (**Table 60**).

The median follow-up time was 138 months. Adjuvant endocrine therapy was offered to 2388 patients while 173 patients did not receive any type of adjuvant therapy [612]. In this study, 188 patients received both endocrine therapy and chemotherapy. The expression levels PR and Ki67 were available for the whole cohort using full-face sections as described in detail in chapter 2.5.4. H&E-stained slides were scanned into high-definition digital images through high-resolution scanning (0.19 um/ pixel) with a high throughput scanner (Pannoramic 250 Flash III; 3DHistech, Budapest, Hungary).

Table 60: Clinicopathological characteristics of the study cohort.

Variables	Number (%)
Age (years)	
<50	605 (24)
≥50	1923 (76)
Menopausal state	
Premenopausal	659 (26)
Post-menopausal	1902 (74)
Tumour size (cm)	
≤2	1876 (73)
>2	682 (27)
Histologic tumour grade	
Grade 1	604 (24)
Grade 2	1438 (56)
Grade 3	519 (20)
Histologic tumour types	
No Special Type (NST)	1428 (56)
Lobular	336 (13)
Other special types	168 (7)
NST mixed	629 (24)
Lymph node status	
Negative	1933 (76)
Positive	628 (24)
Lymhovascular invasion	
Absent	2164 (85)
Present	397 (15)
Nottingham prognostic index	
Good prognostic group	1428 (56)
Moderate prognostic group	1030 (40)
Poor prognostic group	103 (4)
ER expression levels.	
Low <10%	27 (2)
Intermediate (10-99%)	728 (28)
High (100%)	1788 (70)
PR status	
Negative ≤10%	498 (22)
Positive >10%	1811 (78)
Ki67 expression	
Low ≤10%	1008 (43)
High >10%	1388 (57)

8.2.2 Training of the algorithm

To define the morphologic features of tumours that represent ITH in BC, training of an algorithm was done by the author and a group of experienced pathologists. Different morphological features of BC were annotated in a large subset of cohort. The annotated tumour features included tumour grade components (tubule formation, degree of polymorphism and mitotic figures), different types of stromal features including tumour associated stroma and fibrofatty stroma and both intra-tumour and stromal TILs as described in detail in chapter 2.6.1.1.1 [613, 614].

8.2.3 Computing of ITH features using DL

A DL based pipeline was proposed to quantify different morphological features related to tumour and stroma. The DL part of this chapter was performed by an experienced computer scientist (Dr Wahab Noorul, University of Warwick). Tissue mask was generated for WSI using image thresholding and morphological operations. A DL model was used evaluate the morphological features of both tumour and stroma in invasive tumour [615].

8.2.4 Segmentation and classification of nuclei

(HoVer-Net), which is nuclei segmentation and classification model was used to segment and classify different types of nuclei [616]. A pre-trained version of the model (pretrained on BC subset of the PanNuke dataset) was fine-tuned to get CNNnuc on the target dataset [383]. During

inference on a WSI, a nuclei mask was generated by mapping the type/class of each detected nucleus in WSI (**Figure 8-1)**.

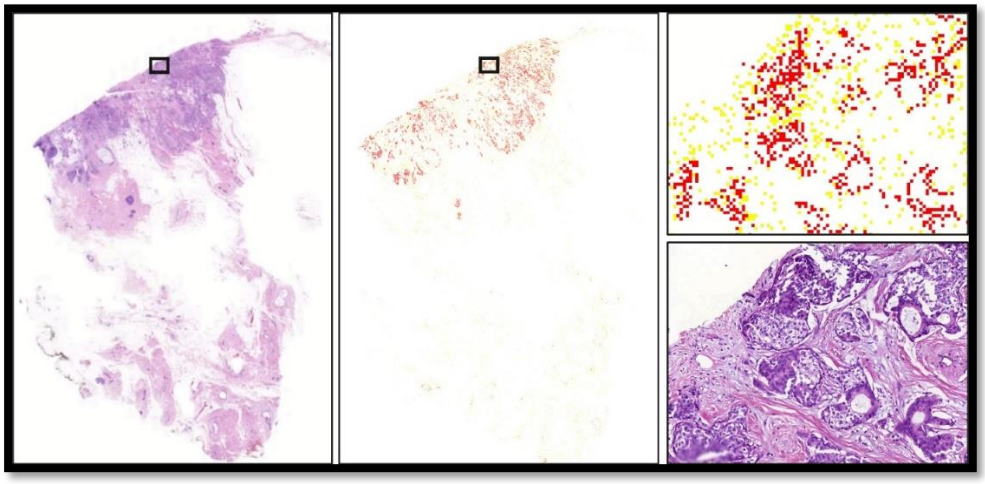


Figure 8-1: Classification for tumour nuclei for calculation of the ITH features.

8.2.5 Region segmentation

In order to segment and define regions of interest, The U-Net (a CNN) model was adopted by adding two additional encoding/decoding blocks for semantic segmentation of both tumour and stromal regions **(Figure 8-2)** and was trained on previously annotated regions [617]. Training parameters for the U-Net included: patch size of 512 × 512 with 96 pixels context on all sides, batch size 8, learning rate of 0.01 and the input was normalised to [0-1]. Different augmentations (random brightness/contrast, random rotate and median blur) were used during training with values of 0.5.

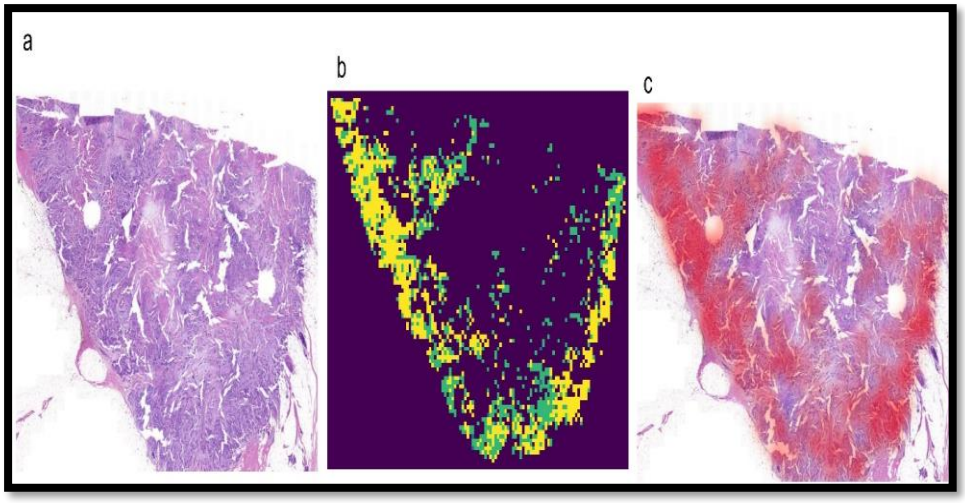


Figure 8-2: Region segmentation
a- invasive tumour regions in the H&E slide, (b&c)- maps used to quantify morphological features of ITH.

8.2.6 Evaluation of the morphological features of ITH.

Using a DL model, features related to tumour and stroma were defined and calculated for each patient to assess ITH. These features were extracted from several patches (each is 62 micrometre size) of each WSI. In addition, they were extracted using hexagonal tiling [618]. A comprehensive quantification of ITH based on these hexagonal grid analytics was performed, then systematically subsampled into a grid. This allowed computation of the total number of variables and producing a co-occurrence matrix which enables quantification of ITH features. These measures include uniformity measures and spatial measures [619], which integrate proximity between features as a key spatial component into the measurement of ITH **(Figure 8-3).** Features were calculated using a standard python library [619]. Features related to tumour differentiation heterogeneity, TILs heterogeneity and stromal heterogeneity were calculated. Colocalization represents the percentage

of each feature in an area. For each feature, several measures were calculated and recorded; colocalization value of 0 means the features do not co-localise where as a value of 1 means the same percentage of the features are found together [619]. The total number of co-occurrences were divided by 2 and these values were then normalised (i.e. each value is divided by the total sum of the values), and the resulting values were termed as cumulative matrix (CM). CM was used to calculate different features such as homogeneity and heterogeneity features such as contrast, dissimilarity and correlation as previously described [620, 621]. The sum of squared elements in the matrix, also known as uniformity which ranged between 0 and 1 was also calculated where a contrast of 0 means a constant image. Energy is quite similar to contrast but with a linear dependent off-diagonal of the CM [622].

While homogeneity measures the closeness of the element's distribution in the CM to the CM diagonal, heterogeneity was calculated as (1-homogeneity) [623, 624]. Tumours differentiation features were mainly related to tumour architecture, tubule formation, mitosis, nuclear and cytologic features. The tumour microenvironment includes stromal features which were related to only the tumour associated stroma within the context of tumour cells

Morphological tumour and stromal features were correlated with the clinicopathologic data to explore which clinical parameters showed heterogeneity and which were less heterogenous. Morphological features of tumour differentiation were normalised to (0-100) and then

summed into one score. The same was performed in terms of stromal and TILS features. To develop a heterogeneity score for each patient, all the extracted features were tested against BCSS in univariate analysis and the significant features were plotted in multivariate analysis and heterogeneity score was calculated from the beta (B) value generated from the multivariate Cox regression analysis for each of the significant variables in the univariate analysis. To classify heterogeneity score into 2 groups, X-tile software was used to determine the best cutoff. Then the heterogeneity score was correlated with the available clinicopathological and outcome data. Also, the clinical value of the heterogeneity score was evaluated in different sub-cohorts (LN+ cohort, adjuvant therapy naive cohort and endocrine therapy and chemotherapy treated cohort).

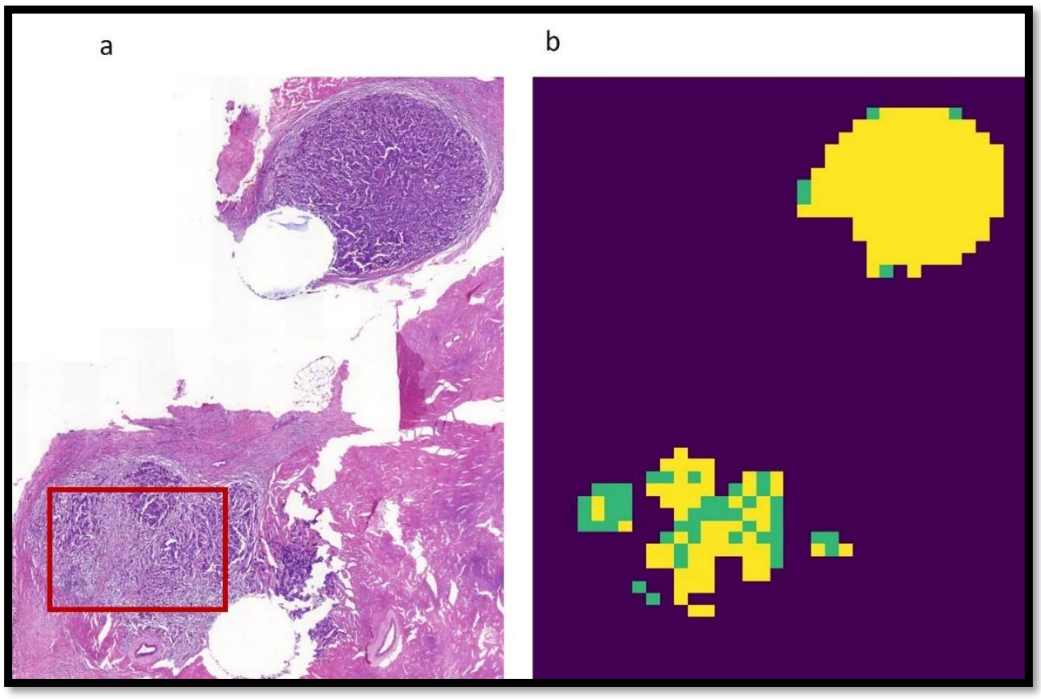


Figure 8-3: Evaluation of morphological features of ITH
a- tumour area selection in H&E image, b- heat map to calculate the heterogeneity features.

8.2.7 Statistical Analysis

Statistical Package for the Social Sciences software v.27.0 (SPSS, Chicago, IL, USA) was used for statistical analysis. Chi-square test was used for analysis of categorical data. Pearson correlation coefficient test was used to assess the correlation between continuous variables. Mean heterogeneity scores across tumour grades and other clinical parameters were compared by ANOVA test. Patients having values greater than the cutoff (>4.9) were deemed as the heterogeneous group, and otherwise they were considered as in the low heterogenous category. For statistical analysis, Ki67 expression levels were categorised into low and high using a 10% cutoff as described in chapter 5. To evaluate the association of heterogeneity score with the levels of ER expression in tumours, ER expression levels were classified into 3 groups, low ER <10%, ER intermediate (11-99%), High expression (100 %). 10% cutoff was used to categorise PR expression levels into low and high groups as previously defined in chapter 6. Considering the individual features of ITH, measures of each feature were combined into one score and this score was categorised into 2 groups based on X-tile for statistical analysis. Outcome analysis was assessed using Kaplan–Meier curves and the log-rank test. A Cox regression model was used for the multivariate analysis. Estimated HR and their 95% confidence intervals (95% CI) were calculated. For all tests, $P < 0.05$ (two-tailed) was considered statistically significant.

8.3 Results

8.3.1 Assessment of ITH features.

Features related to tumour and stroma (n=162) were extracted and quantified from WSI patches to assess ITH in BC using a DL model. Overall, 162 features related to morphological features of tumours and tumour associated microenvironment were identified and quantified. These features were assessed individually then all features were compiled in one score. A significant positive correlation was found between morphological features of tumours ($p<0.0001$). Not only did the tumour features show positive correlation but also, the positive correlation was observed was between the tumour and stromal features ($p<0.0001$).

8.3.2 Association of ITH features and clinicopathological parameters

8.3.2.1 Intra-tumour differentiation heterogeneity

The highest degree of intra-tumour differentiation heterogeneity was seen in histologic grade 3 and to a lesser degree in large size tumours and younger age patients. Low levels of ER expression and higher levels of Ki67 index were associated with increased degree of intra-tumour differentiation heterogeneity. Special tumour types showed the lowest levels of degree of intra-tumour differentiation heterogeneity compared to NST tumour type (10% versus 45% respectively). 13% of invasive lobular carcinoma showed high levels of intra-tumour differentiation heterogeneity. High degree of intra-tumour differentiation heterogeneity was associated with the development of LVI (**Table 61).**

8.3.2.2 Intra-tumour stromal and TILs heterogeneity

High degree of stromal heterogeneity was significantly associated with younger age, premenopausal status, larger tumour size, grade 3 tumours, NST tumour type, LVI, and high Ki67 expression ($p<0.0001$) (**Table 62).** Similarly, the high degree of TILs heterogeneity showed a strong association with clinical parameters characteristic of aggressive tumour behaviour (**Table 63).**

Table 61: Relationship between tumour differentiation heterogeneity clinicopathological parameters

Variables	Tumour differentiation heterogeneity		X^2 *P*-value
	Low (n, %)	High (n, %)	
Age (years)			
<50	314 (52)	291 (48)	45.4
≥50	1289 (67)	634 (33)	**<0.0001**
Menopausal state			
Premenopausal	357 (54)	302 (46)	33.5
Post-menopausal	1270 (67)	435 (33)	**<0.0001**
Tumour size			
≤2cm	1312 (70)	564 (30)	123.3
>2cm	314 (46)	386 (54)	**<0.0001**
Tumour grade			
Grade 1	543 (90)	40 (10)	618.2
Grade 2	980 (68)	458 (32)	**<0.0001**
Grade 3	104 (20)	415 (80)	
Histologic tumour types			
No special type (NST)	738 (52)	690 (48)	228.3
Lobular	286 (85)	50 (15)	**<0.0001**
Other special types	154 (92)	14 (8)	
NST mixed	449 (71)	180 (29)	
Lymph node status			
Negative	1299 (67)	638 (33)	45.9
Positive	328 (52)	300 (48)	**<0.0001**
Lymphovascular invasion			
Absent	1476 (68)	688 (32)	131.8
Present	151 (38)	246 (62)	**<0.0001**
Nottingham prognostic index			
Good prognostic group	1144 (80)	286 (20)	416.4
Moderate prognostic group	466 (45)	564 (55)	**<0.0001**
Poor prognostic group	17 (17)	86 (83)	
ER expression levels.			
Low <10%	6 (22)	21 (78)	40.9
Intermediate (10-99%)	417 (57)	311 (43)	**<0.0001**
High (100%)	1196 (67)	592 (33)	
PR status			0.62
Low ≤10%	305 (61)	193 (39)	0.43
High>10%	1144 (63)	667 (37)	
Ki67 expression			
Low ≤10%	814 (81)	194 (19)	236.8
High >10%	694 (50)	694 (50)	**<0.0001**

Significant P values are in **bold.**

Table 62: **Relationship between stromal heterogeneity and clinicopathological parameters**

Variables	Stromal heterogeneity		X^2
	Low (n, %)	High (n, %)	***P*-value**
Age (years)			
<50	290 (48)	315 (52)	42.4
≥50	1208 (63)	714 (37)	**<00001**
Menopausal state			
Premenopausal	322 (49)	337 (51)	37.6
Post-menopausal	1188 (63)	713 (37)	**<00001**
Tumour size			
≤2cm	1288 (69)	588 (31)	268.6
>2cm	222 (33)	459 (67)	**<00001**
Tumour grade			
Grade 1	439 (73)	165 (27)	131.9
Grade 2	868 (60	570 (40)	**<0.0001**
Grade 3	203 (39)	315 (61)	
Histologic tumour types			
No special type (NST)	765 (54)	663 (46)	51.6
Lobular	219 (65)	117 (35)	**<0.0001**
Other special types	131 (78)	37 (22)	
NST mixed	395 (68)	37 (22)	
Lymph node status			43.2
Negative	1210 (67)	722 (37)	
Positive	300 (48)	328 (52)	**<0.0001**
Lymphovascular invasion			
Absent	1334 (62)	829 (38)	41.7
Present	176 (44)	221 (56)	**<0.0001**
Nottingham prognostic index			
Good prognostic group	1019 (71)	409 (29)	227.1
Moderate prognostic group	469 (46)	560 (54)	**<0.0001**
Poor prognostic group	22 (21)	81 (79)	
ER expression levels.			
Low <10%	19 (70)	8 (30)	1.57
Intermediate (10-99%)	425 (58)	303 (42)	0.46
High (100%)	1057 (59)	730 (41)	
PR status			
Low ≤10%	298 (60)	200 (40)	0.47
High>10%	1052 (58)	758 (42)	0.49
Ki67 expression			
Low ≤10	724 (72)	284 (28)	122.8
High >10	683 (49)	704 (51)	**<0.0001**

Significant P values are in **bold.**

Table 63: **Relationship between TILs heterogeneity and clinicopathological parameters.**

Variables	TILS heterogeneity		X^2
	Low (n, %)	High (n, %)	***P*-value**
Age (years)			
<50	487 (80)	118 (20)	16.19
≥50	1674 (87)	248 (13)	**<0.0001**
Menopausal state			
Premenopausal	531 (81)	128 (19)	17.1
Post-menopausal	1657 (87)	244 (13)	**<0.0001**
Tumour size			
≤2cm	1752 (93)	123 (7)	355.4
>2cm	435 (64)	247 (36)	**<0.0001**
Tumour grade			
Grade 1	393 (95)	21 (5)	99.8
Grade 2	909 (90)	101 (10)	**<0.0001**
Grade 3	253 (73)	96 (27)	
Histologic tumour types			
No special type (NST)	1179 (83)	248 (17)	28.7
Lobular	289 (86)	47 (14)	**<0.0001**
Other special types	160 (95)	8 (5)	
NST mixed	560 (89)	69 (11)	
Lymph node status			
Negative	1701 (88)	231 (12)	42.03
Positive	487 (77)	141 (23)	**<0.0001**
Lymphovascular invasion			
Absent	1900 (88)	263 (12)	63.19
Present	288 (73)	109 (27)	**<0.0001**
Nottingham prognostic index			
Good prognostic group	1159 (95)	68 (5)	307.4
Moderate prognostic group	780 (76)	250 (24)	**<0.0001**
Poor prognostic group	49 (48)	54 (52)	
ER expression levels.			
Low <10%	20 (74)	7 (26)	5.8
Intermediate (10-99%)	609 (84)	118 (16)	0.05
High (100%)	1545 (86)	243 (14)	
PR status			
Low ≤10%	413 (83)	85 (17)	3.3
High>10%	1560 (86)	250 (14)	0.06
Ki67 expression			
Low ≤10	949 (94)	59 (6)	103.2
High >10	1101 (79)	286 (21)	**<0.0001**

Significant p values are in **bold.**

8.3.3 Overall heterogeneity score

The mean ± SD of heterogeneity score was 10.4 ± 25.13, median was 2.54 and ranged from -14.24 to 195.87. The mean ± SD of heterogeneity score in grade 3 tumours was 15.4 ± 15.7 compared to 2.9 ± 17.9 in grade 1 tumours. There was a significant difference in the mean of heterogeneity score between grade 1 and grade 2. Similarly, a significant difference was observed between grade 2 and grade 3. In addition, a statistically significant difference was observed between grade 1 and grade 3 in terms of heterogeneity score. In terms of tumour size, there was a significant difference in the mean heterogeneity score between larger tumour size and smaller tumour size (<2cm). In terms of histologic types, the highest heterogeneity score was observed in NST tumour type with a mean ± SD of 12.5 ± 25.2 while the lowest score was found in lobular carcinoma and other special tumour types (**Table 64)**.

With regard to ER expression levels, a high heterogeneity score was associated with low ER expression tumours with a mean ± SD of 9.4 ± 21.6 compared to 7.1 ± 22.1 in high ER expressed tumours. Similarly, low PR expression tumours had a higher heterogeneity score compared to the high PR expressed tumours. High proliferative tumours (high Ki67 expression) showed a higher heterogeneity score with mean ± SD of 11.9 ± 23.3 compared to 2.0 ± 17.7 in low proliferative tumours **(Table 64).**

A high heterogeneity score showed a significant association with younger age, premenopausal status, larger tumour size, high grade tumours, NST tumour type, presence of LVI and poor NPI. In addition,

high heterogeneity score showed a significant association with low PR expressed tumours and high proliferative tumours (p<0.001) (**Table 65).**

Table 64: **Mean and standard deviation (Mean ± SD) of heterogeneity score in the different pathological parameters**

Pathological parameters	**Mean ± SD of heterogeneity score**
Tumour size (cm)	
≤2	4.5 ± 17
>2	19.8 ± 31.1
Tumour Grade	
Grade 1	2.9 ± 17.9
Grade 2	7.4 ± 22.3
Grade 3	15. 4±15.7
Histologic tumour types	
No Special Type (NST)	12.5 ± 25.2
Lobular	4.4 ± 25.1
Other special types	0.39 ±15.8
NST mixed	11.1 ± 25.9
Lymph node status	
Negative	0.39 ± 15.8
Positive	11.1 ± 25.9)
Lymhovascular invasion	
Absent	5.3 ± 27.3
Present	7.2 ± 21.5
ER expression levels.	
Low <10%	9.4 ± 21.6
Intermediate (10-99%)	8.3± 21.5
High (100%)	7.1 ± 22.1
PR status	
Low ≤10%	11.5 ± 24.7
High>10%	6.9 ± 24.1
Ki67 expression	
Low ≤10%	2.0 ± 17.7
High >10%	11.9 ± 23.3

Table 65: **Relationship between overall heterogeneity score and clinicopathological parameters in the study cohort**

Variables	Heterogeneity score		X^2
	Low (n, %)	High (n, %)	*P*-value
Age (years)			
<50	396 (66)	208 (34)	7.7
≥50	452 (52)	411 (48)	**0.005**
Menopausal state			
Premenopausal	312 (47)	347 (53)	22.2
Post-menopausal	1102 (58)	800 (42)	**<00001**
Tumour size			
≤2cm	882 (64)	498 (36)	73.1
>2cm	157(40)	237 (60)	**<00001**
Tumour grade			
Grade 1	396 (66)	208 (34)	80.4
Grade 2	814 (57)	624 (43)	**<0.0001**
Grade 3	204 (39)	315 (61)	
Histologic tumour types			
No special type (NST)	716 (50)	712 (50)	78.7
Lobular	240 (71)	96 (29)	**<0.0001**
Other special types	126 (75)	42 (25)	
NST mixed	332 (53)	297 (47)	
Lymph node status			
Negative	11145 (59)	788 (41)	51.6
Positive	269 (43)	359 (57)	**<0.0001**
Lymphovascular invasion			
Absent	1258 (58)	906 (42)	48.1
Present	156 (39)	241 (61)	**<0.0001**
Nottingham prognostic index			154.4
Good prognostic group	932 (65)	496 (35)	**<0.0001**
Moderate prognostic group	461 (45)	569 (55)	
Poor prognostic group	21 (20)	82 (80)	
ER expression levels.			
Low <10%	10 (62)	6 (38)	1.9
Intermediate (10-99%)	263 (56)	207 (44)	0.38
High (100%)	762 (60)	518 (40)	
PR status			20.2
Low ≤10%	316 (64)	182 (36)	**<0.0001**
High>10%	944 (52)	867 (48)	
Ki67 expression			
Low ≤10%	683 (68)	325 (32)	115.7
High >10%	633 (46)	755 (54)	**<0.0001**

Significant p values are in **bold**

8.3.4 Outcome analysis

A high degree of intra-tumour differentiation heterogeneity showed significant association with shorter patient survival (p<0.0001) (**Figure 8-4).** Similarly, high degree of stromal and TILs heterogeneity had strong associations of poor patient survival (both p<0.0001) (**Figure** 8-6**)**.

There was a strong association between high overall heterogeneity score and poor patient survival in terms of both BCSS and DMFS (p<0.0001) **(Figure 8-7).** By stratifying patients according to tumour grade, heterogeneity score dichotomised patients with each of grade 2 or grade 3 tumours into 2 distinct prognostic groups (p=0.03 and 0.01, respectively). Notably, looking at patients with moderate NPI, the heterogeneity score categorised those patients into 2 distinct prognostic groups (p=0.045) (**Figure 8-8**).

In terms of therapy response, there was a significant association between a high heterogeneity score and poor outcome in patients who received endocrine therapy (p<0.0001). In contrast, there was no significant difference in outcome in patients with low versus high heterogeneity score who received both endocrine therapy and chemotherapy (p=0.867) (**Figure 8-9)**

In multivariate analysis, including overall heterogeneity score, tumour grade, LVI and tumour size, a high heterogeneity score was independently associated with poor BCSS and short DMFS (p=0.001 and p<0.0001, respectively) (**Table 66**).

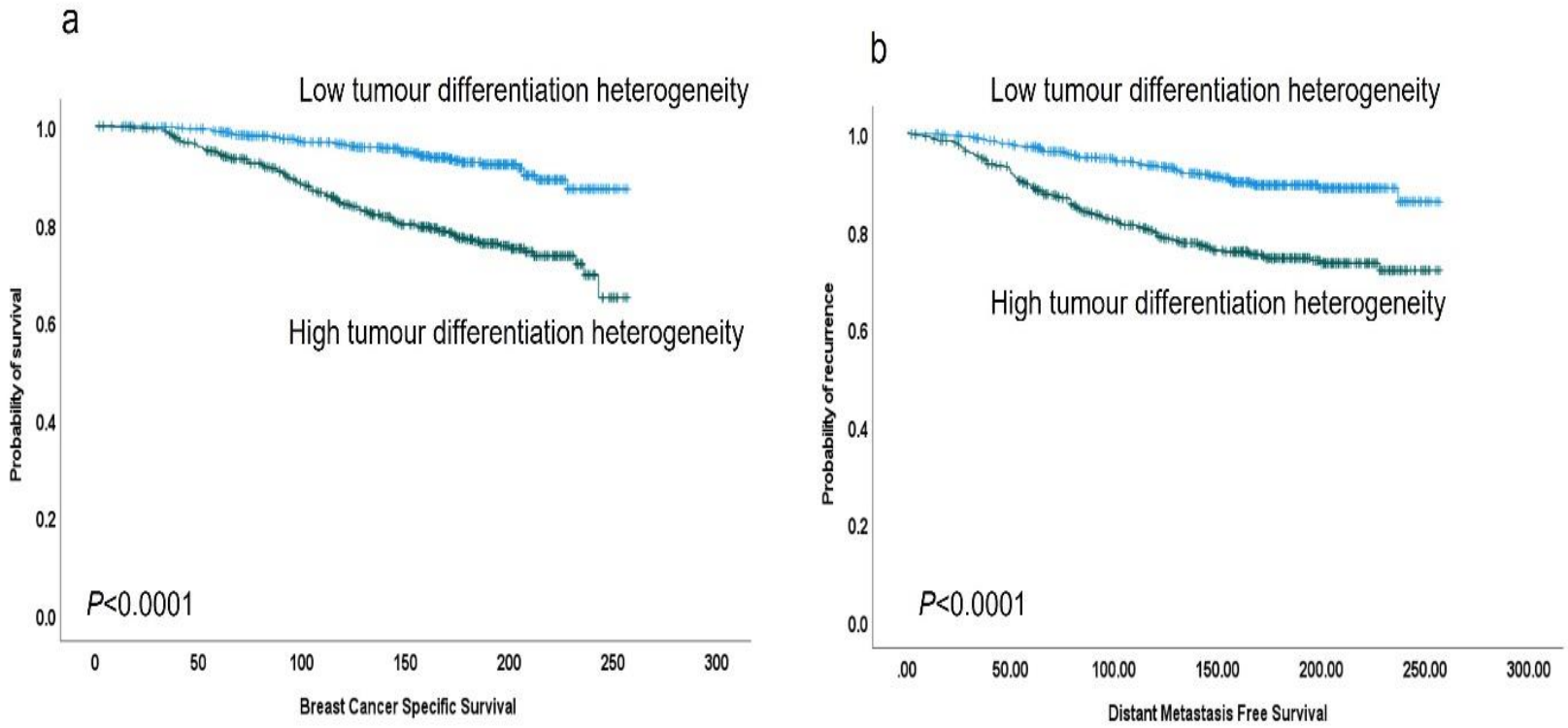


Figure 8-4: Kaplan Meier curves show the association of tumour differentiation heterogeneity with BCSS (a) and DMFS (b).

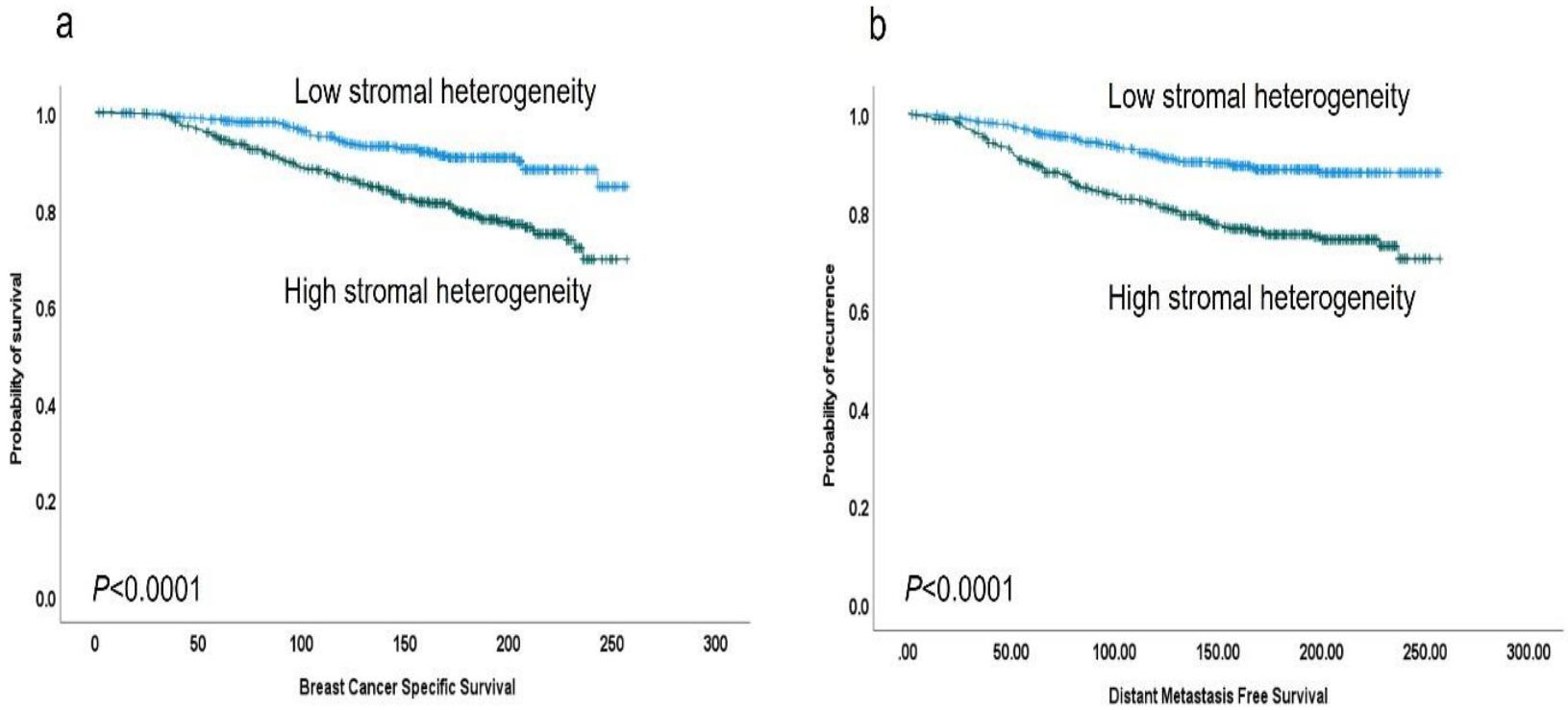


Figure 8-5: Kaplan Meier curves show the association of stromal heterogeneity with BCSS (a) and DMFS (b).

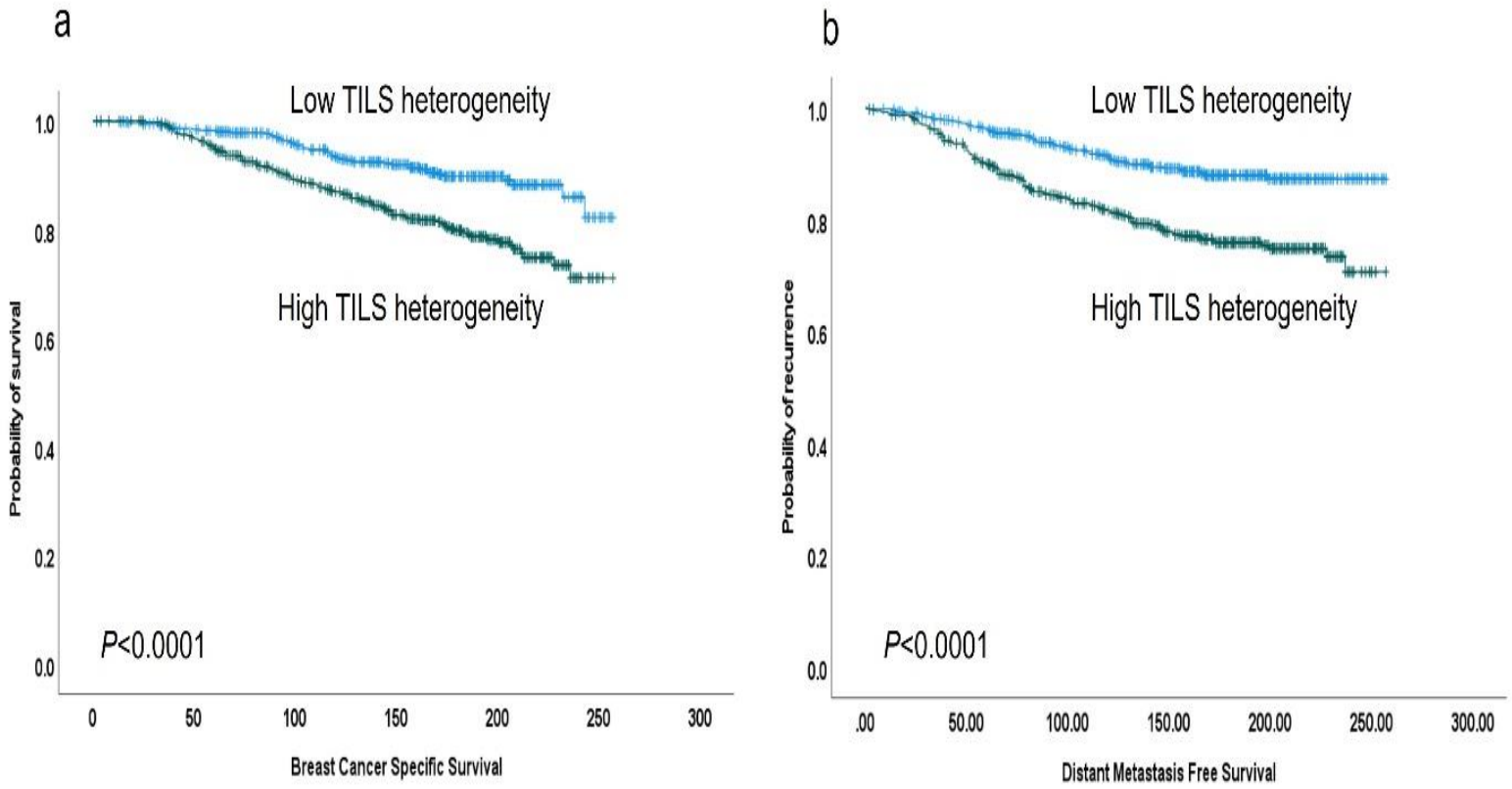


Figure 8-6: Kaplan Meier curves show the association of TILs heterogeneity with BCSS (a) and DMFS (b).

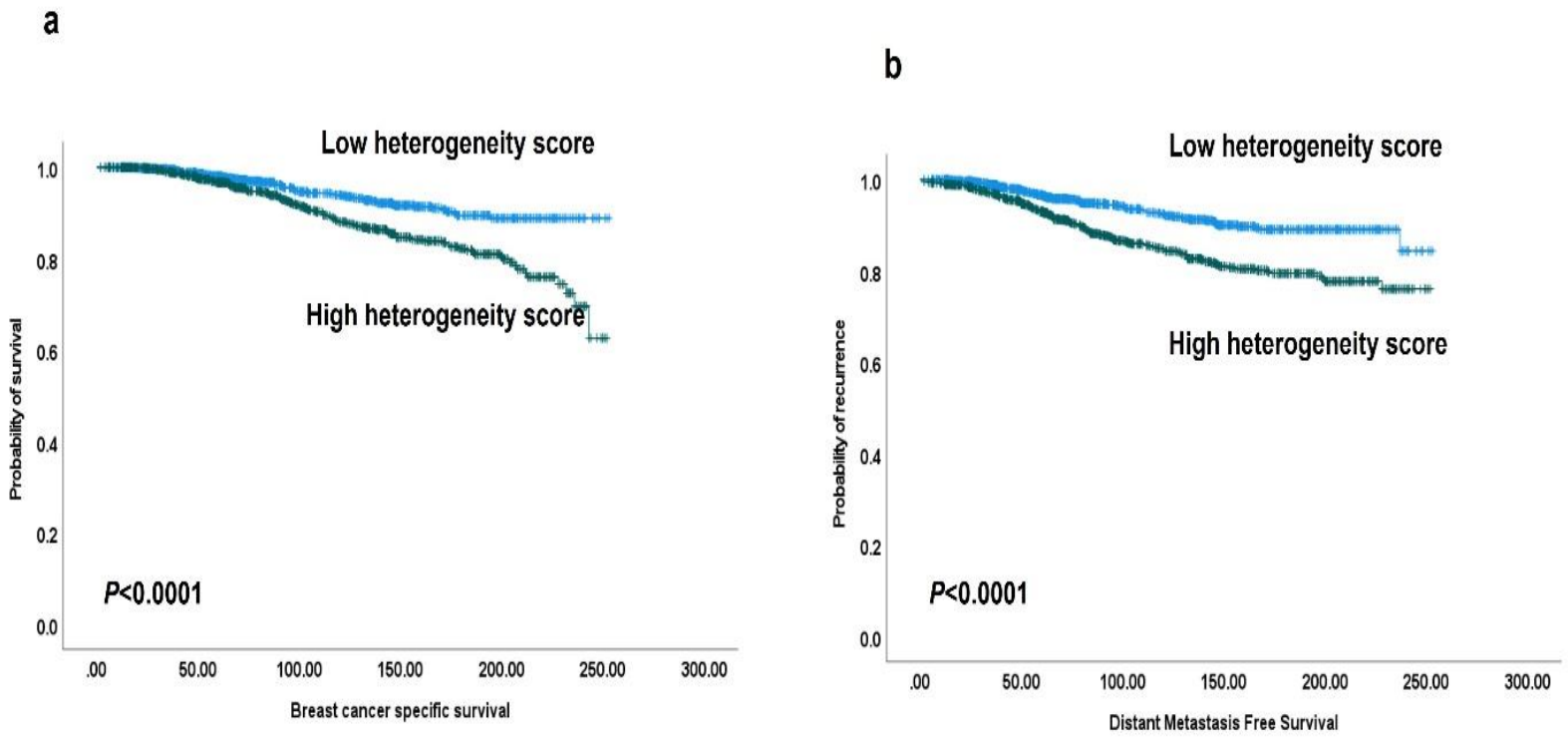


Figure 8-7: Kaplan Meier curves show the association of heterogeneity score with BCSS (a) and DMFS (b).

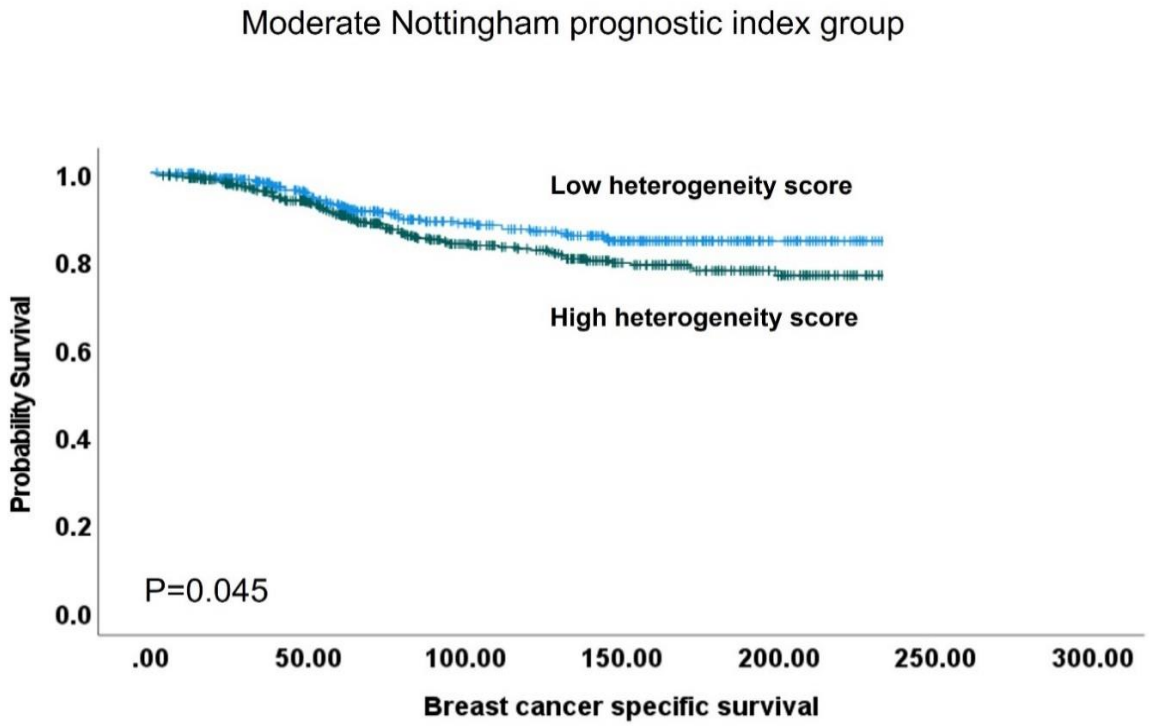


Figure 8-8: Kaplan Meier curve shows the association of heterogeneity score with BCSS in intermediate risk patients (moderate Nottingham prognostic index).

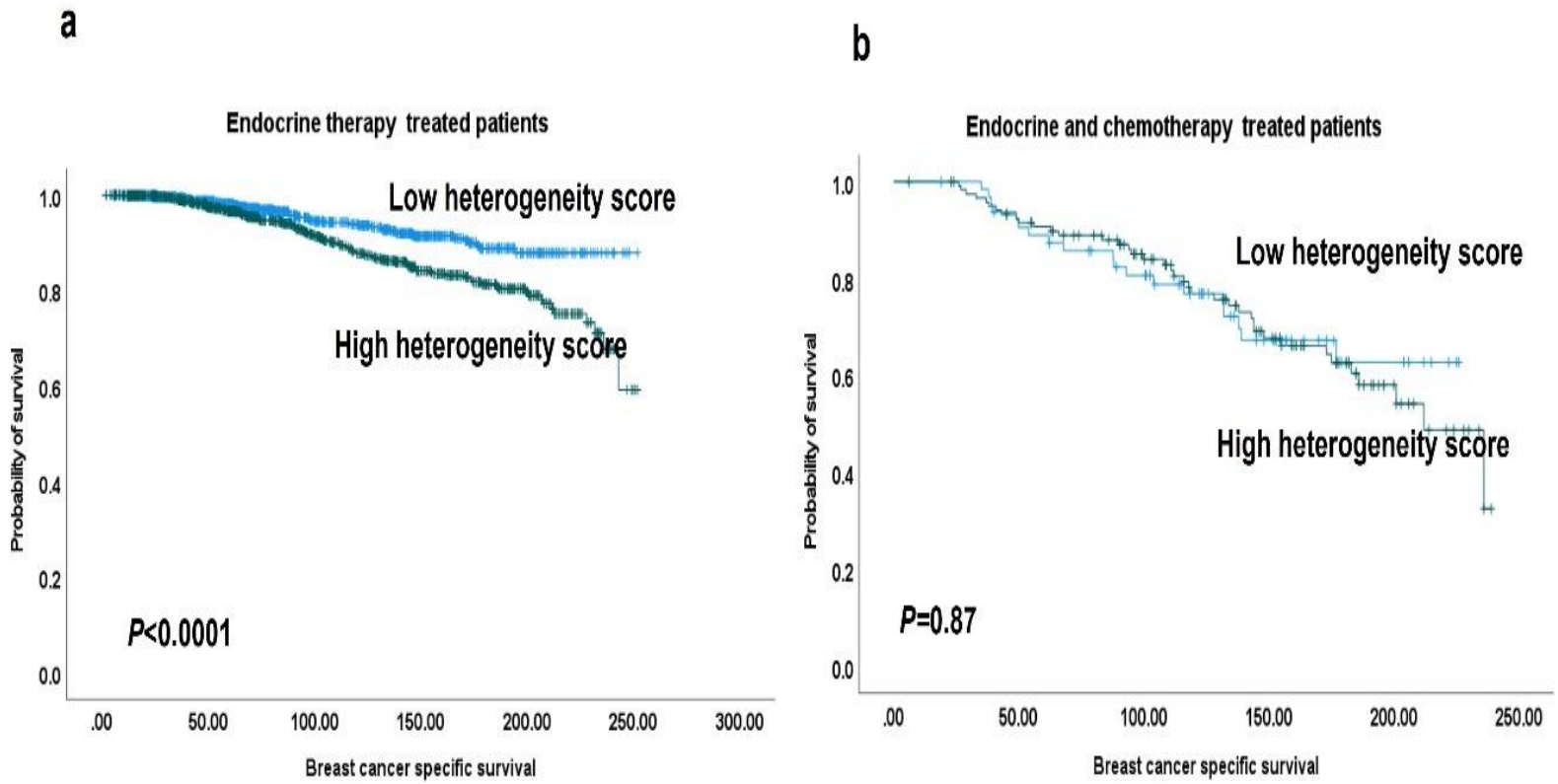


Figure 8-9: Kaplan Meier curves show the association of heterogeneity score and adjuvant therapy High heterogeneity score was associated with shorter BCSS (a) in endocrine therapy treated patients while no difference in outcome was found in both endocrine therapy and chemotherapy treated patients (b).

Table 66: Multivariate analysis of overall intra-tumour heterogeneity score in luminal BC

Parameters	**BCSS**			**DMFS**		
	Hazard ratio	95% CI	*P-value*	Hazard ratio	95% CI	*P-value*
Heterogeneity score	2.2	1.3-3.5	***0.001***	2.3	1.5-3.5	***<0.0001***
Tumour size	1.3	0.8-1.9	*0.31*	1.3	0.8-2.1	*0.17*
Grade	2.6	1.8-3.9	***<0.0001***	1.8	1.3-2.5	***0.001***
LVI	1.6	0.9-2.6	*0.062*	1.5	1.5-3.7	***<0.0001***

95%CI, 95% confidence interval. Significant p values are in **bold**.

8.3.5 Intra-tumoural heterogeneity score in sub-cohorts

8.3.5.1 LN positive sub-cohort:

The mean ± SD of heterogeneity score in LN + tumours was 18.6± 31.2, median was 9.6 (range -14.2 to 195.9), compared to LN- tumours whereas the mean ± SD of heterogeneity score was 7.7± 22.2, median was 0.8 (range -14.2 to 178). Low heterogeneity score was observed in 43% (269/628) of patients while high score was found in 57%. A high heterogeneity score had a significant association with younger age, larger tumour size, high grade, LVI and poor NPI (**Table 67).** However, no association was found between heterogeneity score and outcome in those patients (p=0.55).

Table 67: **Relationship between Heterogeneity score and clinicopathological parameters in LN positive BC**

Variables	Heterogeneity score		X^2 *P*-value
	Low (n, %)	High (n, %)	
Age (years)			
<50	65 (33)	130 (67)	11.1
≥50	204 (48)	225 (52)	**<0.0001**
Menopausal state			
Premenopausal	76 (38)	126 (62)	13.3
Post-menopausal	193 (45)	233 (55)	0.06
Tumour size			
≤2cm	190 (54)	164 (46)	38.8
>2cm	78 (29)	193 (71)	**<0.0001**
Tumour grade			
Grade 1	58 (54)	49 (46)	20.1
Grade 2	163 (46)	192 (54)	**<0.0001**
Grade 3	48 (29)	118 (71)	
Histologic tumour types			
No special type (NST)	148 (41)	217 (59)	5.7
Lobular	36 (55)	29 (45)	0.1
Other special types	10 (53)	9 (47)	
NST mixed	75 (73)	107 (27)	
Lymphovascular invasion			
Absent	196 (47)	220 (53)	9.2
Present	73 (34)	139 (66)	**0.002**
Nottingham prognostic index			32.8
Good prognostic group	51 (61)	33 (39)	**<0.0001**
Moderate prognostic group	179 (45)	244 (55)	
Poor prognostic group	21 (20)	82 (80)	
ER expression levels.			
Low <10%	5 (83)	1 (17)	9.3
Intermediate (10-99%)	76 (36)	135 (64)	**0.009**
High (100%)	184 (46)	219 (54)	
PR status			0.02
Low ≤10%	48 (42)	66 (58)	0.8
High>10%	189 (41)	269 (59)	
Ki67 expression			22.1
Low ≤10%	126 (54)	106 (46)	**<0.0001**
High >10%	126 (34)	236 (66)	

ER, oestrogen receptor. PR, Progesterone receptor. Significant p values are in **bold**

8.3.5.2 Endocrine therapy naïve sub-cohort:

The mean ± SD of heterogeneity score in endocrine therapy naïve patients was 4.8± 19.1, median was -1.1 (range -13.8 to 127.6). Low heterogeneity score was observed 64% (111/173) of patients while high score was found in 36%. In those group of patients, high heterogeneity score had a significant association with younger age, larger tumour size, NST and NST mixed tumour type, LN status and poor NPI. In those patients, no association was found between heterogeneity score and patient outcome (p=0.24).

8.4 DISCUSSION

BC is a highly heterogeneous disease caused by a variety of distinct genetic alterations in mammary epithelial cells, leading to vastly different disease manifestations in individual patients [625]. In this study, a large cohort of luminal BC, which is considered as an indeterminate challenging group of BCS was used. This uniform cohort, which is less likely to be heterogenous was selected to eliminate the impact of molecular classes as HER2 enriched and TNBC have different morphology. Currently, all luminal BC patients are treated with endocrine therapy; however, some patients eventually acquire resistance to targeted therapies and relapse [626, 627]. This resistance may be a direct consequence of pre-existing ITH [624]. Tumour heterogeneity determines cancer progression, treatment outcomes, and patient prognosis [628]. However, the precise way to assess ITH remains a challenge owing to its complexity.

With the help of AI, it can be possible to record every aspect of the spatial and cellular distribution [629]. Although several methods have been utilised for assessment of ITH in BC either through using genetic or epigenetic assays, those methods are expensive, time consuming and remain challenging for routine clinical practice [352, 626]. In addition, these methods have been studied at the molecular level not at the morphological aspect. Although morphological heterogeneity is common and being an integral part of clinical diagnostics, it has not been fully investigated. The intra-tumoral distributions of HR, HER2, and Ki-67, assessed by IHC, have been characterised by many studies [530, 630]. AI studies have demonstrated benefits in terms of capacity, precision, and accuracy for biomarker assessment. In addition, AI can address the issue with ITH because AI based algorithms generate spatial coordinates of biomarker expression within a tissue context and enable application of spatial analytics for various ITH aspects [631]. The cause of ITH has been shown to be related to clonal evolution, whereby subclones of cells with a growth advantage will persist, survive and ultimately expand, less fit subclones are unable to compete against the others [632]. Additionally, with heterogeneity in tumour cells, extensive heterotypic cooperation of cancer cells with cells occurs and these interactions jointly dictate the process of tumour progression [633].

Previous studies showed that histopathologic samples from different tumour areas in the same tumour had varying characteristics features [634]. Recent advances in computational pathology and DL have enabled discovering and quantification an enormous amount of information [365].

Another merit of AI tools is that it can measure features related to spatial distribution of tumours cells in the context of the tissue. This can aid in the identification and measurement of morphologic ITH to provide a robust prognostic and predictive factor to guide patients` treatment. The concepts of ITH and tumour-host interaction within the tumour microenvironment have been used in oncology and pathology to designate many crucial aspects of tumour biology for selecting personalised cancer therapies [635]. Heterogeneity in morphological features, especially histologic tumour gradeis one of the well-known heterogenous morphological features in routine practice. Therefore, the assessment of grade depends on the evaluation of all slides of tumours to pick the less differentiated areas and to assess mitosis figures and nuclear pleomorphism [62]. However, subjective visual quantification of the grade heterogeneity in BC is challenging and this subjectivity in assessment decrease the reliability of its use in routine practice [354]. A common issue for the methods of ITH assessment is discordance between observers which represents a challenge to the validation of ITH assessment approaches [365].

In this study, AI identified and extracted features related to tumour and its associated microenvironment from WSIs. Several morphological features showed morphological heterogeneity including tumour differentiation, tumour associated stroma and TILs. This suggests that tumours with pre-existing ITH can show multiple morphological features and could help stratification of patients according to the abundance of these features.

Grade 3 tumours and NST tumour type showed higher degree of heterogeneity, which suggest that the underlying ITH in those aggressive tumour characteristics plays a biological role in their behaviour. It was found that 13% of invasive lobular carcinoma had high levels of intra-tumour differentiation heterogeneity, which may reflect the existence of more aggressive variants with this tumour type such as pleomorphic and solid lobular carcinomas. Patients with low and intermediate ER expression had higher ITH compared to patients with high ER expression. This could explain that the good response of the tumours with high ER expression to endocrine therapy due to the less degree of heterogeneity. High heterogeneity score showed a significant association with parameters characteristic of aggressive tumour behaviour, which prove the concept of ITH role in determining the behaviour of tumours. Similarly. It was shown that high heterogeneity score defined by algorithm, using whole-exome sequencing data, had association with larger tumour size and NST tumour type [636].

A significant association was found between high heterogeneity score and poor patient outcome. This can confirm the growing evidence that BC heterogeneity can influence patients prognosis [637]. Also, It has been suggested that heterogeneity in luminal BC was associated with poor prognosis [636]. Notably, patients with higher heterogeneity score predicted poor response for endocrine therapy while patients with low heterogeneity score responded well to endocrine therapy. On the other hand, there was no significant difference in outcome between patients with high heterogeneity score versus those with low score. This confirms

that the response to chemotherapy is better in patients with high degree of heterogeneity and imply that heterogeneity score can be used as surrogate for the addition of chemotherapy. Earlier studies indicated that pre-existing intratumor heterogeneity enhances the probability of some tumour cells surviving treatment-induced elimination that could leads to therapy failure [626, 630].

Conclusion: based on the previous findings, AI-based tools can robustly assess morphological features representative of ITH in early-stage BC. Heterogeneity score can be used as a prognostic surrogate marker of luminal BC and can be used to guide therapy decisions for early-stage BC patients.

Chapter 9 General Discussion

The current study has clearly confirmed the importance of a constellation of different clinicopathological factors, proteomics and genetics in the prediction of BC behaviour and how to refine the prognosis of BC using DP.

9.1 Background

BC, the most common cancer in women, is regarded as an extraordinarily heterogeneous disease with a diverse variety of histological entities, clinical behaviour, outcome, and responses to therapy, with substantial influences on patients' prognosis and management [352]. Worldwide, BC is the second most common cancer and the fifth cause of cancer related deaths overall [638]. The majority of BC-related morbidities are attributable to the extra-mammary spread of the tumour both regionally and systemically [13]. BC patients' risk stratification and prediction of the usefulness of different therapeutic protocols currently utilise patient and tumour clinicopathological criteria [1]. It was reported that substantial proportions of patients with apparently similar clinical and pathological features may show distinctive outcomes and variable responses to the same therapeutic regimens. This can be attributed to the inter-tumour and intra-tumour heterogeneity [2, 610].

Adjuvant chemotherapy and hormonal treatment improve survival in BC patients but have potentially serious side effects [639]. Therefore, adjuvant treatment should only be given to high-risk patients, which

requires robust prognostic factors to indicate those with high-risk of recurrence and additional factors to predict response to treatment [640]. It is widely accepted that the engagement between histopathological classification, and assessment of molecular prognostic biomarkers; ER, PR, and HER2 status remains the mainstay of BC management [641]. A multitude of prognostic factors have been identified for BC; many of them are directly or indirectly related to cellular proliferation. Mitosis and Ki67 are crucial indicators of the tumour proliferation rate in BC [642]. Although mitotic count is the most important constituent of Nottingham histological grade, there are some issues with reproducibility of grading due to the lack of strict protocols and these can be amplified with the use of digital slides [643]. Similarly, immunohistochemical assessment of proliferation using the Ki67 has a potential use in both research and clinical management of BC. However, lack of consistency and low reproducibility have limited the value of Ki67 in routine clinical practice [644].

9.2 Hypothesis, aims, and methods of the study.

This thesis hypothesised that the prognosis in BC patients, particularly the development of metastatic disease, according to the widely accepted concept is an early cancer phenomenon that could be inferred from studying primary cancer. Significant improvement in patients' risk stratification and personalised treatment could be achieved by refinement of those currently used clinical and pathological parameters. In addition, it was hypothesised that the use of DP can enable reliable

and reproducible tools to improve patient outcomes. A set of aims were put forward to explore the hypothesis including (a) refining the definition of mitosis in WSI, (b) exploration of the role of atypical mitosis in BC, (c), identification of potential prognostic markers in BC, (d) studying different methods of Ki67 assessment in BC, (e) assessing the clinical value of PR in BC and (f) evaluation of ITH in BC using AI tools. To achieve these aims, a large well-characterised series of primary invasive BC was used. Digitalisation of histological glass slides from the Nottingham cohort was performed using a high-throughput automated scanner. In addition, the online available cohorts, including The TCGA cohort were also utilised in this study.

9.3 Cardinal findings of the study

9.3.1 Eyeballing assessment of mitotic figures in BC

In this study, it was hypothesised that the challenges associated with mitosis identification resulted from the difficulty in identification of the different shapes of mitotic figures and how to differentiate true mitotic figures from mimickers. These challenges increased with the era of DP. Previously, it was reported that the lack of refined criteria to assess mitosis using WSI may lead to inconsistent reporting which would have clinical implications [106, 136]. In this study, hairy like projection was not evident for all mitotic figures; however, irregular, ragged border nuclear outline was appreciated in most of the mitotic figures in WSI and it can be a valuable distinguishing feature to differentiate mitoses from their mimickers. Most of the mitotic figures show the presence of abundant

eosinophilic cytoplasm. However, some showed deep eosinophilic cytoplasm which was similar to the cytoplasm of the apoptotic bodies. Most of the mitotic figures had no specific relation to nearby neoplastic cells. Although dark hyperchromatic dense tumour cells are increasingly seen with WSI, absence of the irregular ragged border in these cells can classify them as mimickers. Some studies recommend exclusion of prophase in counting mitosis as it is easily confused with other mimickers [400, 410]. However, with the refined criteria in the current study, recognising prophase mitosis became much easier where prophase mitosis was the second most common type of mitoses in BC. Therefore, excluding prophase from counting would impact the overall mitotic account and hence the overall grading. An average of 20% reduction in mitotic count was seen using WSI compared to LM. This reduction in mitotic count led to developing a refined cutoff of mitosis for grading in WSI.

9.3.2 The role of atypical mitosis in BC

Although, mitotic count is a combination typical and atypical mitosis, the presence of atypical mitotic figures is generally acknowledged to be a feature of malignancy and provides prognostic value in many types of cancer [147]. In this study, high atypical mitoses showed a significant association with clinicopathological parameters characteristics of aggressive tumour behaviours and with poor outcome in luminal BC patients. Furthermore, a good prognostic and predictive indictor was developed in this study which is atypical-to-typical mitoses ratio. This

ratio not only showed a significant association with poor patient outcome, but also showed an association with poor prognosis in TNBC patients who received chemotherapy. Conversely, overall mitosis failed to show similar association. Differential gene expression analysis identified a large number of genes that were significantly associated with high atypical mitoses. Notably, some of these genes were involved in chromosomal localisation, segregation, and DNA breakdown. Similarly, other studies have linked atypical mitoses with genomic instability, telomere dysfunction and chromosomal abnormalities [143]. Also, presence of atypical mitosis was considered as a distinct event that is possibly involved in carcinogenesis [430].

9.3.3 Assessment of proliferation in BC using Ki67

Although, Ki67 is an essential indicator of cellular proliferation in malignancy, there is no standard method yet to evaluate its level of expression [204]. In addition, several methods have been used to evaluate Ki67 expression in BC which lower the degree of concordance between pathologists reporting Ki67 [209]. This thesis has explored the detailed methodology for Ki67, and the results showed that quantification of 1000 cells is the most accurate method to predict patients` outcome. Ki67 predicted the response to endocrine therapy and adding chemotherapy to high Ki67 cases can improve the outcome. To refine the Ki67 cutoff in luminal BC, commonly used cutoffs for Ki67 were compared where 10% was found to be the best cutoff to dichotomise luminal BC patients into two distinct prognostic groups based on patient outcome.

This was similar to previous study [221]. This study showed that automatic based Ki67 scores had positive correlation with manual scores and showed a significant association with patient outcome. In addition, the current study explored the dynamic relationship between mitotic counts and Ki67. There was a significant positive correlation between mitosis and Ki67 expression similar to previous reports [472, 473]. In addition, most of the BC cases showed overlapping of Ki67 and mitosis hotspots. This suggests that both indicators can be assessed in certain areas which reflect the most biologically active part of the tumour, which drives the outcome of the disease.

9.3.4 The Role of PR in BC

PR is an estrogen-regulated gene, and its synthesis in normal and cancer cells requires ER [287]. Although, PR negativity in ER-positive BC was considered as one of the high recurrence risks, there has been equivocal data about the role of PR expression in luminal BC [305, 512]. In this study, low PR expression was significantly associated with pathological parameters characterises of aggressive tumour behaviour including high tumour grade and poor NPI group. Low PR showed independent association with poor patient outcome consistent with previous studies [302, 305, 497, 522, 523]. In addition, low PR predicted resistance to endocrine therapy and this suggests the need for addition of adjuvant chemotherapy. With head-to-head comparison, low PR provided independent prognostic value superior to Ki67, especially in endocrine therapy treated patients. In the current study, the optimal cutoff, to

categorise PR expression levels into two distinct prognostic groups was 10% in both full-face section and core needle biopsies.

9.3.5 Identification of potential prognostic markers in BC

Biomarkers that aid in prognosis and prediction of BC are essential for timely identification and appropriate control of the disease throughout treatment [534]. Protein molecules are usually utilised as biomarkers in the diagnosis and prognosis of BC [645]. In this thesis, the prognostic and predictive values of H.H1, PLK1 and MCM7, which showed a significant association with atypical mitosis, were assessed in a large number of BC using TMA. In the current study, high expression of H.H1 was associated with parameters characteristic with aggressive tumour behaviour including high grade and LN metastasis. It was also associated with poor patient outcome. Similarly, a previous study showed that H.H1 is associated with high proliferation in prostate cancer [323]. It was shown that high H.H1 expression is associated with avoidance of G1 cellular senescence in cancer [323]. Similarly, high *H.H1* mRNA expression showed a significant association of poor prognosis of BC. Notably, high expression of H.H1 showed a significant association with poor outcome in TNBC similar to a previous study [576].

PLK1 is a key oncogenic regulator of completion of G2–M phase of the cell cycle and it is highly overexpressed in various cancer cells, including BC [586, 646]. In the current study, not only, was there a significant association between high PLK1 expression and TNBC molecular subtype, but also high PLK1 expression showed a significant

association with poor outcome in patients with this molecular subtype. However, looking at luminal BC, low PLK1 expression was associated with poor prognosis. The different functions of PLK1 in the different molecular classes of BC could be the cause of this difference. For example, PLK1 mediates transcription of genes involved in tumour-suppressive functions in luminal BC [591, 592]. On the other hand, PLK1 may act as an oncogene in TNBC where the inhibition of PLK1 *in vitro* studies, decreases the viability of TNBC cell lines [585].

Enabling replicative immortality is an important hallmark of cancer cells and requires unlimited potential in DNA replication [647]. MCM7 is an important subunit of the MCM family, and it is involved in DNA replication and tumour formation and progression [347]. In the current study, high MCM7 expression was significantly associated with poor prognosis of BC at both transcriptomic and proteomic levels. It has been shown that inhibition of MCM7 significantly inhibited cellular proliferation *in vitro* and *in vivo* studies [595]. In the current study, high MCM7 expression showed a significant association with poor outcome in luminal BC patients who received endocrine therapy.

9.3.6 Assessment of ITH in early-stage BC using AI

ITH represents the diversity exists in tumour cell subpopulations, including cancer and stromal cells across different regions of a tumour and it largely affects patient prognosis and treatment options [3]. Although several methods have been used to assess ITH, these methods can be difficult to be used in the routine clinical practice as

most of them are using genetic or epigenetic assays [648]. In this thesis, heterogeneity in morphological features of BC was assessed using DL model whereas different features related to tumour and its associated microenvironment were extracted from WSIs. It was found that grade 3 tumours and NST tumour type showed higher degree of ITH. However, only 13% of lobular carcinoma had high ITH. In addition, patients with the highest ER expression showed the lowest degree of ITH. In terms of proliferation, patients who had high Ki67 expression showed high degree of ITH. In this study, high degree of heterogeneity was significantly associated with poor prognosis in patients who received adjuvant therapy and this association was independent of other confounders.

9.4 Final conclusions

With the era of DP and AI, prognostic stratification of BC patients and prediction of survival for tailored therapy would be improved. BC progression results from the contribution of different biological pathways including hormonal receptors, growth factors and molecular signalling pathways. Increased cancer cell proliferation as evidenced by increased tumour growth fraction can directly related to BC progression and optimal assessment of Ki67 as surrogate marker of proliferation can improve the prognosis BC. In addition, AI-based tools can help to decrease the subjectivity in proliferation assessment in BC. Refining the definition of mitosis in BC helps to improve the concordance between pathologists and helps to improve grading in the clinical practice, especially with the

digitalisation of the pathology laboratories. IHC assessment of the cellular and molecular derangements is cost-effective, reproducible and clinically applicable, therefore targeting these molecular allies, through specific therapies could halt BC progression particularly those targeting specific proteins.

9.5 Strengths of the study

1- The study was designed to use a large cohort of BC patients. In this cohort, the clinicopathological parameters and clinical outcome data were available and prospectively maintained. Therefore, findings of observational studies and the studied tissue biomarkers were linked to clinical outcome and different molecular classes.

2- In the current study, the digitalisation of the glass slides for the whole cohort was used to study the biomarkers and to compare the morphological features between glass slides and digital images.

3- In this study, the antibodies were used with positive and negative controls in addition to WB for some of the studied antibodies. In addition, for adequate staining interpretation and scoring, strict adherence to published guidelines of prognostic marker studies was followed [209, 649].

4- The use of the publicly available mRNA cohorts to validate the results of proteomic results found in this study. The available digital images for the TCGA cohort were used to assess the morphological features of tumours as validation methods.

5- Different statistical approaches in determination of the optimal cutoffs of staining positivity of the used markers, rather than using arbitrarily chosen ones.
6- Using the pathologist's eye approach in markers expression scoring and using AI based tools to assess the morphological features that could be subvisual to human eye and compare between the two methods.
7- The use of data as both continuous and categorical was helpful in confirmation of the results.
8- Using bioinformatic analysis to study the association of the studied variables with biological and molecular pathways.
9- Several published studies were derived from this thesis.

9.6 Potential limitations

1- The data were only assessed in one centre and no multicentre validation.

2- Although, the processes of staining and interpretation of the expression of biomarkers are quite subjective, a standard approach was adopted in staining and independent confirmation by additional assessors were used to minimise subjectivity. Inter-observer agreement in marker interpretation was evaluated and excellent concordance was attained.

3- Some of the biomarkers used in this study were performed on TMAs. However, a subset of full-face sections was used as a

training. However, TMA results have repeatedly been reported to show excellent concordance with full face sections in BC. In addition, for Ki67 and PR which show heterogeneity of expression, full-face sections were used study the expression of those markers.

4- Evaluation of automated PR using AI based tools was not completed because of time.

5- In this study, some of the patients were diagnosed before 2000 whereas therapeutic methods used at the time tended to be different from those currently in practice and many modalities are continually evolving.

9.7 Future Plans

1- Validation studies for mitosis refined cutoff based on patient outcome is warranted.

2- Further studies for validation of the proposed methods of Ki67 would improve its use in the clinical practice.

3- Multicentre validation for the current studied markers is warranted, especially in validation of using 10% for PR expression categorisation for risk stratification of luminal BC patients.

4- Complete a panel of potential prognostic markers in BC including SMC1, NEIL3, NEIL1, Spi-B transcription factor and SKA1.

5- To increase the robustness of this study results, further functional studies are required to validate the protein expression results. For example, *in vitro* knockdown of the identified genes particularly the oncogenes would be of value to explore the practicality of therapeutic consideration of the target genes.

6- Using multiple cores per case in the TMA could be more beneficial for better assessment to overcome the core loss and the heterogeneity if any.

References

1. Rakha, E.A., G.M. Tse, and C.M. Quinn, *An update on the pathological classification of breast cancer.* Histopathology, 2023. **82**(1): p. 5-16.
2. Feng, Y., et al., *Breast cancer development and progression: Risk factors, cancer stem cells, signaling pathways, genomics, and molecular pathogenesis.* Genes Dis, 2018. **5**(2): p. 77-106.
3. Lüönd, F., S. Tiede, and G. Christofori, *Breast cancer as an example of tumour heterogeneity and tumour cell plasticity during malignant progression.* Br J Cancer, 2021. **125**(2): p. 164-175.
4. Eccles, S.A., et al., *Critical research gaps and translational priorities for the successful prevention and treatment of breast cancer.* Breast Cancer Research, 2013. **15**(5): p. R92.
5. Admoun, C. and H.N. Mayrovitz, *The Etiology of Breast Cancer*, in *Breast Cancer*, H.N. Mayrovitz, Editor. 2022, Exon Publications

Copyright: The Authors.; The author confirms that the materials included in this chapter do not violate copyright laws. Where relevant, appropriate permissions have been obtained from the original copyright holder(s), and all original sources have been appropriately acknowledged or referenced.: Brisbane (AU).

6. Lakhtakia, R. and I. Burney, *A Brief History of Breast Cancer: Part III - Tumour biology lays the foundation for medical oncology.* Sultan Qaboos Univ Med J, 2015. **15**(1): p. e34-8.
7. Arnold, M., et al., *Current and future burden of breast cancer: Global statistics for 2020 and 2040.* Breast, 2022. **66**: p. 15-23.
8. Bray, F., et al., *Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.* CA: A Cancer Journal for Clinicians, 2018. **68**(6): p. 394-424.
9. Lakhtakia, R., *A Brief History of Breast Cancer: Part I: Surgical domination reinvented.* Sultan Qaboos Univ Med J, 2014. **14**(2): p. e166-9.
10. Lima, S.M., R.D. Kehm, and M.B. Terry, *Global breast cancer incidence and mortality trends by region, age-groups, and fertility patterns.* EClinicalMedicine, 2021. **38**: p. 100985.
11. Heer, E., et al., *Global burden and trends in premenopausal and postmenopausal breast cancer: a population-based study.* Lancet Glob Health, 2020. **8**(8): p. e1027-e1037.
12. Wilkinson, L. and T. Gathani, *Understanding breast cancer as a global health concern.* Br J Radiol, 2022. **95**(1130): p. 20211033.
13. Francies, F.Z., et al., *Breast cancer in low-middle income countries: abnormality in splicing and lack of targeted treatment options.* Am J Cancer Res, 2020. **10**(5): p. 1568-1591.

14. Liu, N., et al., *Burden, trends, and risk factors for breast cancer in China from 1990 to 2019 and its predictions until 2034: an up-to-date overview and comparison with those in Japan and South Korea.* BMC Cancer, 2022. **22**(1): p. 826.
15. Tsuzuki, S., et al., *Skeletal complications in cancer patients with bone metastases.* Int J Urol, 2016. **23**(10): p. 825-832.
16. Lelekakis, M., et al., *A novel orthotopic model of breast cancer metastasis to bone.* Clin Exp Metastasis, 1999. **17**(2): p. 163-70.
17. Tommasi, C., et al., *Long-Term Effects of Breast Cancer Therapy and Care: Calm after the Storm?* J Clin Med, 2022. **11**(23).
18. Stoicescu, E.A., et al., *Ocular adverse effects of anti-cancer chemotherapy.* J Med Life, 2023. **16**(6): p. 818-821.
19. Roberts, S.A. and W.H. Frishman, *Cardiotoxicity of breast cancer chemotherapy.* Cardiol Rev, 2023.
20. Rojas, K., et al., *Endocrine Therapy for Surgeons: Practical Pearls for Managing Menopausal, Bone Loss and Sexual Adverse Effects.* Ann Surg Oncol, 2023.
21. Kurokawa, Y.K., et al., *Modeling trastuzumab-related cardiotoxicity in vitro using human stem cell-derived cardiomyocytes.* Toxicol Lett, 2018. **285**: p. 74-80.
22. Modi, N.D., et al., *Patient-reported outcomes predict survival and adverse events following anticancer treatment initiation in advanced HER2-positive breast cancer.* ESMO Open, 2022. **7**(3): p. 100475.
23. Redig, A.J. and S.S. McAllister, *Breast cancer as a systemic disease: a view of metastasis.* J Intern Med, 2013. **274**(2): p. 113-26.
24. Riggio, A.I., K.E. Varley, and A.L. Welm, *The lingering mysteries of metastatic recurrence in breast cancer.* British Journal of Cancer, 2021. **124**(1): p. 13-26.
25. *Menarche, menopause, and breast cancer risk: individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies.* Lancet Oncol, 2012. **13**(11): p. 1141-51.
26. Duffy, S.W., et al., *Mammographic density and breast cancer risk in breast screening assessment cases and women with a family history of breast cancer.* Eur J Cancer, 2018. **88**: p. 48-56.
27. Mehrgou, A. and M. Akouchekian, *The importance of BRCA1 and BRCA2 genes mutations in breast cancer development.* Med J Islam Repub Iran, 2016. **30**: p. 369.
28. Zattarin, E., et al., *Breast cancers arising in subjects with germline BRCA1 or BRCA2 mutations: Different biological and clinical entities with potentially diverse therapeutic opportunities.* Crit Rev Oncol Hematol, 2023. **190**: p. 104109.
29. Lima, Z.S., et al., *Recent advances of therapeutic targets based on the molecular signature in breast cancer: genetic mutations and implications for current treatment paradigms.* Journal of Hematology & Oncology, 2019. **12**(1): p. 38.

30. Miller, B.J., et al., *Pooled analysis of loss of heterozygosity in breast cancer: a genome scan provides comparative evidence for multiple tumor suppressors and identifies novel candidate regions.* Am J Hum Genet, 2003. **73**(4): p. 748-67.
31. Zhang, X. and T. Sjöblom, *Targeting Loss of Heterozygosity: A Novel Paradigm for Cancer Therapy.* Pharmaceuticals (Basel), 2021. **14**(1).
32. Xu, J., L. Huang, and J. Li, *DNA aneuploidy and breast cancer: a meta-analysis of 141,163 cases.* Oncotarget, 2016. **7**(37): p. 60218-60229.
33. Rodrigues-Ferreira, S. and C. Nahmias, *From tumorigenesis to cell death: the aneuploidy paradox.* Mol Cell Oncol, 2020. **7**(2): p. 1709390.
34. Lebok, P., et al., *p16 overexpression and 9p21 deletion are linked to unfavorable tumor phenotype in breast cancer.* Oncotarget, 2016. **7**(49): p. 81322-81331.
35. Shahbandi, A., H.D. Nguyen, and J.G. Jackson, *TP53 Mutations and Outcomes in Breast Cancer: Reading beyond the Headlines.* Trends Cancer, 2020. **6**(2): p. 98-110.
36. Soliman, N.A. and S.M. Yussif, *Ki-67 as a prognostic marker according to breast cancer molecular subtype.* Cancer Biol Med, 2016. **13**(4): p. 496-504.
37. Li, J.J., J.Y. Tsang, and G.M. Tse, *Tumor Microenvironment in Breast Cancer-Updates on Therapeutic Implications and Pathologic Assessment.* Cancers (Basel), 2021. **13**(16).
38. Kuhn, E., et al. *Updates on Lymphovascular Invasion in Breast Cancer*. Biomedicines, 2023. **11**, DOI: 10.3390/biomedicines11030968.
39. Nathanson, S.D., et al., *Breast cancer metastasis through the lympho-vascular system.* Clin Exp Metastasis, 2018. **35**(5-6): p. 443-454.
40. Ji, R.C., *Lymph Nodes and Cancer Metastasis: New Perspectives on the Role of Intranodal Lymphatic Sinuses.* Int J Mol Sci, 2016. **18**(1).
41. Rahman, M. and S. Mohammed, *Breast cancer metastasis and the lymphatic system.* Oncol Lett, 2015. **10**(3): p. 1233-1239.
42. David Nathanson, S., et al., *Clinicopathological Evaluation of the Potential Anatomic Pathways of Systemic Metastasis from Primary Breast Cancer Suggests an Orderly Spread Through the Regional Lymph Nodes.* Ann Surg Oncol, 2020. **27**(12): p. 4810-4818.
43. Reticker-Flynn, N.E., et al., *Lymph node colonization induces tumor-immune tolerance to promote distant metastasis.* Cell, 2022. **185**(11): p. 1924-1942.e23.
44. Kim, J.Y., et al., *The prognostic significance of the lymph node ratio in axillary lymph node positive breast cancer.* J Breast Cancer, 2011. **14**(3): p. 204-12.

45. Ngai, V., et al., *Non-invasive predictors of axillary lymph node burden in breast cancer: a single-institution retrospective analysis.* Breast Cancer Res Treat, 2022. **195**(2): p. 161-169.
46. Dudas, J., et al., *Epithelial to Mesenchymal Transition: A Mechanism that Fuels Cancer Radio/Chemoresistance.* Cells, 2020. **9**(2).
47. Obi, E.N., et al., *Biomarker Analysis of Formalin-Fixed Paraffin-Embedded Clinical Tissues Using Proteomics.* Biomolecules, 2023. **13**(1).
48. Kokkat, T.J., et al., *Archived formalin-fixed paraffin-embedded (FFPE) blocks: A valuable underexploited resource for extraction of DNA, RNA, and protein.* Biopreserv Biobank, 2013. **11**(2): p. 101-6.
49. Zheng, Y., *Study Design Considerations for Cancer Biomarker Discoveries.* J Appl Lab Med, 2018. **3**(2): p. 282-289.
50. Ellis, I.O., et al., *An international unified approach to reporting and grading invasive breast cancer. An overview of the International Collaboration on Cancer Reporting (ICCR) initiative.* Histopathology, 2023. **82**(1): p. 189-197.
51. Sluijter, C.E., et al., *The effects of implementing synoptic pathology reporting in cancer diagnosis: a systematic review.* Virchows Arch, 2016. **468**(6): p. 639-49.
52. Aurilio, G., et al., *A meta-analysis of oestrogen receptor, progesterone receptor and human epidermal growth factor receptor 2 discordance between primary breast cancer and metastases.* European Journal of Cancer, 2014. **50**(2): p. 277-289.
53. Davey, M.G., et al., *Ki-67 as a Prognostic Biomarker in Invasive Breast Cancer.* Cancers (Basel), 2021. **13**(17).
54. Chrzanowska, N.M., J. Kowalewski, and M.A. Lewandowska, *Use of Fluorescence In Situ Hybridization (FISH) in Diagnosis and Tailored Therapies in Solid Tumors.* Molecules, 2020. **25**(8).
55. Kiyose, S., et al., *Chromogenic in situ hybridization (CISH) to detect HER2 gene amplification in breast and gastric cancer: Comparison with immunohistochemistry (IHC) and fluorescence in situ hybridization (FISH).* Pathology International, 2012. **62**(11): p. 728-734.
56. Rana, M.K., et al., *Evolution of Frozen Section in Carcinoma Breast: Systematic Review.* Int J Breast Cancer, 2022. **2022**: p. 4958580.
57. Narod, S.A. and V. Sopik, *Is invasion a necessary step for metastases in breast cancer?* Breast Cancer Res Treat, 2018. **169**(1): p. 9-23.
58. Cserni, G., *Histological type and typing of breast carcinomas and the WHO classification changes over time.* Pathologica, 2020. **112**(1): p. 25-41.
59. Tan, P.H., et al., *The 2019 World Health Organization classification of tumours of the breast.* Histopathology, 2020. **77**(2): p. 181-185.

60. Makki, J., *Diversity of Breast Carcinoma: Histological Subtypes and Clinical Relevance.* Clin Med Insights Pathol, 2015. **8**: p. 23-31.
61. Allison, K.H., *Prognostic and predictive parameters in breast pathology: a pathologist's primer.* Modern Pathology, 2021. **34**: p. 94-106.
62. van Dooijeweert, C., P.J. van Diest, and I.O. Ellis, *Grading of invasive breast carcinoma: the way forward.* Virchows Arch, 2022. **480**(1): p. 33-43.
63. Rakha, E.A., et al., *Prognostic significance of Nottingham histologic grade in invasive breast carcinoma.* J Clin Oncol, 2008. **26**(19): p. 3153-8.
64. Elston, C.W. and I.O. Ellis, *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**(5): p. 403-410.
65. Zhou, L., M. Rueda, and A. Alkhateeb, *Classification of Breast Cancer Nottingham Prognostic Index Using High-Dimensional Embedding and Residual Neural Network.* Cancers (Basel), 2022. **14**(4).
66. Takahashi, H., et al., *Molecular Biological Features of Nottingham Histological Grade 3 Breast Cancers.* Ann Surg Oncol, 2020. **27**(11): p. 4475-4485.
67. Langer, I., et al., *Axillary recurrence rate in breast cancer patients with negative sentinel lymph node (SLN) or SLN micrometastases: prospective analysis of 150 patients after SLN biopsy.* Ann Surg, 2005. **241**(1): p. 152-8.
68. Tonellotto, F., et al., *Impact of Number of Positive Lymph Nodes and Lymph Node Ratio on Survival of Women with Node-Positive Breast Cancer.* Eur J Breast Health, 2019. **15**(2): p. 76-84.
69. Green, A.R., et al., *Nottingham Prognostic Index Plus: Validation of a clinical decision making tool in breast cancer in an independent series.* J Pathol Clin Res, 2016. **2**(1): p. 32-40.
70. Lee, A.H. and I.O. Ellis, *The Nottingham prognostic index for invasive carcinoma of the breast.* Pathol Oncol Res, 2008. **14**(2): p. 113-5.
71. Kerin, E.P., et al., *Comparison of the Nottingham Prognostic Index and OncotypeDX© recurrence score in predicting outcome in estrogen receptor positive breast cancer.* The Breast, 2022. **66**: p. 227-235.
72. Zhong, Y.-M., F. Tong, and J. Shen, *Lympho-vascular invasion impacts the prognosis in breast-conserving surgery: a systematic review and meta-analysis.* BMC Cancer, 2022. **22**(1): p. 102.
73. Kuhn, E., et al., *Updates on Lymphovascular Invasion in Breast Cancer.* Biomedicines, 2023. **11**(3).
74. Kariri, Yousif A., et al., *Molecular Complexity of Lymphovascular Invasion: The Role of Cell Migration in Breast Cancer as a Prototype.* Pathobiology, 2020. **87**(4): p. 218-231.

75. Ryu, Y.J., et al., *Lymphovascular invasion can be better than pathologic complete response to predict prognosis in breast cancer treated with neoadjuvant chemotherapy.* Medicine (Baltimore), 2018. **97**(30): p. e11647.
76. Saponaro, M., et al., *Extended Adjuvant Endocrine Treatment in Luminal Breast Cancers in the Era of Genomic Tests.* Int J Mol Sci, 2022. **23**(21).
77. Rakha, E.A., et al., *Breast cancer prognostic classification in the molecular era: the role of histological grade.* Breast Cancer Research, 2010. **12**(4): p. 207.
78. Tang, J., et al., *An estrogen receptor (ER)-related signature in predicting prognosis of ER-positive breast cancer following endocrine treatment.* J Cell Mol Med, 2019. **23**(8): p. 4980-4990.
79. Sudhesh Dev, S., et al., *Receptor Tyrosine Kinases and Their Signaling Pathways as Therapeutic Targets of Curcumin in Cancer.* Front Pharmacol, 2021. **12**: p. 772510.
80. Hsu, J.L. and M.C. Hung, *The role of HER2, EGFR, and other receptor tyrosine kinases in breast cancer.* Cancer Metastasis Rev, 2016. **35**(4): p. 575-588.
81. Iqbal, N. and N. Iqbal, *Human Epidermal Growth Factor Receptor 2 (HER2) in Cancers: Overexpression and Therapeutic Implications.* Mol Biol Int, 2014. **2014**: p. 852748.
82. Gemmete, J.J. and S.K. Mukherji, *Trastuzumab (herceptin).* AJNR Am J Neuroradiol, 2011. **32**(8): p. 1373-4.
83. *Trastuzumab for early-stage, HER2-positive breast cancer: a meta-analysis of 13 864 women in seven randomised trials.* Lancet Oncol, 2021. **22**(8): p. 1139-1150.
84. Sørlie, T., et al., *Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications.* Proc Natl Acad Sci U S A, 2001. **98**(19): p. 10869-74.
85. Makretsov, N.A., et al., *Hierarchical clustering analysis of tissue microarray immunostaining data identifies prognostically significant groups of breast carcinoma.* Clin Cancer Res, 2004. **10**(18 Pt 1): p. 6143-51.
86. Desmedt, C., C. Sotiriou, and M.J. Piccart-Gebhart, *Development and validation of gene expression profile signatures in early-stage breast cancer.* Cancer Invest, 2009. **27**(1): p. 1-10.
87. Tsang, J.Y.S. and G.M. Tse, *Molecular Classification of Breast Cancer.* Adv Anat Pathol, 2020. **27**(1): p. 27-35.
88. Rakha, E.A., et al., *Triple-negative breast cancer: distinguishing between basal and nonbasal subtypes.* Clin Cancer Res, 2009. **15**(7): p. 2302-10.
89. Singh, D.D. and D.K. Yadav, *TNBC: Potential Targeting of Multiple Receptors for a Therapeutic Breakthrough, Nanomedicine, and Immunotherapy.* Biomedicines, 2021. **9**(8).
90. Mazo, C., et al., *Multi-Gene Prognostic Signatures and Prediction of Pathological Complete Response to Neoadjuvant Chemotherapy in ER-positive, HER2-negative Breast Cancer Patients.* Cancers (Basel), 2020. **12**(5).

91. Sinn, P., et al., *Multigene Assays for Classification, Prognosis, and Prediction in Breast Cancer: a Critical Review on the Background and Clinical Utility.* Geburtshilfe Frauenheilkd, 2013. **73**(9): p. 932-940.
92. Arango, B.A., C.L. Rivera, and S. Glück, *Gene expression profiling in breast cancer.* Am J Transl Res, 2013. **5**(2): p. 132-8.
93. Cognetti, F., et al., *Multigene tests for breast cancer: the physician’s perspective.* Oncotarget; Vol 12, No 9, 2021.
94. Curigliano, G., et al., *PROCURE European consensus on breast cancer multigene signatures in early breast cancer management.* npj Breast Cancer, 2023. **9**(1): p. 8.
95. Shi, X., H. Yi, and S. Ma, *Measures for the degree of overlap of gene signatures and applications to TCGA.* Brief Bioinform, 2015. **16**(5): p. 735-44.
96. Su, J., et al., *Development of prognostic signature and nomogram for patients with breast cancer.* Medicine (Baltimore), 2019. **98**(11): p. e14617.
97. Moscalu, M., et al., *Histopathological Images Analysis and Predictive Modeling Implemented in Digital Pathology-Current Affairs and Perspectives.* Diagnostics (Basel), 2023. **13**(14).
98. Pantanowitz, L., et al., *Validating Whole Slide Imaging for Diagnostic Purposes in Pathology Guideline from the College of American Pathologists Pathology and Laboratory Quality Center.* Archives of pathology & laboratory medicine, 2013. **137**.
99. Niazi, M.K.K., A.V. Parwani, and M.N. Gurcan, *Digital pathology and artificial intelligence.* The Lancet Oncology, 2019. **20**(5): p. e253-e261.
100. Melo, R.C.N., et al., *Whole Slide Imaging and Its Applications to Histopathological Studies of Liver Disorders.* Frontiers in Medicine, 2020. **6**: p. 310.
101. Mukhopadhyay, S., et al., *Whole Slide Imaging Versus Microscopy for Primary Diagnosis in Surgical Pathology: A Multicenter Blinded Randomized Noninferiority Study of 1992 Cases (Pivotal Study).* Am J Surg Pathol, 2018. **42**(1): p. 39-52.
102. Bauer, T., et al., *Validation of Whole Slide Imaging for Primary Diagnosis in Surgical Pathology.* Archives of pathology & laboratory medicine, 2013.
103. Tabata, K., et al., *Whole-slide imaging at primary pathological diagnosis: Validation of whole-slide imaging-based primary pathological diagnosis at twelve Japanese academic institutes.* Pathol Int, 2017. **67**(11): p. 547-554.
104. Indu, M., R. Rathy, and M.P. Binu, *"Slide less pathology": Fairy tale or reality?* Journal of oral and maxillofacial pathology : JOMFP, 2016. **20**(2): p. 284-288.
105. Mahmud, H., et al., *What influences algorithmic decision-making? A systematic literature review on algorithm aversion.* Technological Forecasting and Social Change, 2022. **175**: p. 121390.

106. Williams, B.J., et al., *Digital pathology for the primary diagnosis of breast histopathological specimens: an innovative validation and concordance study on digital pathology validation and training.* Histopathology, 2018. **72**(4): p. 662-671.
107. Barisoni, L., et al., *Digital pathology and computational image analysis in nephropathology.* Nat Rev Nephrol, 2020. **16**(11): p. 669-685.
108. Pantanowitz, L.X., et al., *Review of the current state of whole slide imaging in pathology.* Journal of Pathology Informatics, 2011. **2**.
109. García-Rojo, M., et al., *New European Union Regulations Related to Whole Slide Image Scanners and Image Analysis Software.* Journal of pathology informatics, 2019. **10**: p. 2-2.
110. Paul, D., et al., *Artificial intelligence in drug discovery and development.* Drug Discov Today, 2021. **26**(1): p. 80-93.
111. Abels, E., et al., *Computational pathology definitions, best practices, and recommendations for regulatory guidance: a white paper from the Digital Pathology Association.* J Pathol, 2019. **249**(3): p. 286-294.
112. Han, Z., et al., *Breast Cancer Multi-classification from Histopathological Images with Structured Deep Learning Model.* Sci Rep, 2017. **7**(1): p. 4172.
113. Cruz-Roa, A., et al., *Accurate and reproducible invasive breast cancer detection in whole-slide images: A Deep Learning approach for quantifying tumor extent.* Sci Rep, 2017. **7**: p. 46450.
114. Nateghi, R., H. Danyali, and M.S. Helfroush, *A deep learning approach for mitosis detection: Application in tumor proliferation prediction from whole slide images.* Artif Intell Med, 2021. **114**: p. 102048.
115. Balkenhol, M.C.A., et al., *Deep learning assisted mitotic counting for breast cancer.* Lab Invest, 2019. **99**(11): p. 1596-1606.
116. Huang, S.C., et al., *Deep neural network trained on gigapixel images improves lymph node metastasis detection in clinical settings.* Nat Commun, 2022. **13**(1): p. 3347.
117. Liu, Y., et al., *Applications of Artificial Intelligence in Breast Pathology.* Archives of Pathology & Laboratory Medicine, 2023. **147**(9): p. 1003-1013.
118. Alakwaa, F.M., K. Chaudhary, and L.X. Garmire, *Deep Learning Accurately Predicts Estrogen Receptor Status in Breast Cancer Metabolomics Data.* Journal of Proteome Research, 2018. **17**(1): p. 337-347.
119. Shamai, G., et al., *Artificial Intelligence Algorithms to Assess Hormonal Status From Tissue Microarrays in Patients With Breast Cancer.* JAMA Network Open, 2019. **2**(7): p. e197700-e197700.
120. Hu, Y., et al., *Artificial intelligence and its applications in digital hematopathology.* Blood Sci, 2022. **4**(3): p. 136-142.
121. Mustafa Yousif, D.S.M., Keluo Yao, *Artificial intelligence is the key driver for digital pathology adoption.* Clinical laporatory, 2021. **5**.

122. Beresford, M.J., G.D. Wilson, and A. Makris, *Measuring proliferation in breast cancer: practicalities and applications.* Breast cancer research : BCR, 2006. **8**(6): p. 216-216.
123. Hassan, N., et al., *SETD3 acts as a prognostic marker in breast cancer patients and modulates the viability and invasion of breast cancer cells.* Scientific Reports, 2020. **10**(1): p. 2262.
124. Veta, M., et al., *Assessment of algorithms for mitosis detection in breast cancer histopathology images.* Medical Image Analysis, 2015. **20**(1): p. 237-248.
125. Banin Hirata, B.K., et al., *Molecular Markers for Breast Cancer: Prediction on Tumor Behavior.* Disease Markers, 2014. **2014**: p. 513158.
126. van Diest, P.J., E. van der Wall, and J.P.A. Baak, *Prognostic value of proliferation in invasive breast cancer: a review.* Journal of Clinical Pathology, 2004. **57**(7): p. 675.
127. Dayal, J.H.S., et al., *Multiparameter DNA content analysis identifies distinct groups in primary breast cancer.* British Journal of Cancer, 2013. **108**(4): p. 873-880.
128. Aleskandarany, M.A., et al., *Prognostic value of proliferation assay in the luminal, HER2-positive, and triple-negative biologic classes of breast cancer.* Breast Cancer Res, 2012. **14**(1): p. R3.
129. Juríková, M., et al., *Ki67, PCNA, and MCM proteins: Markers of proliferation in the diagnosis of breast cancer.* Acta Histochem, 2016. **118**(5): p. 544-52.
130. Mercadante, A.A. and A. Kasi, *Genetics, Cancer Cell Cycle Phases*, in *StatPearls*. 2023, StatPearls Publishing

Copyright © 2023, StatPearls Publishing LLC.: Treasure Island (FL) ineligible companies. Disclosure: Anup Kasi declares no relevant financial relationships with ineligible companies.

131. Hunt, T., K. Nasmyth, and B. Novák, *The cell cycle.* Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 2011. **366**(1584): p. 3494-3497.
132. Alberts B, J.A., Lewis J, et al. Molecular Biology of the Cell. 4th edition, *Molecular Biology of the Cell. 4th edition*. An Overview of the Cell Cycle, ed. t. edition. 2002, New York: Garland Science.
133. Olar, A., et al., *Mitotic Index is an Independent Predictor of Recurrence-Free Survival in Meningioma.* Brain Pathol, 2015. **25**(3): p. 266-75.
134. Pan, X., et al., *Mitosis detection techniques in H&E stained breast cancer pathological images: A comprehensive review.* Computers & Electrical Engineering, 2021. **91**: p. 107038.
135. Kato, T.A. and J.S. Haskins, *Mitotic Index Analysis.* Methods Mol Biol, 2023. **2519**: p. 17-26.
136. Al-Janabi, S., et al., *Evaluation of mitotic activity index in breast cancer using whole slide digital images.* PLoS One, 2013. **8**(12): p. e82576.
137. Elston, C.W. and I.O. Ellis, *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**(5): p. 403-10.
138. Donovan, T.A., et al., *Mitotic Figures—Normal, Atypical, and Imposters: A Guide to Identification.* Veterinary Pathology, 2020. **58**(2): p. 243-257.
139. Rizzo, D.C., *Fundamentals of Anatomy and Physiology*. Cengage learning. Vol. 2016. 2016.
140. Johnson, A.e.a., *The Cell.* 5th edn ed. 2008, Abingdon: : Garlan Science.
141. Potapova, T. and G.J. Gorbsky, *The Consequences of Chromosome Segregation Errors in Mitosis and Meiosis.* Biology (Basel), 2017. **6**(1).
142. Gisselsson, D., *Classification of chromosome segregation errors in cancer.* Chromosoma, 2008. **117**(6): p. 511-9.
143. Yoko, M., et al., *Morphological Markers of Chromosomal Instability*, in *Chromosomal Abnormalities*, L.L. Marcelo and S. Sonia, Editors. 2017, IntechOpen: Rijeka. p. Ch. 2.
144. Ankle, M.R., A.D. Kale, and S. Charantimath, *Comparison of staining of mitotic figures by haematoxylin and eosin-and crystal violet stains, in oral epithelial dysplasia and squamous cell carcinoma.* Indian J Dent Res, 2007. **18**(3): p. 101-5.
145. Kalatova, B., et al., *Tripolar mitosis in human cells and embryos: occurrence, pathophysiology and medical implications.* Acta Histochem, 2015. **117**(1): p. 111-25.
146. Matsuda, Y., et al., *Mitotic index and multipolar mitosis in routine histologic sections as prognostic markers of pancreatic cancers: A clinicopathological study.* Pancreatology, 2016. **16**(1): p. 127-32.
147. Ohashi, R., et al., *Prognostic utility of atypical mitoses in patients with breast cancer: A comparative study with Ki67 and phosphohistone H3.* J Surg Oncol, 2018. **118**(3): p. 557-567.
148. Tvedten, H., *Atypical mitoses: morphology and classification.* Veterinary Clinical Pathology, 2009. **38**(4): p. 418-420.
149. van Diest, P.J., et al., *Reproducibility of mitosis counting in 2,469 breast cancer specimens: results from the Multicenter Morphometric Mammary Carcinoma Project.* Hum Pathol, 1992. **23**(6): p. 603-7.
150. Hartwell, L.H., et al., *From molecular to modular cell biology.* Nature, 1999. **402**(6761 Suppl): p. C47-52.
151. Emad, A.R., et al., *Review of the national external quality assessment (EQA) scheme for breast pathology in the UK.* Journal of Clinical Pathology, 2017. **70**(1): p. 51.
152. Asmaa, I., et al., *Assessment of mitotic activity in breast cancer: revisited in the digital pathology era.* Journal of Clinical Pathology, 2022. **75**(6): p. 365.
153. Bonert, M. and A.J. Tate, *Mitotic counts in breast cancer should be standardized with a uniform sample area.* Biomed Eng Online, 2017. **16**(1): p. 28.

154. Aubreville, M., et al. *Augmented Mitotic Cell Count Using Field of Interest Proposal*. in *Bildverarbeitung für die Medizin 2019*. 2019. Wiesbaden: Springer Fachmedien Wiesbaden.
155. Tabata, K., et al., *Validation of mitotic cell quantification via microscopy and multiple whole-slide scanners.* Diagnostic Pathology, 2019. **14**(1): p. 65.
156. Veta, M., et al., *Mitosis Counting in Breast Cancer: Object-Level Interobserver Agreement and Comparison to an Automatic Method.* PloS one, 2016. **11**(8): p. e0161286-e0161286.
157. Bertram, C.A., et al., *Computerized Calculation of Mitotic Count Distribution in Canine Cutaneous Mast Cell Tumor Sections: Mitotic Count Is Area Dependent.* Veterinary Pathology, 2019. **57**(2): p. 214-226.
158. Bertram, C.A., et al., *A large-scale dataset for mitotic figure assessment on whole slide images of canine cutaneous mast cell tumor.* Scientific Data, 2019. **6**(1): p. 274.
159. Al-Janabi, S., et al., *Evaluation of mitotic activity index in breast cancer using whole slide digital images.* PloS one, 2013. **8**(12): p. e82576-e82576.
160. Davidson, T., et al., *Breast cancer prognostic factors in the digital era: Comparison of Nottingham grade using whole slide images and glass slides.* Journal of Pathology Informatics, 2019. **10**(1): p. 11-11.
161. Wei, B.-R., et al., *Agreement in Histological Assessment of Mitotic Activity Between Microscopy and Digital Whole Slide Images Informs Conversion for Clinical Diagnosis.* Academic pathology, 2019. **6**: p. 2374289519859841-2374289519859841.
162. Rakha, E.A., et al., *Breast cancer histologic grading using digital microscopy: concordance and outcome association.* J Clin Pathol, 2018. **71**(8): p. 680-686.
163. Sun, X. and P.D. Kaufman, *Ki-67: more than a proliferation marker.* Chromosoma, 2018. **127**(2): p. 175-186.
164. Nielsen, T.O., et al., *Assessment of Ki67 in Breast Cancer: Updated Recommendations From the International Ki67 in Breast Cancer Working Group.* JNCI: Journal of the National Cancer Institute, 2021. **113**(7): p. 808-819.
165. Gerdes, J., et al., *Cell cycle analysis of a cell proliferation-associated human nuclear antigen defined by the monoclonal antibody Ki-67.* J Immunol, 1984. **133**(4): p. 1710-5.
166. Gerdes, J., et al., *Production of a mouse monoclonal antibody reactive with a human nuclear antigen associated with cell proliferation.* Int J Cancer, 1983. **31**(1): p. 13-20.
167. Scholzen, T. and J. Gerdes, *The Ki-67 protein: from the known and the unknown.* J Cell Physiol, 2000. **182**(3): p. 311-22.
168. Bahaddin, M.M., *A comparative study between Ki67 positive versus Ki67 negative females with breast cancer: Cross sectional study.* Annals of Medicine and Surgery, 2020. **60**: p. 232-235.
169. Verheijen, R., et al., *Ki-67 detects a nuclear matrix-associated proliferation-related antigen. II. Localization in mitotic cells and*

association with chromosomes. J Cell Sci, 1989. **92 (Pt 4)**: p. 531-40.
170. Yerushalmi, R., et al., *Ki67 in breast cancer: prognostic and predictive potential.* The Lancet Oncology, 2010. **11**(2): p. 174-183.
171. Sobecki, M., et al., *The cell proliferation antigen Ki-67 organises heterochromatin.* Elife, 2016. **5**: p. e13722.
172. Booth, D.G., et al., *Ki-67 is a PP1-interacting protein that organises the mitotic chromosome periphery.* Elife, 2014. **3**: p. e01641.
173. Andrés Sánchez, N., D. Fisher, and L. Krasinska, *Physiological functions and roles in cancer of the proliferation marker Ki-67.* Journal of Cell Science, 2022. **135**.
174. Barbareschi, M., et al., *Quantitative growth fraction evaluation with MIB1 and Ki67 antibodies in breast carcinomas.* Am J Clin Pathol, 1994. **102**(2): p. 171-5.
175. Veronese, S.M., C. Maisano, and J. Scibilia, *Comparative prognostic value of Ki-67 and MIB-1 proliferation indices in breast cancer.* Anticancer Res, 1995. **15**(6b): p. 2717-22.
176. Zabaglo, L., et al., *Comparative validation of the SP6 antibody to Ki67 in breast cancer.* J Clin Pathol, 2010. **63**(9): p. 800-4.
177. Ekholm, M., et al., *Immunohistochemical assessment of Ki67 with antibodies SP6 and MIB1 in primary breast cancer: a comparison of prognostic value and reproducibility.* Histopathology, 2014. **65**(2): p. 252-60.
178. Garwain, O., et al., *The chromatin-binding domain of Ki-67 together with p53 protects human chromosomes from mitotic damage.* Proc Natl Acad Sci U S A, 2021. **118**(32).
179. Bullwinkel, J., et al., *Ki-67 protein is associated with ribosomal RNA transcription in quiescent and proliferating cells.* J Cell Physiol, 2006. **206**(3): p. 624-35.
180. Vanneste, E., et al., *Chromosome instability is common in human cleavage-stage embryos.* Nat Med, 2009. **15**(5): p. 577-83.
181. Dowsett, M., et al., *Assessment of Ki67 in Breast Cancer: Recommendations from the International Ki67 in Breast Cancer Working Group.* JNCI: Journal of the National Cancer Institute, 2011. **103**(22): p. 1656-1664.
182. Faratian, D., et al., *Membranous and cytoplasmic staining of Ki67 is associated with HER2 and ER status in invasive breast carcinoma.* Histopathology, 2009. **54**(2): p. 254-7.
183. Cserni, G., *Analysis of membranous Ki-67 staining in breast cancer and surrounding breast epithelium.* Virchows Arch, 2018. **473**(2): p. 145-153.
184. Lindboe, C.F. and S.H. Torp, *Comparison of Ki-67 equivalent antibodies.* J Clin Pathol, 2002. **55**(6): p. 467-71.
185. Stec, R., et al., *Intensity of Nuclear Staining for Ki-67, p53 and Survivin as a New Prognostic Factor in Non-muscle Invasive Bladder Cancer.* Pathology & Oncology Research, 2020. **26**(2): p. 1211-1219.

186. Focke, C.M., T. Decker, and P.J. van Diest, *Intratumoral heterogeneity of Ki67 expression in early breast cancers exceeds variability between individual tumours.* Histopathology, 2016. **69**(5): p. 849-861.
187. Yersal, O. and S. Barutca, *Biological subtypes of breast cancer: Prognostic and therapeutic implications.* World J Clin Oncol, 2014. **5**(3): p. 412-24.
188. Aleskandarany, M.A., et al., *Impact of intratumoural heterogeneity on the assessment of Ki67 expression in breast cancer.* Breast Cancer Res Treat, 2016. **158**(2): p. 287-95.
189. Lashen, A., et al., *Ki67 assessment in invasive luminal breast cancer: A comparative study between different scoring methods.* Histopathology, 2022. **n/a**(n/a).
190. Kalvala, J., et al., *Concordance between core needle biopsy and surgical excision specimens for Ki-67 in breast cancer – a systematic review of the literature.* Histopathology, 2022. **80**(3): p. 468-484.
191. Romero, Q., et al., *Ki67 proliferation in core biopsies versus surgical samples - a model for neo-adjuvant breast cancer studies.* BMC Cancer, 2011. **11**: p. 341.
192. Gandini, S., et al., *Association of molecular subtypes with Ki-67 changes in untreated breast cancer patients undergoing pre-surgical trials.* Ann Oncol, 2014. **25**(3): p. 618-623.
193. Kwok, T.C., et al., *Histological grading of breast cancer on needle core biopsy: the role of immunohistochemical assessment of proliferation.* Histopathology, 2010. **57**(2): p. 212-9.
194. Wolff, A.C., et al., *Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update.* J Clin Oncol, 2013. **31**(31): p. 3997-4013.
195. Hammond, M.E., et al., *American Society of Clinical Oncology/College Of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer.* J Clin Oncol, 2010. **28**(16): p. 2784-95.
196. Arima, N., et al., *Abstract P1-02-01: Pre-analytical setting is critical for an assessment of the Ki-67 labeling index for breast cancer.* Cancer Research, 2013. **73**(24 Supplement): p. P1-02-01.
197. Combs, S.E., et al., *Loss of antigenicity with tissue age in breast cancer.* Lab Invest, 2016. **96**(3): p. 264-9.
198. Nielsen, T.O., et al., *Assessment of Ki67 in Breast Cancer: Updated Recommendations From the International Ki67 in Breast Cancer Working Group.* JNCI: Journal of the National Cancer Institute, 2020.
199. Li, L.T., et al., *Ki67 is a promising molecular target in the diagnosis of cancer (review).* Mol Med Rep, 2015. **11**(3): p. 1566-72.

200. Dowsett, M., et al., *Assessment of Ki67 in breast cancer: recommendations from the International Ki67 in Breast Cancer working group.* J Natl Cancer Inst, 2011. **103**(22): p. 1656-64.
201. Hashmi, A.A., et al., *Ki67 index in intrinsic breast cancer subtypes and its association with prognostic parameters.* BMC Research Notes, 2019. **12**(1): p. 605.
202. Kitson, S., et al., *Ki-67 in endometrial cancer: scoring optimization and prognostic relevance for window studies.* Modern Pathology, 2017. **30**(3): p. 459-468.
203. Aleskandarany, M.A., et al., *Impact of intratumoural heterogeneity on the assessment of Ki67 expression in breast cancer.* Breast Cancer Research and Treatment, 2016. **158**(2): p. 287-295.
204. Kadivar, M. and F. Aram, *Assessment of Ki67 in Breast Cancer: A Comparison Between the Eye-10 Method, Stepwise Counting Strategy, and International System of Ki67 Evaluation.* Iran J Pathol, 2020. **15**(1): p. 13-18.
205. Gudlaugsson, E., et al., *Comparison of the effect of different techniques for measurement of Ki67 proliferation on reproducibility and prognosis prediction accuracy in breast cancer.* Histopathology, 2012. **61**(6): p. 1134-44.
206. Laurinavicius, A., et al., *A methodology to ensure and improve accuracy of Ki67 labelling index estimation by automated digital image analysis in breast cancer tissue.* Breast Cancer Res, 2014. **16**(2): p. R35.
207. Mu, K., et al., *A standardized method for quantifying proliferation by Ki-67 and cyclin A immunohistochemistry in breast cancer.* Annals of Diagnostic Pathology, 2015. **19**(4): p. 243-248.
208. Lashen, A.G., et al., *Assessment of proliferation in breast cancer: cell cycle or mitosis? An observational study.* Histopathology, 2021. **79**(6): p. 1087-1098.
209. Nielsen, T.O., et al., *Assessment of Ki67 in Breast Cancer: Updated Recommendations From the International Ki67 in Breast Cancer Working Group.* J Natl Cancer Inst, 2021. **113**(7): p. 808-819.
210. Zhong, F., et al., *A Comparison of Visual Assessment and Automated Digital Image Analysis of Ki67 Labeling Index in Breast Cancer.* PLoS One, 2016. **11**(2): p. e0150505.
211. Shui, R., et al., *An Interobserver Reproducibility Analysis of Ki67 Visual Assessment in Breast Cancer.* PLOS ONE, 2015. **10**(5): p. e0125131.
212. Aman, N.A., et al., *Immunohistochemical Evaluation of Ki-67 and Comparison with Clinicopathologic Factors in Breast Carcinomas.* Asian Pac J Cancer Prev, 2019. **20**(1): p. 73-79.
213. Goldhirsch, A., 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**(8): p. 1736-47.

214. Cheang, M.C., et al., *Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer.* J Natl Cancer Inst, 2009. **101**(10): p. 736-50.
215. Coates, A.S., et al., *Tailoring therapies--improving the management of early breast cancer: St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2015.* Ann Oncol, 2015. **26**(8): p. 1533-46.
216. Petrelli, F., et al., *Prognostic value of different cut-off levels of Ki-67 in breast cancer: a systematic review and meta-analysis of 64,196 patients.* Breast Cancer Res Treat, 2015. **153**(3): p. 477-91.
217. Abubakar, M., et al., *Prognostic value of automated KI67 scoring in breast cancer: a centralised evaluation of 8088 patients from 10 study groups.* Breast Cancer Research, 2016. **18**(1): p. 104.
218. Noske, A., et al., *Risk stratification in luminal-type breast cancer: Comparison of Ki-67 with EndoPredict test results.* Breast, 2020. **49**: p. 101-107.
219. Pathmanathan, N. and R.L. Balleine, *Ki67 and proliferation in breast cancer.* J Clin Pathol, 2013. **66**(6): p. 512-6.
220. Van den Berg, E.J., et al., *KI67 immunohistochemistry quantification in breast carcinoma: A comparison of visual estimation, counting and Immunoratio©.* Applied immunohistochemistry & molecular morphology: AIMM, 2021. **29**(2): p. 105.
221. Aleskandarany, M.A., et al., *MIB1/Ki-67 labelling index can classify grade 2 breast cancer into two clinically distinct subgroups.* Breast Cancer Res Treat, 2011. **127**(3): p. 591-9.
222. Wu, Q., et al., *Prognostic Value of Ki-67 in Patients With Resected Triple-Negative Breast Cancer: A Meta-Analysis.* Front Oncol, 2019. **9**: p. 1068.
223. Zhu, X., et al., *The prognostic and predictive potential of Ki-67 in triple-negative breast cancer.* Scientific Reports, 2020. **10**(1): p. 225.
224. Santisteban, M., et al., *Ki67: a time-varying biomarker of risk of breast cancer in atypical hyperplasia.* Breast Cancer Res Treat, 2010. **121**(2): p. 431-7.
225. Smith, I., et al., *Long-term outcome and prognostic value of Ki67 after perioperative endocrine therapy in postmenopausal women with hormone-sensitive early breast cancer (POETIC): an open-label, multicentre, parallel-group, randomised, phase 3 trial.* Lancet Oncol, 2020. **21**(11): p. 1443-1454.
226. Giordano, S.H., R.A. Freedman, and M.R. Somerfield, *Abemaciclib With Endocrine Therapy in the Treatment of High-Risk Early Breast Cancer: ASCO Optimal Adjuvant Chemotherapy and Targeted Therapy Guideline Rapid Recommendation Update.* J Clin Oncol, 2022. **40**(3): p. 307-309.
227. Rakha, E.A., et al., *Assessment of Predictive Biomarkers in Breast Cancer: Challenges and Updates.* Pathobiology, 2022. **89**(5): p. 263-277.

228. Rakha, E.A., et al., *Breast cancer prognostic classification in the molecular era: the role of histological grade.* Breast Cancer Res, 2010. **12**(4): p. 207.
229. Pinder, S.E., et al., *Assessment of the new proliferation marker MIB1 in breast carcinoma using image analysis: associations with other prognostic factors and survival.* Br J Cancer, 1995. **71**(1): p. 146-9.
230. Kanyılmaz, G., et al., *Prognostic Importance of Ki-67 in Breast Cancer and Its Relationship with Other Prognostic Factors.* Eur J Breast Health, 2019. **15**(4): p. 256-261.
231. Liang, Q., et al., *Effect of Ki-67 Expression Levels and Histological Grade on Breast Cancer Early Relapse in Patients with Different Immunohistochemical-based Subtypes.* Scientific Reports, 2020. **10**(1): p. 7648.
232. Inwald, E.C., 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 Res Treat, 2013. **139**(2): p. 539-52.
233. Petric, M., et al., *Correlation between Ki67 and histological grade in breast cancer patients treated with preoperative chemotherapy.* Asian Pac J Cancer Prev, 2014. **15**(23): p. 10277-80.
234. Lashen, A., et al., *The characteristics and clinical significance of atypical mitosis in breast cancer.* Modern Pathology, 2022. **35**(10): p. 1341-1348.
235. Ibrahim, A., et al., *Defining the area of mitoses counting in invasive breast cancer using whole slide image.* Modern Pathology, 2022. **35**(6): p. 739-748.
236. Pathmanathan, N., et al., *The prognostic value of Ki67 in systemically untreated patients with node-negative breast cancer.* Journal of Clinical Pathology, 2014. **67**(3): p. 222.
237. Mathieu, M.C., et al., *The poor responsiveness of infiltrating lobular breast carcinomas to neoadjuvant chemotherapy can be explained by their biological profile.* Eur J Cancer, 2004. **40**(3): p. 342-51.
238. Yerushalmi, R., et al., *Ki67 in breast cancer: prognostic and predictive potential.* Lancet Oncol, 2010. **11**(2): p. 174-83.
239. Dirican, E. and E. Kiliç, *A Machine Learning Approach for the Association of ki-67 Scoring with Prognostic Factors.* J Oncol, 2018. **2018**: p. 1912438.
240. Yin, Y., et al., *The levels of Ki-67 positive are positively associated with lymph node metastasis in invasive ductal breast cancer.* Cell Biochem Biophys, 2014. **70**(2): p. 1145-51.
241. Asadi, M., et al., *Evaluation of Involvement of Axillary Lymph Nodes with Ki-67 Expression in Patients with Breast Cancer.* Int J Cancer Manag, 2018. **11**(8): p. e66567.
242. Chu, Z., et al., *Association between axillary lymph node status and Ki67 labeling index in triple-negative medullary breast carcinoma.* Japanese Journal of Clinical Oncology, 2015. **45**(7): p. 637-641.

243. Tawfik, K., et al., *Ki-67 expression in axillary lymph node metastases in breast cancer is prognostically significant.* Human Pathology, 2013. **44**(1): p. 39-46.
244. Asvadi, T., et al., *Ki-67 status in patients with primary breast cancer and its relationship with other prognostic factors.* Biomedical Research and Therapy, 2019. **6**: p. 2986-2991.
245. Wiesner, F.G., et al., *Ki-67 as a prognostic molecular marker in routine clinical use in breast cancer patients.* Breast, 2009. **18**(2): p. 135-41.
246. Liu, S., et al., *Measures of cell turnover (proliferation and apoptosis) and their association with survival in breast cancer.* Clin Cancer Res, 2001. **7**(6): p. 1716-23.
247. Payandeh, M., et al., *Correlations between HER2 Expression and Other Prognostic Factors in Breast Cancer: Inverse Relations with the Ki-67 Index and P53 Status.* Asian Pacific journal of cancer prevention: APJCP, 2016. **17**: p. 1015-1018.
248. Shokouh, T.Z., A. Ezatollah, and P. Barand, *Interrelationships Between Ki67, HER2/neu, p53, ER, and PR Status and Their Associations With Tumor Grade and Lymph Node Involvement in Breast Carcinoma Subtypes: Retrospective-Observational Analytical Study.* Medicine (Baltimore), 2015. **94**(32): p. e1359.
249. Ragab, H.M., et al., *Assessment of Ki-67 as a potential biomarker in patients with breast cancer.* Journal of Genetic Engineering and Biotechnology, 2018. **16**(2): p. 479-484.
250. Yin, Y., et al., *The Levels of Ki-67 Positive are Positively Associated with Lymph Node Metastasis in Invasive Ductal Breast Cancer.* Cell Biochemistry and Biophysics, 2014. **70**(2): p. 1145-1151.
251. Inic, Z., et al., *Difference between Luminal A and Luminal B Subtypes According to Ki-67, Tumor Size, and Progesterone Receptor Negativity Providing Prognostic Information.* Clin Med Insights Oncol, 2014. **8**: p. 107-11.
252. Miyashita, M., et al., *Histopathological subclassification of triple negative breast cancer using prognostic scoring system: five variables as candidates.* Virchows Archiv : an international journal of pathology, 2011. **458**(1): p. 65-72.
253. Fallah, P., et al., *Can high Ki67 predict distant recurrence in early-stage breast cancer with low Oncotype Dx score?* Journal of Clinical Oncology, 2021. **39**(15_suppl): p. e12561-e12561.
254. Turner, B.M., et al., *Use of modified Magee equations and histologic criteria to predict the Oncotype DX recurrence score.* Modern Pathology, 2015. **28**(7): p. 921-931.
255. Lashen, A., et al., *Evaluation Oncotype DX® 21-Gene Recurrence Score and Clinicopathological Parameters: A single institutional experience.* Histopathology, 2023.
256. Klein, M.E., et al., *Prediction of the Oncotype DX recurrence score: use of pathology-generated equations derived by linear regression analysis.* Modern Pathology, 2013. **26**(5): p. 658-664.

257. Paik, S., et al., *A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer.* N Engl J Med, 2004. **351**(27): p. 2817-26.
258. Sgroi, D.C., et al., *Prediction of late distant recurrence in patients with oestrogen-receptor-positive breast cancer: a prospective comparison of the breast-cancer index (BCI) assay, 21-gene recurrence score, and IHC4 in the TransATAC study population.* Lancet Oncol, 2013. **14**(11): p. 1067-1076.
259. Mulla, N., et al., *Determine the relationship between Ki 67 and Oncotype DX.* Journal of Clinical Oncology, 2016. **34**(15_suppl): p. e12017-e12017.
260. Namuche, F., et al., *214P - Ki67 as an important predictor for oncotype Dx recurrence score risk groups in early breast cancer.* Annals of Oncology, 2018. **29**: p. viii68.
261. Sahebjam, S., et al., *Ki 67 is a major, but not the sole determinant of Oncotype Dx recurrence score.* Br J Cancer, 2011. **105**(9): p. 1342-5.
262. Rakha, E.A., et al., *Prognostic stratification of oestrogen receptor-positive HER2-negative lymph node-negative class of breast cancer.* Histopathology, 2017. **70**(4): p. 622-631.
263. Zhu, X., et al., *The prognostic and predictive potential of Ki-67 in triple-negative breast cancer.* Sci Rep, 2020. **10**(1): p. 225.
264. de Azambuja, E., et al., *Ki-67 as prognostic marker in early breast cancer: a meta-analysis of published studies involving 12,155 patients.* Br J Cancer, 2007. **96**(10): p. 1504-13.
265. Criscitiello, C., et al., *High Ki-67 score is indicative of a greater benefit from adjuvant chemotherapy when added to endocrine therapy in Luminal B HER2 negative and node-positive breast cancer.* The Breast, 2014. **23**(1): p. 69-75.
266. Nielsen, T.O., et al., *High-Risk Premenopausal Luminal A Breast Cancer Patients Derive no Benefit from Adjuvant Cyclophosphamide-based Chemotherapy: Results from the DBCG77B Clinical Trial.* Clinical Cancer Research, 2017. **23**(4): p. 946.
267. Pondé, N.F., D. Zardavas, and M. Piccart, *Progress in adjuvant systemic therapy for breast cancer.* Nature Reviews Clinical Oncology, 2019. **16**(1): p. 27-44.
268. Teven, C.M., et al., *Systemic Therapy for Early-Stage Breast Cancer: What the Plastic Surgeon Should Know.* Eplasty, 2017. **17**: p. e7.
269. Viale, G., et al., *Which patients benefit most from adjuvant aromatase inhibitors? Results using a composite measure of prognostic risk in the BIG 1-98 randomized trial.* Ann Oncol, 2011. **22**(10): p. 2201-7.
270. Penault-Llorca, F., et al., *Ki67 expression and docetaxel efficacy in patients with estrogen receptor-positive breast cancer.* J Clin Oncol, 2009. **27**(17): p. 2809-15.
271. Karlsson, P., et al., *Long-term results of International Breast Cancer Study Group Trial VIII: adjuvant chemotherapy plus*

goserelin compared with either therapy alone for premenopausal patients with node-negative breast cancer. Ann Oncol, 2011. **22**(10): p. 2216-26.
272. Arana Echarri, A., et al., *A Phenomic Perspective on Factors Influencing Breast Cancer Treatment: Integrating Aging and Lifestyle in Blood and Tissue Biomarker Profiling.* Frontiers in Immunology, 2021. **11**: p. 3640.
273. Dowsett, M., et al., *Prognostic value of Ki67 expression after short-term presurgical endocrine therapy for primary breast cancer.* J Natl Cancer Inst, 2007. **99**(2): p. 167-70.
274. Chen, X., et al., *The predictive value of Ki-67 before neoadjuvant chemotherapy for breast cancer: a systematic review and meta-analysis.* Future Oncology, 2017. **13**(9): p. 843-857.
275. von Minckwitz, G., et al., *Ki67 measured after neoadjuvant chemotherapy for primary breast cancer.* Clinical cancer research : an official journal of the American Association for Cancer Research, 2013. **19**(16): p. 4521-4531.
276. Chen, C., et al., *Decrease in the Ki67 index during neoadjuvant chemotherapy predicts favorable relapse-free survival in patients with locally advanced breast cancer.* Cancer Biol Med, 2019. **16**(3): p. 575-586.
277. Tao, M., et al., *Ki-67 labeling index is a predictive marker for a pathological complete response to neoadjuvant chemotherapy in breast cancer: A meta-analysis.* Medicine (Baltimore), 2017. **96**(51): p. e9384.
278. Zhou, B., D.Q. Yang, and F. Xie, *Biological markers as predictive factors of response to neoadjuvant taxanes and anthracycline chemotherapy in breast carcinoma.* Chin Med J (Engl), 2008. **121**(5): p. 387-91.
279. Wei, Y., et al., *[Association between hormone receptors and response to neoadjuvant anthracycline-based chemotherapy in breast cancer patients].* Beijing Da Xue Xue Bao Yi Xue Ban, 2007. **39**(5): p. 481-3.
280. Feng, M., et al., *Automated quantitative analysis of Ki-67 staining and HE images recognition and registration based on whole tissue sections in breast carcinoma.* Diagn Pathol, 2020. **15**(1): p. 65.
281. Suciu, C., et al., *Semi-automated evaluation of Ki-67 index in invasive ductal carcinoma of the breast.* Oncol Lett, 2014. **7**(1): p. 107-114.
282. Niazi, M.K.K., A.V. Parwani, and M.N. Gurcan, *Digital pathology and artificial intelligence.* Lancet Oncol, 2019. **20**(5): p. e253-e261.
283. Tizhoosh, H.R. and L. Pantanowitz, *Artificial Intelligence and Digital Pathology: Challenges and Opportunities.* J Pathol Inform, 2018. **9**: p. 38.
284. Stålhammar, G., et al., *Digital image analysis outperforms manual biomarker assessment in breast cancer.* Mod Pathol, 2016. **29**(4): p. 318-29.

285. Abubakar, M., et al., *Prognostic value of automated KI67 scoring in breast cancer: a centralised evaluation of 8088 patients from 10 study groups.* Breast Cancer Res, 2016. **18**(1): p. 104.
286. Serna, G., et al., *Sequential immunohistochemistry and virtual image reconstruction using a single slide for quantitative KI67 measurement in breast cancer.* Breast, 2020. **53**: p. 102-110.
287. Li, Z., et al., *The Role of Progesterone Receptors in Breast Cancer.* Drug Des Devel Ther, 2022. **16**: p. 305-314.
288. Allred, D.C., *Issues and updates: evaluating estrogen receptor-alpha, progesterone receptor, and HER2 in breast cancer.* Mod Pathol, 2010. **23 Suppl 2**: p. S52-9.
289. Peng, J., S. Sengupta, and V.C. Jordan, *Potential of selective estrogen receptor modulators as treatments and preventives of breast cancer.* Anticancer Agents Med Chem, 2009. **9**(5): p. 481-99.
290. Li, Y., et al., *Clinicopathological Characteristics and Breast Cancer-Specific Survival of Patients With Single Hormone Receptor-Positive Breast Cancer.* JAMA Netw Open, 2020. **3**(1): p. e1918160.
291. Mohammed, H., et al., *Progesterone receptor modulates ERα action in breast cancer.* Nature, 2015. **523**(7560): p. 313-7.
292. Rakha, E.A., J.S. Reis-Filho, and I.O. Ellis, *Combinatorial biomarker expression in breast cancer.* Breast Cancer Res Treat, 2010. **120**(2): p. 293-308.
293. Allison, K.H., et al., *Estrogen and Progesterone Receptor Testing in Breast Cancer: American Society of Clinical Oncology/College of American Pathologists Guideline Update.* Arch Pathol Lab Med, 2020. **144**(5): p. 545-563.
294. Kurozumi, S., et al., *ER, PgR, Ki67, p27(Kip1), and histological grade as predictors of pathological complete response in patients with HER2-positive breast cancer receiving neoadjuvant chemotherapy using taxanes followed by fluorouracil, epirubicin, and cyclophosphamide concomitant with trastuzumab.* BMC Cancer, 2015. **15**: p. 622.
295. Boland, M.R., et al., *Impact of progesterone receptor status on response to neoadjuvant chemotherapy in estrogen receptor-positive breast cancer patients.* J Surg Oncol, 2020. **122**(5): p. 861-868.
296. Alferez, D.G., et al., *The Role of Steroid Hormones in Breast and Effects on Cancer Stem Cells.* Curr Stem Cell Rep, 2018. **4**(1): p. 81-94.
297. Ellis, I.O., et al., *Pathology reporting of breast disease in surgical excision specimens incorporating the dataset for histological reporting of breast cancer*. June 2016, London: The Royal College of Pathologists. 160.
298. Tran, B. and P.L. Bedard, *Luminal-B breast cancer and novel therapeutic targets.* Breast Cancer Res, 2011. **13**(6): p. 221.

299. Creighton, C.J., et al., *Molecular profiles of progesterone receptor loss in human breast tumors.* Breast Cancer Res Treat, 2009. **114**(2): p. 287-99.
300. Selli, C., J.M. Dixon, and A.H. Sims, *Accurate prediction of response to endocrine therapy in breast cancer patients: current and future biomarkers.* Breast Cancer Res, 2016. **18**(1): p. 118.
301. Kang, Y.J., et al., *Ki-67 Expression is a Significant Prognostic Factor Only When Progesterone Receptor Expression is Low in Estrogen Receptor-Positive and HER2-Negative Early Breast Cancer.* J Oncol, 2019. **2019**: p. 7386734.
302. Bravaccini, S., et al., *The impact of progesterone receptor expression on prognosis of patients with rapidly proliferating, hormone receptor-positive early breast cancer: a post hoc analysis of the IBIS 3 trial.* Therapeutic Advances in Medical Oncology, 2020. **12**: p. 1758835919888999.
303. Bravaccini, S., et al., *The prognostic role of progesterone receptor in patients with rapidly proliferating, hormone receptor-positive early breast cancer.* Journal of Clinical Oncology, 2019. **37**(15_suppl): p. 545-545.
304. Yoder, R., et al., *Impact of low versus negative estrogen/progesterone receptor status on clinico-pathologic characteristics and survival outcomes in HER2-negative breast cancer.* npj Breast Cancer, 2022. **8**(1): p. 80.
305. Purdie, C.A., et al., *Progesterone receptor expression is an independent prognostic variable in early breast cancer: a population-based study.* Br J Cancer, 2014. **110**(3): p. 565-72.
306. Lumachi, F., D.A. Santeufemia, and S.M. Basso, *Current medical treatment of estrogen receptor-positive breast cancer.* World J Biol Chem, 2015. **6**(3): p. 231-9.
307. Sanford, R.A., et al., *High incidence of germline BRCA mutation in patients with ER low-positive/PR low-positive/HER-2 neu negative tumors.* Cancer, 2015. **121**(19): p. 3422-7.
308. Prat, A., et al., *Prognostic significance of progesterone receptor-positive tumor cells within immunohistochemically defined luminal A breast cancer.* J Clin Oncol, 2013. **31**(2): p. 203-9.
309. Yao, N., et al., *Prognostic Impact of Progesterone Receptor Status in Chinese Estrogen Receptor Positive Invasive Breast Cancer Patients.* J Breast Cancer, 2017. **20**(2): p. 160-169.
310. Goldhirsch, A., et al., *Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013.* Ann Oncol, 2013. **24**(9): p. 2206-23.
311. Tellez-Gabriel, M., et al., *Tumour Heterogeneity: The Key Advantages of Single-Cell Analysis.* Int J Mol Sci, 2016. **17**(12).
312. Nalejska, E., E. Mączyńska, and M.A. Lewandowska, *Prognostic and predictive biomarkers: tools in personalized oncology.* Mol Diagn Ther, 2014. **18**(3): p. 273-84.
313. Morgan, M.A. and A. Shilatifard, *Chromatin signatures of cancer.* Genes Dev, 2015. **29**(3): p. 238-49.

314. Scaffidi, P., *Histone H1 alterations in cancer.* Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms, 2016. **1859**(3): p. 533-539.
315. Horwitz, K.B. and W.L. McGuire, *Estrogen control of progesterone receptor in human breast cancer. Correlation with nuclear processing of estrogen receptor.* Journal of Biological Chemistry, 1978. **253**(7): p. 2223-2228.
316. Balhorn, R., R. Chalkley, and D. Granner, *Lysine-rich histone phosphorylation. A positive correlation with cell replication.* Biochemistry, 1972. **11**(6): p. 1094-8.
317. Gréen, A., et al., *Histone H1 interphase phosphorylation becomes largely established in G1 or early S phase and differs in G1 between T-lymphoblastoid cells and normal T cells.* Epigenetics Chromatin, 2011. **4**: p. 15.
318. Marzluff, W.F., *Metazoan replication-dependent histone mRNAs: a distinct set of RNA polymerase II transcripts.* Curr Opin Cell Biol, 2005. **17**(3): p. 274-80.
319. Choi, J., et al., *DNA Methylation and Histone H1 Jointly Repress Transposable Elements and Aberrant Intragenic Transcripts.* Molecular Cell, 2020. **77**(2): p. 310-323.e7.
320. Roque, A., I. Ponte, and P. Suau, *The subtype-specific role of histone H1.0 in cancer cell differentiation and intratumor heterogeneity.* Translational Cancer Research, 2017: p. S414-S417.
321. Yusufova, N., et al., *Histone H1 loss drives lymphoma by disrupting 3D chromatin architecture.* Nature, 2021. **589**(7841): p. 299-305.
322. Funayama, R., et al., *Loss of linker histone H1 in cellular senescence.* J Cell Biol, 2006. **175**(6): p. 869-80.
323. Sato, S., et al., *Histone H1 expression in human prostate cancer tissues and cell lines.* Pathology International, 2012. **62**(2): p. 84-92.
324. Markant, S.L., et al., *Targeting sonic hedgehog-associated medulloblastoma through inhibition of Aurora and Polo-like kinases.* Cancer Res, 2013. **73**(20): p. 6310-22.
325. Takaki, T., et al., *Polo-like kinase 1 reaches beyond mitosis--cytokinesis, DNA damage response, and development.* Curr Opin Cell Biol, 2008. **20**(6): p. 650-60.
326. Schmucker, S. and I. Sumara, *Molecular dynamics of PLK1 during mitosis.* Mol Cell Oncol, 2014. **1**(2): p. e954507.
327. He, Z., et al., *Polo-like kinase 1 contributes to the tumorigenicity of BEL-7402 hepatoma cells via regulation of Survivin expression.* Cancer Lett, 2011. **303**(2): p. 92-8.
328. Park, J.S., et al., *Induction of antitumor immunity using dendritic cells electroporated with Polo-like kinase 1 (Plk1) mRNA in murine tumor models.* Cancer Sci, 2011. **102**(8): p. 1448-54.
329. Gutteridge, R.E., et al., *Plk1 Inhibitors in Cancer Therapy: From Laboratory to Clinics.* Mol Cancer Ther, 2016. **15**(7): p. 1427-35.

330. Gjertsen, B.T. and P. Schöffski, *Discovery and development of the Polo-like kinase inhibitor volasertib in cancer therapy.* Leukemia, 2015. **29**(1): p. 11-9.
331. Cunningham, C.E., et al., *The CINs of Polo-Like Kinase 1 in Cancer.* Cancers (Basel), 2020. **12**(10).
332. Luo, J., N.L. Solimini, and S.J. Elledge, *Principles of cancer therapy: oncogene and non-oncogene addiction.* Cell, 2009. **136**(5): p. 823-37.
333. Jang, Y.J., Y.S. Kim, and W.H. Kim, *Oncogenic effect of Polo-like kinase 1 expression in human gastric carcinomas.* Int J Oncol, 2006. **29**(3): p. 589-94.
334. Takahashi, T., et al., *Polo-like kinase 1 (PLK1) is overexpressed in primary colorectal cancers.* Cancer Sci, 2003. **94**(2): p. 148-52.
335. He, Z.L., et al., *Overexpression of polo-like kinase1 predicts a poor prognosis in hepatocellular carcinoma patients.* World J Gastroenterol, 2009. **15**(33): p. 4177-82.
336. Weichert, W., et al., *Polo-like kinase 1 is overexpressed in prostate cancer and linked to higher tumor grades.* Prostate, 2004. **60**(3): p. 240-5.
337. Weichert, W., et al., *Polo-like kinase isoform expression is a prognostic factor in ovarian carcinoma.* Br J Cancer, 2004. **90**(4): p. 815-21.
338. Weichert, W., et al., *Polo-like kinase isoforms in breast cancer: expression patterns and prognostic implications.* Virchows Arch, 2005. **446**(4): p. 442-50.
339. King, S.I., et al., *Immunohistochemical detection of Polo-like kinase-1 (PLK1) in primary breast cancer is associated with TP53 mutation and poor clinical outcom.* Breast Cancer Res, 2012. **14**(2): p. R40.
340. Donizy, P., et al., *Augmented expression of Polo-like kinase 1 is a strong predictor of shorter cancer-specific overall survival in early stage breast cancer at 15-year follow-up.* Oncol Lett, 2016. **12**(3): p. 1667-1674.
341. Takeshita, T., et al., *High expression of polo-like kinase 1 is associated with TP53 inactivation, DNA repair deficiency, and worse prognosis in ER positive Her2 negative breast cancer.* Am J Transl Res, 2019. **11**(10): p. 6507-6521.
342. Song, H.Y., et al., *DNA replication: Mechanisms and therapeutic interventions for diseases.* MedComm (2020), 2023. **4**(1): p. e210.
343. Forsburg, S.L., *Eukaryotic MCM proteins: beyond replication initiation.* Microbiol Mol Biol Rev, 2004. **68**(1): p. 109-31.
344. Toyokawa, G., et al., *Minichromosome Maintenance Protein 7 is a potential therapeutic target in human cancer and a novel prognostic marker of non-small cell lung cancer.* Mol Cancer, 2011. **10**: p. 65.
345. Nowińska, K. and P. Dzięgiel, *[The role of MCM proteins in cell proliferation and tumorigenesis].* Postepy Hig Med Dosw (Online), 2010. **64**: p. 627-35.

346. Padmanabhan, V., et al., *DNA replication regulation protein Mcm7 as a marker of proliferation in prostate cancer.* J Clin Pathol, 2004. **57**(10): p. 1057-62.
347. Qu, K., et al., *MCM7 promotes cancer progression through cyclin D1-dependent signaling and serves as a prognostic marker for patients with hepatocellular carcinoma.* Cell Death Dis, 2017. **8**(2): p. e2603.
348. Wei, Q., et al., *Phosphorylation of Minichromosome Maintenance Protein 7 (MCM7) by Cyclin/Cyclin-dependent Kinase Affects Its Function in Cell Cycle Regulation*.* Journal of Biological Chemistry, 2013. **288**(27): p. 19715-19725.
349. Ren, B., et al., *MCM7 amplification and overexpression are associated with prostate cancer progression.* Oncogene, 2006. **25**(7): p. 1090-8.
350. Qiu, Y.T., et al., *MCM7 amplification and overexpression promote cell proliferation, colony formation and migration in esophageal squamous cell carcinoma by activating the AKT1/mTOR signaling pathway.* Oncol Rep, 2017. **37**(6): p. 3590-3596.
351. Aleskandarany, M.A., et al., *Tumour Heterogeneity of Breast Cancer: From Morphology to Personalised Medicine.* Pathobiology, 2018. **85**(1-2): p. 23-34.
352. Fumagalli, C. and M. Barberis, *Breast Cancer Heterogeneity.* Diagnostics (Basel), 2021. **11**(9).
353. McQuerry, J.A., et al., *Mechanisms and clinical implications of tumor heterogeneity and convergence on recurrent phenotypes.* J Mol Med (Berl), 2017. **95**(11): p. 1167-1178.
354. Wahab, N., et al., *AI-enabled routine H&E image based prognostic marker for early-stage luminal breast cancer.* NPJ Precis Oncol, 2023. **7**(1): p. 122.
355. Natrajan, R., et al., *Tiling Path Genomic Profiling of Grade 3 Invasive Ductal Breast Cancers.* Clinical Cancer Research, 2009. **15**(8): p. 2711-2722.
356. Loibl, S., et al., *Mutational Diversity and Therapy Response in Breast Cancer: A Sequencing Analysis in the Neoadjuvant GeparSepto Trial.* Clin Cancer Res, 2019. **25**(13): p. 3986-3995.
357. Navin, N., et al., *Inferring tumor progression from genomic heterogeneity.* Genome Res, 2010. **20**(1): p. 68-80.
358. Martelotto, L.G., et al., *Breast cancer intra-tumor heterogeneity.* Breast Cancer Research, 2014. **16**(3): p. 210.
359. Beyes, S., N.G. Bediaga, and A. Zippo, *An Epigenetic Perspective on Intra-Tumour Heterogeneity: Novel Insights and New Challenges from Multiple Fields.* Cancers (Basel), 2021. **13**(19).
360. Park, S.Y., et al., *Cellular and genetic diversity in the progression of in situ human breast carcinomas to an invasive phenotype.* J Clin Invest, 2010. **120**(2): p. 636-44.
361. Ramón y Cajal, S., et al., *Clinical implications of intratumor heterogeneity: challenges and opportunities.* Journal of Molecular Medicine, 2020. **98**(2): p. 161-177.

362. Denisov, E.V., et al., *Intratumoral morphological heterogeneity of breast cancer: neoadjuvant chemotherapy efficiency and multidrug resistance gene expression.* Scientific Reports, 2014. **4**(1): p. 4709.
363. Anderson, N.M. and M.C. Simon, *The tumor microenvironment.* Curr Biol, 2020. **30**(16): p. R921-r925.
364. Wu, T. and Y. Dai, *Tumor microenvironment and therapeutic response.* Cancer Lett, 2017. **387**: p. 61-68.
365. Laurinavicius, A., et al., *Machine-Learning–Based Evaluation of Intratumoral Heterogeneity and Tumor-Stroma Interface for Clinical Guidance.* The American Journal of Pathology, 2021. **191**(10): p. 1724-1731.
366. Gerlinger, M., et al., *Intratumor heterogeneity and branched evolution revealed by multiregion sequencing.* N Engl J Med, 2012. **366**(10): p. 883-892.
367. Navin, N., et al., *Tumour evolution inferred by single-cell sequencing.* Nature, 2011. **472**(7341): p. 90-4.
368. Shah, S.P., et al., *The clonal and mutational evolution spectrum of primary triple-negative breast cancers.* Nature, 2012. **486**(7403): p. 395-9.
369. Yang, F., et al., *Intratumor heterogeneity predicts metastasis of triple-negative breast cancer.* Carcinogenesis, 2017. **38**(9): p. 900-909.
370. Russnes, H.G., et al., *Insight into the heterogeneity of breast cancer through next-generation sequencing.* J Clin Invest, 2011. **121**(10): p. 3810-8.
371. Bera, K., et al., *Artificial intelligence in digital pathology — new tools for diagnosis and precision oncology.* Nature Reviews Clinical Oncology, 2019. **16**(11): p. 703-715.
372. Hewitt, S.M., *Tissue microarrays as a tool in the discovery and validation of predictive biomarkers.* Methods Mol Biol, 2012. **823**: p. 201-14.
373. Gologan, D., et al., *Quality Assurance and Cost Reduction in Histopathology Laboratories Using Tissue Microarrays.* Vet Sci, 2023. **10**(4).
374. Sauter, G. and M. Mirlacher, *Tissue microarrays for predictive molecular pathology.* J Clin Pathol, 2002. **55**(8): p. 575-6.
375. Dancau, A.M., et al., *Tissue Microarrays.* Methods Mol Biol, 2016. **1381**: p. 53-65.
376. Boros, M., et al., *Intratumoral Heterogeneity for Ki-67 Index in Invasive Breast Carcinoma: A Study on 131 Consecutive Cases.* Appl Immunohistochem Mol Morphol, 2017. **25**(5): p. 338-340.
377. Allott, E.H., et al., *Intratumoral heterogeneity as a source of discordance in breast cancer biomarker classification.* Breast Cancer Research, 2016. **18**(1): p. 68.
378. Muftah, A.A., et al., *Ki67 expression in invasive breast cancer: the use of tissue microarrays compared with whole tissue sections.* Breast Cancer Res Treat, 2017. **164**(2): p. 341-348.

379. Zaha, D.C., *Significance of immunohistochemistry in breast cancer.* World J Clin Oncol, 2014. **5**(3): p. 382-92.
380. Camp, R.L., M. Dolled-Filhart, and D.L. Rimm, *X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization.* Clin Cancer Res, 2004. **10**(21): p. 7252-9.
381. Dolled-Filhart, M., et al., *Quantitative In situ Analysis of β-Catenin Expression in Breast Cancer Shows Decreased Expression Is Associated with Poor Outcome.* Cancer Research, 2006. **66**(10): p. 5487-5494.
382. Gutman, D.A., et al., *The Digital Slide Archive: A Software Platform for Management, Integration, and Analysis of Histology for Cancer Research.* Cancer Res, 2017. **77**(21): p. e75-e78.
383. Makhlouf, S., et al., *Evaluation of tumour infiltrating lymphocytes in luminal breast cancer using artificial intelligence.* British Journal of Cancer, 2023.
384. Narayanan, P.L., et al., *Unmasking the immune microecology of ductal carcinoma in situ with deep learning.* NPJ Breast Cancer, 2021. **7**(1): p. 19.
385. Graham, S., et al., *Hover-Net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images.* Med Image Anal, 2019. **58**: p. 101563.
386. Gamper, J., et al. *PanNuke: An Open Pan-Cancer Histology Dataset for Nuclei Instance Segmentation and Classification*. in *European Congress on Digital Pathology*. 2019.
387. Paul, D., *The systemic hallmarks of cancer.* Journal of Cancer Metastasis and Treatment, 2020. **6**: p. 29.
388. Ayob, A.Z. and T.S. Ramasamy, *Cancer stem cells as key drivers of tumour progression.* Journal of Biomedical Science, 2018. **25**(1): p. 20.
389. Feitelson, M.A., et al., *Sustained proliferation in cancer: Mechanisms and novel therapeutic targets.* Seminars in Cancer Biology, 2015. **35**: p. S25-S54.
390. Ayat Gamal, L., et al., *Expression, assessment and significance of Ki67 expression in breast cancer: an update.* Journal of Clinical Pathology, 2023: p. jcp-2022-208731.
391. Cree, I.A., et al., *Counting mitoses: SI(ze) matters!* Mod Pathol, 2021. **34**(9): p. 1651-1657.
392. Veta, M., et al., *Mitosis Counting in Breast Cancer: Object-Level Interobserver Agreement and Comparison to an Automatic Method.* PLoS One, 2016. **11**(8): p. e0161286.
393. Lashen, A., et al., *Visual assessment of mitotic figures in breast cancer: a comparative study between light microscopy and whole slide images.* Histopathology, 2021. **n/a**(n/a).
394. Zhang, B., S. Kirov, and J. Snoddy, *WebGestalt: an integrated system for exploring gene sets in various biological contexts.* Nucleic Acids Res, 2005. **33**(Web Server issue): p. W741-8.
395. van Bergeijk, S.A., et al., *Deep learning supported mitoses counting on whole slide images: A pilot study for validating breast*

cancer grading in the clinical workflow. J Pathol Inform, 2023. **14**: p. 100316.
396. Ibrahim, A., et al., *Assessment of mitotic activity in breast cancer: revisited in the digital pathology era.* Journal of Clinical Pathology, 2022. **75**(6): p. 365.
397. Pantanowitz, L., et al., *Accuracy and efficiency of an artificial intelligence tool when counting breast mitoses.* Diagnostic Pathology, 2020. **15**(1): p. 80.
398. Wei, B.R., et al., *Agreement in Histological Assessment of Mitotic Activity Between Microscopy and Digital Whole Slide Images Informs Conversion for Clinical Diagnosis.* Acad Pathol, 2019. **6**: p. 2374289519859841.
399. Ibrahim, A., et al., *Defining the area of mitoses counting in invasive breast cancer using whole slide image.* Mod Pathol, 2022. **35**(6): p. 739-748.
400. Donovan, T.A., et al., *Mitotic Figures-Normal, Atypical, and Imposters: A Guide to Identification.* Vet Pathol, 2021. **58**(2): p. 243-257.
401. Saraste, A. and K. Pulkki, *Morphologic and biochemical hallmarks of apoptosis.* Cardiovascular Research, 2000. **45**(3): p. 528-537.
402. Yadav, K.S., et al., *Assessment of interobserver variability in mitotic figure counting in different histological grades of oral squamous cell carcinoma.* J Contemp Dent Pract, 2012. **13**(3): p. 339-44.
403. Lester, S.C., et al., *Protocol for the Examination of Specimens From Patients With Invasive Carcinoma of the Breast.* Archives of Pathology & Laboratory Medicine, 2009. **133**(10): p. 1515-1538.
404. Tan, P.H., et al., *The 2019 World Health Organization classification of tumours of the breast.* Histopathology, 2020. **77**(2): p. 181-185.
405. Al-Janabi, S., A. Huisman, and P.J. Van Diest, *Digital pathology: current status and future perspectives.* Histopathology, 2012. **61**(1): p. 1-9.
406. Chowdhury, N., et al., *Interobserver variation in breast cancer grading: a statistical modeling approach.* Anal Quant Cytol Histol, 2006. **28**(4): p. 213-8.
407. Tandon, A., et al., *Image analysis assisted study of mitotic figures in oral epithelial dysplasia and squamous cell carcinoma using differential stains.* J Oral Biol Craniofac Res, 2016. **6**(Suppl 1): p. S18-s23.
408. Veta, M., et al., *Assessment of algorithms for mitosis detection in breast cancer histopathology images.* Med Image Anal, 2015. **20**(1): p. 237-48.
409. Shima, D.T., et al., *Partitioning of the Golgi apparatus during mitosis in living HeLa cells.* J Cell Biol, 1997. **137**(6): p. 1211-28.
410. Meyer, J.S., et al., *Breast carcinoma malignancy grading by Bloom-Richardson system vs proliferation index: reproducibility of grade and advantages of proliferation index.* Mod Pathol, 2005. **18**(8): p. 1067-78.

411. Wang, Z., *Regulation of Cell Cycle Progression by Growth Factor-Induced Cell Signaling.* Cells, 2021. **10**(12).
412. Ohashi, R., et al., *Prognostic utility of atypical mitoses in patients with breast cancer: A comparative study with Ki67 and phosphohistone H3.* Journal of Surgical Oncology, 2018. **118**(3): p. 557-567.
413. Comings, D.E., *Arrangement of chromatin in the nucleus.* Hum Genet, 1980. **53**(2): p. 131-43.
414. Godek, K.M. and D.A. Compton, *Quantitative methods to measure aneuploidy and chromosomal instability.* Methods Cell Biol, 2018. **144**: p. 15-32.
415. Prayson, R.A., *Cell Proliferation and Tumors of the Central Nervous System Part 1: Evaluation of Mitotic Activity.* Journal of Neuropathology & Experimental Neurology, 2002. **61**(6): p. 501-509.
416. Ganem, N.J. and D. Pellman, *Linking abnormal mitosis to the acquisition of DNA damage.* Journal of Cell Biology, 2012. **199**(6): p. 871-881.
417. Shekhar, M.P., et al., *Rad6 overexpression induces multinucleation, centrosome amplification, abnormal mitosis, aneuploidy, and transformation.* Cancer Res, 2002. **62**(7): p. 2115-24.
418. Zaleski, M., et al., *Mitotic activity in noninvasive papillary urothelial carcinoma: its value in predicting tumor recurrence and comparison with the contemporary 2-tier grading system.* Human Pathology, 2019. **84**: p. 275-282.
419. Matsuda, Y., et al., *Mitotic index and multipolar mitosis in routine histologic sections as prognostic markers of pancreatic cancers: A clinicopathological study.* Pancreatology, 2016. **16**(1): p. 127-132.
420. Habougit, C., et al., *Histopathologic features predict survival in diffuse pleural malignant mesothelioma on pleural biopsies.* Virchows Arch, 2017. **470**(6): p. 639-646.
421. Lashen, A., et al., *Visual assessment of mitotic figures in breast cancer: a comparative study between light microscopy and whole slide images.* Histopathology, 2021.
422. Ciriello, G., et al., *Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer.* Cell, 2015. **163**(2): p. 506-19.
423. Lohse, M., et al., *RobiNA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics.* Nucleic Acids Res, 2012. **40**(Web Server issue): p. W622-7.
424. Robinson, M.D., D.J. McCarthy, and G.K. Smyth, *edgeR: a Bioconductor package for differential expression analysis of digital gene expression data.* Bioinformatics, 2010. **26**(1): p. 139-40.
425. Lin, G., et al., *VennPainter: A Tool for the Comparison and Identification of Candidate Genes Based on Venn Diagrams.* PLoS One, 2016. **11**(4): p. e0154315.

426. Van Leeuwen, A.M., et al., *Atypical mitotic figures and the mitotic index in cervical intraepithelial neoplasia.* Virchows Arch, 1995. **427**(2): p. 139-44.
427. Takubo, K., et al., *Changes of telomere length with aging.* Geriatr Gerontol Int, 2010. **10 Suppl 1**: p. S197-206.
428. Takubo, K., et al., *Q-FISH analysis of telomere and chromosome instability in the oesophagus with and without squamous cell carcinoma in situ.* J Pathol, 2010. **221**(2): p. 201-9.
429. Montgomery, E., et al., *Analysis of anaphase figures in routine histologic sections distinguishes chromosomally unstable from chromosomally stable malignancies.* Cancer Biol Ther, 2003. **2**(3): p. 248-52.
430. Batistatou, A., *Mitoses and cancer.* Medical Hypotheses, 2004. **63**(2): p. 281-282.
431. Hayward, D.G., et al., *The centrosomal kinase Nek2 displays elevated levels of protein expression in human breast cancer.* Cancer Res, 2004. **64**(20): p. 7370-6.
432. Pihan, G.A., et al., *Centrosome defects and genetic instability in malignant tumors.* Cancer Res, 1998. **58**(17): p. 3974-85.
433. van Steenhoven, J.E.C., et al., *Assessment of tumour proliferation by use of the mitotic activity index, and Ki67 and phosphohistone H3 expression, in early-stage luminal breast cancer.* Histopathology, 2020. **77**(4): p. 579-587.
434. Khan, M.S., et al., *A comparison of Ki-67 and mitotic count as prognostic markers for metastatic pancreatic and midgut neuroendocrine neoplasms.* Br J Cancer, 2013. **108**(9): p. 1838-45.
435. Kriegsmann, M. and A. Warth, *What is better/reliable, mitosis counting or Ki67/MIB1 staining?* Transl Lung Cancer Res, 2016. **5**(5): p. 543-546.
436. Del Rosario Taco Sanchez, M., et al., *Digital quantification of KI-67 in breast cancer.* Virchows Arch, 2019. **474**(2): p. 169-176.
437. Morris, A.H. and J.P.A. Ioannidis, *Limitations of medical research and evidence at the patient-clinician encounter scale.* Chest, 2013. **143**(4): p. 1127-1135.
438. Vollmer, S., et al., *Machine learning and artificial intelligence research for patient benefit: 20 critical questions on transparency, replicability, ethics, and effectiveness.* BMJ, 2020. **368**: p. l6927.
439. Stålhammar, G., et al., *Digital image analysis of Ki67 in hot spots is superior to both manual Ki67 and mitotic counts in breast cancer.* Histopathology, 2018. **72**(6): p. 974-989.
440. Ishibashi, N., et al., *Correlation between the Ki-67 proliferation index and response to radiation therapy in small cell lung cancer.* Radiation Oncology, 2017. **12**(1): p. 16.
441. Faviana, P., et al., *Proposal for a New Diagnostic Histopathological Approach in the Evaluation of Ki-67 in GEP-NETs.* Diagnostics (Basel), 2022. **12**(8).
442. Wang, H.J., C.Y. Li, and Y.K. Wang, *Diagnostic implications of Ki-67 expression in adipocytes and lipoblasts: an*

immunohistochemical study. Int J Clin Exp Pathol, 2014. **7**(12): p. 8899-904.
443. Li, L., et al., *Artificial intelligence-assisted interpretation of Ki-67 expression and repeatability in breast cancer.* Diagnostic Pathology, 2022. **17**(1): p. 20.
444. Robertson, S., et al., *Prognostic potential of automated Ki67 evaluation in breast cancer: different hot spot definitions versus true global score.* Breast Cancer Research and Treatment, 2020. **183**(1): p. 161-175.
445. Skjervold, A.H., et al., *Visual and digital assessment of Ki-67 in breast cancer tissue - a comparison of methods.* Diagn Pathol, 2022. **17**(1): p. 45.
446. Maranta, A.F., et al., *Do YOU know the Ki-67 index of your breast cancer patients? Knowledge of your institution's Ki-67 index distribution and its robustness is essential for decision-making in early breast cancer.* Breast, 2020. **51**: p. 120-126.
447. Skjervold, A.H., et al., *Visual and digital assessment of Ki-67 in breast cancer tissue - a comparison of methods.* Diagnostic Pathology, 2022. **17**(1): p. 45.
448. Ades, F., et al., *Luminal B Breast Cancer: Molecular Characterization, Clinical Management, and Future Perspectives.* Journal of Clinical Oncology, 2014. **32**(25): p. 2794-2803.
449. Jang, M.H., et al., *A comparison of Ki-67 counting methods in luminal Breast Cancer: The Average Method vs. the Hot Spot Method.* PLoS One, 2017. **12**(2): p. e0172031.
450. McCarty, K.S., Jr., et al., *Estrogen receptor analyses. Correlation of biochemical and immunohistochemical methods using monoclonal antireceptor antibodies.* Arch Pathol Lab Med, 1985. **109**(8): p. 716-21.
451. Mudduwa, L.K., *Quick score of hormone receptor status of breast carcinoma: correlation with the other clinicopathological prognostic parameters.* Indian J Pathol Microbiol, 2009. **52**(2): p. 159-63.
452. Tobin, N.P., et al., *Multi-level gene expression signatures, but not binary, outperform Ki67 for the long term prognostication of breast cancer patients.* Molecular Oncology, 2014. **8**(3): p. 741-752.
453. Varga, Z., et al., *Ki-67 assessment in early breast cancer: SAKK28/12 validation study on the IBCSG VIII and IBCSG IX cohort.* Scientific Reports, 2019. **9**(1): p. 13534.
454. Bustreo, S., et al., *Optimal Ki67 cut-off for luminal breast cancer prognostic evaluation: a large case series study with a long-term follow-up.* Breast Cancer Res Treat, 2016. **157**(2): p. 363-371.
455. Lester, S.C., et al., *Protocol for the examination of specimens from patients with invasive carcinoma of the breast.* Arch Pathol Lab Med, 2009. **133**(10): p. 1515-38.
456. Hoadley, K.A., et al., *Multiplatform analysis of 12 cancer types reveals molecular classification within and across tissues of origin.* Cell, 2014. **158**(4): p. 929-944.

457. Wang, K., et al., *MapSplice: Accurate mapping of RNA-seq reads for splice junction discovery.* Nucleic Acids Research, 2010. **38**(18): p. e178-e178.
458. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.* Genome Biol, 2014. **15**(12): p. 550.
459. Song, T.H., et al., *Simultaneous Cell Detection and Classification in Bone Marrow Histology Images.* IEEE J Biomed Health Inform, 2019. **23**(4): p. 1469-1476.
460. Lu, W., et al., *AI-based intra-tumor heterogeneity score of Ki67 expression as a prognostic marker for early-stage ER+/HER2− breast cancer.* The Journal of Pathology: Clinical Research, 2023. **n/a**(n/a).
461. Awadelkarim, K., et al., *Ki-67 Labeling Index in Primary Invasive Breast Cancer from Sudanese Patients: A Pilot Study.* ISRN Pathology, 2012. **2012**: p. 6 pages.
462. Basile, M.L., F.S. Kuga, and F. Del Carlo Bernardi, *Comparation of the quantification of the proliferative index KI67 between eyeball and semi-automated digital analysis in gastro-intestinal neuroendrocrine tumors.* Surgical and Experimental Pathology, 2019. **2**(1): p. 21.
463. Hida, A.I., et al., *Visual assessment of Ki67 using a 5-grade scale (Eye-5) is easy and practical to classify breast cancer subtypes with high reproducibility.* J Clin Pathol, 2015. **68**(5): p. 356-61.
464. Fulawka, L. and A. Halon, *Ki-67 evaluation in breast cancer: The daily diagnostic practice.* Indian Journal of Pathology and Microbiology, 2017. **60**(2): p. 177-184.
465. Mu, K., et al., *A standardized method for quantifying proliferation by Ki-67 and cyclin A immunohistochemistry in breast cancer.* Ann Diagn Pathol, 2015. **19**(4): p. 243-8.
466. Dzulkifli, F.A., M.Y. Mashor, and H. Jaafar. *An overview of recent counting methods for Ki67 IHC staining*. 2019.
467. Arima, N., et al., *A Comparison of the Hot Spot and the Average Cancer Cell Counting Methods and the Optimal Cutoff Point of the Ki-67 Index for Luminal Type Breast Cancer.* Oncology, 2015. **90**(1): p. 43-50.
468. Arima, N., et al., *A Comparison of the Hot Spot and the Average Cancer Cell Counting Methods and the Optimal Cutoff Point of the Ki-67 Index for Luminal Type Breast Cancer.* Oncology, 2016. **90**(1): p. 43-50.
469. Ahmed, S.T., et al., *Proliferative Index (Ki67) for Prediction in Breast Duct Carcinomas.* Asian Pac J Cancer Prev, 2018. **19**(4): p. 955-959.
470. Ankle, M., et al., *Comparison of staining of mitotic figures by haematoxylin and eosin-and crystal violet stains, in oral epithelial dysplasia and squamous cell carcinoma.* Indian Journal of Dental Research, 2007. **18**(3): p. 101-105.

471. Adams, D.L., et al., *Mitosis in circulating tumor cells stratifies highly aggressive breast carcinomas.* Breast Cancer Res, 2016. **18**(1): p. 44.
472. Weidner, N., D.H. Moore, 2nd, and R. Vartanian, *Correlation of Ki-67 antigen expression with mitotic figure index and tumor grade in breast carcinomas using the novel "paraffin"-reactive MIB1 antibody.* Hum Pathol, 1994. **25**(4): p. 337-42.
473. Trihia, H., et al., *Ki-67 expression in breast carcinoma.* Cancer, 2003. **97**(5): p. 1321-1331.
474. Herbert, K.J., et al., *Targeting TOPK sensitises tumour cells to radiation-induced damage by enhancing replication stress.* Cell Death & Differentiation, 2020.
475. Lee, L.H., H. Yang, and G. Bigras, *Current breast cancer proliferative markers correlate variably based on decoupled duration of cell cycle phases.* Sci Rep, 2014. **4**: p. 5122.
476. Rossi, L., et al., *Prognostic impact of discrepant Ki67 and mitotic index on hormone receptor-positive, HER2-negative breast carcinoma.* Br J Cancer, 2015. **113**(7): p. 996-1002.
477. Untch, M., et al., *13th st. Gallen international breast cancer conference 2013: primary therapy of early breast cancer evidence, controversies, consensus - opinion of a german team of experts (zurich 2013).* Breast Care (Basel), 2013. **8**(3): p. 221-9.
478. Noske, A., et al., *Risk stratification in luminal-type breast cancer: Comparison of Ki-67 with EndoPredict test results.* The Breast, 2020. **49**: p. 101-107.
479. Gluz, O., et al., *West German Study Group Phase III PlanB Trial: First Prospective Outcome Data for the 21-Gene Recurrence Score Assay and Concordance of Prognostic Markers by Central and Local Pathology Assessment.* J Clin Oncol, 2016. **34**(20): p. 2341-9.
480. Copur, M.S., et al., *The relationship of oncotype Dx Recurrence Score (RS) with Ki 67 in early stage breast cancer patients in a community based cancer center in rural central Nebraska.* Journal of Clinical Oncology, 2022. **40**(16_suppl): p. e12566-e12566.
481. Sahebjam, S., et al., *Ki 67 is a major, but not the sole determinant of Oncotype Dx recurrence score.* British Journal of Cancer, 2011. **105**(9): p. 1342-1345.
482. Lee, J., et al., *Ki-67, 21-Gene Recurrence Score, Endocrine Resistance, and Survival in Patients With Breast Cancer.* JAMA Network Open, 2023. **6**(8): p. e2330961-e2330961.
483. Garutti, M., et al., *Definition of High-Risk Early Hormone-Positive HER2-Negative Breast Cancer: A Consensus Review.* Cancers (Basel), 2022. **14**(8).
484. Sittenfeld, S.M.C., et al., *A multi-institutional prediction model to estimate the risk of recurrence and mortality after mastectomy for T1-2N1 breast cancer.* Cancer, 2022. **128**(16): p. 3057-3066.
485. Diana, A., et al., *Prognostic Relevance of Progesterone Receptor Levels in Early Luminal-Like HER2 Negative Breast Cancer*

Subtypes: A Retrospective Analysis. Front Oncol, 2022. **12**: p. 813462.
486. Järvinen, T.A., et al., *Estrogen receptor beta is coexpressed with ERalpha and PR and associated with nodal status, grade, and proliferation rate in breast cancer.* Am J Pathol, 2000. **156**(1): p. 29-35.
487. Allison, K.H., et al., *Estrogen and Progesterone Receptor Testing in Breast Cancer: ASCO/CAP Guideline Update.* Journal of Clinical Oncology, 2020. **38**(12): p. 1346-1366.
488. Gellersen, B., M.S. Fernandes, and J.J. Brosens, *Non-genomic progesterone actions in female reproduction.* Hum Reprod Update, 2009. **15**(1): p. 119-38.
489. Prat, A., et al., *Prognostic significance of progesterone receptor-positive tumor cells within immunohistochemically defined luminal A breast cancer.* J Clin Oncol, 2013. **31**(2): p. 203-9.
490. Anderson, E., *The role of oestrogen and progesterone receptors in human mammary development and tumorigenesis.* Breast Cancer Res, 2002. **4**(5): p. 197-201.
491. Rozeboom, B., N. Dey, and P. De, *ER+ metastatic breast cancer: past, present, and a prescription for an apoptosis-targeted future.* Am J Cancer Res, 2019. **9**(12): p. 2821-2831.
492. Van Asten, K., et al., *Prognostic Value of the Progesterone Receptor by Subtype in Patients with Estrogen Receptor-Positive, HER-2 Negative Breast Cancer.* Oncologist, 2019. **24**(2): p. 165-171.
493. Grassadonia, A., et al., *Long-term outcome of breast cancer patients with pathologic N3a lymph node stage.* Breast, 2017. **32**: p. 79-86.
494. Bardou, V.-J., et al., *Progesterone Receptor Status Significantly Improves Outcome Prediction Over Estrogen Receptor Status Alone for Adjuvant Endocrine Therapy in Two Large Breast Cancer Databases.* Journal of Clinical Oncology, 2003. **21**(10): p. 1973-1979.
495. Purdie, C.A., et al., *Progesterone receptor expression is an independent prognostic variable in early breast cancer: a population-based study.* British Journal of Cancer, 2014. **110**(3): p. 565-572.
496. Campbell, E.J., et al., *The combined endocrine receptor in breast cancer, a novel approach to traditional hormone receptor interpretation and a better discriminator of outcome than ER and PR alone.* Br J Cancer, 2016. **115**(8): p. 967-973.
497. Davies, C., et al., *Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials.* Lancet, 2011. **378**(9793): p. 771-84.
498. Hefti, M.M., et al., *Estrogen receptor negative/progesterone receptor positive breast cancer is not a reproducible subtype.* Breast Cancer Res, 2013. **15**(4): p. R68.

499. Allison, K.H., et al., *Estrogen and Progesterone Receptor Testing in Breast Cancer: ASCO/CAP Guideline Update.* J Clin Oncol, 2020. **38**(12): p. 1346-1366.
500. Ogawa, Y., et al., *Immunohistochemical assessment for estrogen receptor and progesterone receptor status in breast cancer: analysis for a cut-off point as the predictor for endocrine therapy.* Breast Cancer, 2004. **11**(3): p. 267-75.
501. Bravaccini, S., et al., *The impact of progesterone receptor expression on prognosis of patients with rapidly proliferating, hormone receptor-positive early breast cancer: a post hoc analysis of the IBIS 3 trial.* Ther Adv Med Oncol, 2020. **12**: p. 1758835919888999.
502. Asogan, A.B., G.S. Hong, and S.K. Arni Prabhakaran, *Concordance between core needle biopsy and surgical specimen for oestrogen receptor, progesterone receptor and human epidermal growth factor receptor 2 status in breast cancer.* Singapore Med J, 2017. **58**(3): p. 145-149.
503. Arnedos, M., et al., *Discordance between core needle biopsy (CNB) and excisional biopsy (EB) for estrogen receptor (ER), progesterone receptor (PgR) and HER2 status in early breast cancer (EBC).* Ann Oncol, 2009. **20**(12): p. 1948-52.
504. Ahmad, Z., et al., *Artificial intelligence (AI) in medicine, current applications and future role with special emphasis on its potential and promise in pathology: present and future impact, obstacles including costs and acceptance among pathologists, practical and philosophical considerations. A comprehensive review.* Diagnostic Pathology, 2021. **16**(1): p. 24.
505. Försch, S., et al., *Artificial Intelligence in Pathology.* Dtsch Arztebl Int, 2021. **118**(12): p. 194-204.
506. Vijayashree, R., P. Aruthra, and K.R. Rao, *A comparison of manual and automated methods of quantitation of oestrogen/progesterone receptor expression in breast carcinoma.* J Clin Diagn Res, 2015. **9**(3): p. Ec01-5.
507. Litjens, G., et al., *A survey on deep learning in medical image analysis.* Med Image Anal, 2017. **42**: p. 60-88.
508. Muftah, A.A., et al., *Further evidence to support bimodality of oestrogen receptor expression in breast cancer.* Histopathology, 2017. **70**(3): p. 456-465.
509. Andrahennadi, S., et al., *Current Landscape of Targeted Therapy in Hormone Receptor-Positive and HER2-Negative Breast Cancer.* Curr Oncol, 2021. **28**(3): p. 1803-1822.
510. Sehdev, S., et al., *Safety of adjuvant endocrine therapies in hormone receptor-positive early breast cancer.* Curr Oncol, 2009. **16 Suppl 2**(Suppl 2): p. S14-23.
511. Allred, D.C., *Issues and updates: evaluating estrogen receptor-α, progesterone receptor, and HER2 in breast cancer.* Modern Pathology, 2010. **23**(2): p. S52-S59.
512. Payne, S.J., et al., *Predictive markers in breast cancer--the present.* Histopathology, 2008. **52**(1): p. 82-90.

513. Goldhirsch, A., et al., *Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013.* Ann Oncol, 2013. **24**(9): p. 2206-23.
514. Leen, S., et al., *Study assessing the quality of quantification of estrogen receptor protein expression by immunohistochemistry and gene expression in breast cancer.* Patholog Res Int, 2014. **2014**: p. 372653.
515. Nadji, M., et al., *Immunohistochemistry of estrogen and progesterone receptors reconsidered: experience with 5,993 breast cancers.* Am J Clin Pathol, 2005. **123**(1): p. 21-7.
516. Viale, G., et al., *Chemoendocrine Compared With Endocrine Adjuvant Therapies for Node-Negative Breast Cancer: Predictive Value of Centrally Reviewed Expression of Estrogen and Progesterone Receptors—International Breast Cancer Study Group.* Journal of Clinical Oncology, 2008. **26**(9): p. 1404-1410.
517. Onitilo, A.A., et al., *Breast cancer subtypes based on ER/PR and Her2 expression: comparison of clinicopathologic features and survival.* Clin Med Res, 2009. **7**(1-2): p. 4-13.
518. Dunbier, A.K., et al., *Relationship between plasma estradiol levels and estrogen-responsive gene expression in estrogen receptor-positive breast cancer in postmenopausal women.* J Clin Oncol, 2010. **28**(7): p. 1161-7.
519. Rani, A., et al., *Endocrine Resistance in Hormone Receptor Positive Breast Cancer–From Mechanism to Therapy.* Frontiers in Endocrinology, 2019. **10**(245).
520. Thakkar, J.P. and D.G. Mehta, *A review of an unfavorable subset of breast cancer: estrogen receptor positive progesterone receptor negative.* Oncologist, 2011. **16**(3): p. 276-85.
521. Cui, X., et al., *Biology of progesterone receptor loss in breast cancer and its implications for endocrine therapy.* J Clin Oncol, 2005. **23**(30): p. 7721-35.
522. Mook, S., et al., *Calibration and discriminatory accuracy of prognosis calculation for breast cancer with the online Adjuvant! program: a hospital-based retrospective cohort study.* Lancet Oncol, 2009. **10**(11): p. 1070-6.
523. Liu, S., et al., *Progesterone receptor is a significant factor associated with clinical outcomes and effect of adjuvant tamoxifen therapy in breast cancer patients.* Breast Cancer Res Treat, 2010. **119**(1): p. 53-61.
524. Diana, A., et al., *Prognostic Relevance of Progesterone Receptor Levels in Early Luminal-Like HER2 Negative Breast Cancer Subtypes: A Retrospective Analysis.* Frontiers in Oncology, 2022. **12**.
525. Freelander, A., et al., *Molecular Biomarkers for Contemporary Therapies in Hormone Receptor-Positive Breast Cancer.* Genes (Basel), 2021. **12**(2).
526. Sutepvarnon, A., M. Warnnissorn, and V. Srimuninnimit, *Predictive value of Ki67 for adjuvant chemotherapy in node-*

negative, hormone receptor-positive breast cancer. J Med Assoc Thai, 2013. **96 Suppl 2**: p. S60-6.
527. Criscitiello, C., et al., *High Ki-67 score is indicative of a greater benefit from adjuvant chemotherapy when added to endocrine therapy in luminal B HER2 negative and node-positive breast cancer.* Breast, 2014. **23**(1): p. 69-75.
528. Colozza, M., D. Larsimont, and M.J. Piccart, *Progesterone Receptor Testing: Not the Right Time to Be Buried.* Journal of Clinical Oncology, 2005. **23**(16): p. 3867-3868.
529. Morgan, D.A.L., N.A. Refalo, and K.L. Cheung, *Strength of ER-positivity in relation to survival in ER-positive breast cancer treated by adjuvant tamoxifen as sole systemic therapy.* The Breast, 2011. **20**(3): p. 215-219.
530. Zilenaite, D., et al., *Independent Prognostic Value of Intratumoral Heterogeneity and Immune Response Features by Automated Digital Immunohistochemistry Analysis in Early Hormone Receptor-Positive Breast Carcinoma.* Front Oncol, 2020. **10**: p. 950.
531. Torhorst, J., et al., *Tissue microarrays for rapid linking of molecular changes to clinical endpoints.* Am J Pathol, 2001. **159**(6): p. 2249-56.
532. Abele, N., et al., *Noninferiority of Artificial Intelligence–Assisted Analysis of Ki-67 and Estrogen/Progesterone Receptor in Breast Cancer Routine Diagnostics.* Modern Pathology, 2023. **36**(3): p. 100033.
533. Łukasiewicz, S., et al., *Breast Cancer-Epidemiology, Risk Factors, Classification, Prognostic Markers, and Current Treatment Strategies-An Updated Review.* Cancers (Basel), 2021. **13**(17).
534. Afzal, S., et al., *Breast Cancer; Discovery of Novel Diagnostic Biomarkers, Drug Resistance, and Therapeutic Implications.* Front Mol Biosci, 2022. **9**: p. 783450.
535. Neves Rebello Alves, L., et al. *Biomarkers in Breast Cancer: An Old Story with a New End*. Genes, 2023. **14**, DOI: 10.3390/genes14071364.
536. Thenrajan, T., S. Alwarappan, and J. Wilson, *Molecular Diagnosis and Cancer Prognosis-A Concise Review.* Diagnostics (Basel), 2023. **13**(4).
537. Krzyszczyk, P., et al., *The growing role of precision and personalized medicine for cancer treatment.* Technology (Singap World Sci), 2018. **6**(3-4): p. 79-100.
538. Pal, M., et al., *Current advances in prognostic and diagnostic biomarkers for solid cancers: Detection techniques and future challenges.* Biomedicine & Pharmacotherapy, 2022. **146**: p. 112488.
539. Izquierdo-Bouldstridge, A., et al., *Histone H1 depletion triggers an interferon response in cancer cells via activation of heterochromatic repeats.* Nucleic Acids Res, 2017. **45**(20): p. 11622-11642.

540. Espinoza Pereira, K.N., et al., *Histone mutations in cancer.* Biochemical Society Transactions, 2023. **51**(5): p. 1749-1763.
541. Fyodorov, D.V., et al., *Emerging roles of linker histones in regulating chromatin structure and function.* Nat Rev Mol Cell Biol, 2018. **19**(3): p. 192-206.
542. Lai, S., et al., *Molecular and Cellular Functions of the Linker Histone H1.2.* Front Cell Dev Biol, 2021. **9**: p. 773195.
543. Hyun, S.Y., H.I. Hwang, and Y.J. Jang, *Polo-like kinase-1 in DNA damage response.* BMB Rep, 2014. **47**(5): p. 249-55.
544. Liu, Z., Q. Sun, and X. Wang, *PLK1, A Potential Target for Cancer Therapy.* Transl Oncol, 2017. **10**(1): p. 22-32.
545. Al-Qahtani, Q.H., et al., *Kinome inhibition reveals a role for polo-like kinase 1 in targeting post-transcriptional control in cancer.* Mol Oncol, 2021. **15**(8): p. 2120-2139.
546. Kandala, S., et al., *Chronic chromosome instability induced by Plk1 results in immune suppression in breast cancer.* Cell Reports, 2023. **42**(12): p. 113266.
547. Hu, K., et al., *Small interfering RNA library screen identified polo-like kinase-1 (PLK1) as a potential therapeutic target for breast cancer that uniquely eliminates tumor-initiating cells.* Breast Cancer Res, 2012. **14**(1): p. R22.
548. Yu, S., et al., *MCMs in Cancer: Prognostic Potential and Mechanisms.* Anal Cell Pathol (Amst), 2020. **2020**: p. 3750294.
549. Saito, Y., et al., *MCMBP promotes the assembly of the MCM2–7 hetero-hexamer to ensure robust DNA replication in human cells.* eLife, 2022. **11**: p. e77393.
550. Riera, A., et al., *From structure to mechanism-understanding initiation of DNA replication.* Genes Dev, 2017. **31**(11): p. 1073-1088.
551. Kaur, G., et al., *Minichromosome Maintenance Complex (MCM) Genes Profiling and MCM2 Protein Expression in Cervical Cancer Development.* Asian Pac J Cancer Prev, 2019. **20**(10): p. 3043-3049.
552. Kwok, H.F., et al., *Prognostic significance of minichromosome maintenance proteins in breast cancer.* Am J Cancer Res, 2015. **5**(1): p. 52-71.
553. Yousef, E.M., et al., *MCM2: An alternative to Ki-67 for measuring breast cancer cell proliferation.* Modern Pathology, 2017. **30**(5): p. 682-697.
554. Toyokawa, G., et al., *Minichromosome Maintenance Protein 7 is a potential therapeutic target in human cancer and a novel prognostic marker of non-small cell lung cancer.* Molecular Cancer, 2011. **10**(1): p. 65.
555. Zhang, D., et al., *Reliability of Tissue Microarrays in Detecting Protein Expression and Gene Amplification in Breast Cancer.* Modern Pathology, 2003. **16**(1): p. 79-85.
556. Győrffy, B., et al., *An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using*

microarray data of 1,809 patients. Breast Cancer Res Treat, 2010. **123**(3): p. 725-31.
557. Ma, H., et al., *Quantitative measures of estrogen receptor expression in relation to breast cancer-specific mortality risk among white women and black women.* Breast Cancer Res, 2013. **15**(5): p. R90.
558. Lashen, A., et al., *Ki67 assessment in invasive luminal breast cancer: a comparative study between different scoring methods.* Histopathology, 2022. **81**(6): p. 786-798.
559. Hacking, S.M., E. Yakirevich, and Y. Wang, *From Immunohistochemistry to New Digital Ecosystems: A State-of-the-Art Biomarker Review for Precision Breast Cancer Medicine.* Cancers (Basel), 2022. **14**(14).
560. Neves Rebello Alves, L., et al., *Biomarkers in Breast Cancer: An Old Story with a New End.* Genes (Basel), 2023. **14**(7).
561. Zhang, Y. and Z. Zhang, *The history and advances in cancer immunotherapy: understanding the characteristics of tumor-infiltrating immune cells and their therapeutic implications.* Cell Mol Immunol, 2020. **17**(8): p. 807-821.
562. Audia, J.E. and R.M. Campbell, *Histone Modifications and Cancer.* Cold Spring Harb Perspect Biol, 2016. **8**(4): p. a019521.
563. Hergeth, S.P. and R. Schneider, *The H1 linker histones: multifunctional proteins beyond the nucleosomal core particle.* EMBO Rep, 2015. **16**(11): p. 1439-53.
564. Millán-Ariño, L., et al., *Mapping of six somatic linker histone H1 variants in human breast cancer cells uncovers specific features of H1.2.* Nucleic Acids Res, 2014. **42**(7): p. 4474-93.
565. Yang, S.M., et al., *H1 linker histone promotes epigenetic silencing by regulating both DNA methylation and histone H3 methylation.* Proc Natl Acad Sci U S A, 2013. **110**(5): p. 1708-13.
566. Kowalski, A. and J. Pałyga, *Linker histone subtypes and their allelic variants.* Cell Biol Int, 2012. **36**(11): p. 981-96.
567. Kostova, N.N., et al., *Immunohistochemical demonstration of histone H1(0) in human breast carcinoma.* Histochem Cell Biol, 2005. **124**(5): p. 435-43.
568. Brockers, K. and R. Schneider, *Histone H1, the forgotten histone.* Epigenomics, 2019. **11**(4): p. 363-366.
569. Khachaturov, V., et al., *Histone H1.5, a novel prostatic cancer marker: an immunohistochemical study.* Human Pathology, 2014. **45**(10): p. 2115-2119.
570. Telu, K.H., et al., *Alterations of histone H1 phosphorylation during bladder carcinogenesis.* J Proteome Res, 2013. **12**(7): p. 3317-26.
571. Xie, W., et al., *Expression and potential prognostic value of histone family gene signature in breast cancer.* Exp Ther Med, 2019. **18**(6): p. 4893-4903.
572. Skvortsova, K., C. Stirzaker, and P. Taberlay, *The DNA methylation landscape in cancer.* Essays Biochem, 2019. **63**(6): p. 797-811.

573. Lee, Y.Y., et al., *CSF2 Overexpression Is Associated with STAT5 Phosphorylation and Poor Prognosis in Patients with Urothelial Carcinoma.* J Cancer, 2016. **7**(6): p. 711-21.
574. Sancho, M., et al., *Depletion of human histone H1 variants uncovers specific roles in gene expression and cell growth.* PLoS Genet, 2008. **4**(10): p. e1000227.
575. Zlatanova, J. and D. Doenecke, *Histone H1 zero: a major player in cell differentiation?* Faseb j, 1994. **8**(15): p. 1260-8.
576. Liao, R., et al., *HIST1H1B Promotes Basal-Like Breast Cancer Progression by Modulating CSF2 Expression.* Front Oncol, 2021. **11**: p. 780094.
577. Zolota, V., et al., *Epigenetic Alterations in Triple-Negative Breast Cancer-The Critical Role of Extracellular Matrix.* Cancers (Basel), 2021. **13**(4).
578. Lee, S.Y., C. Jang, and K.A. Lee, *Polo-like kinases (plks), a key regulator of cell cycle and new potential target for cancer therapy.* Dev Reprod, 2014. **18**(1): p. 65-71.
579. de Cárcer, G., *The Mitotic Cancer Target Polo-Like Kinase 1: Oncogene or Tumor Suppressor?* Genes (Basel), 2019. **10**(3).
580. de Cárcer, G., et al., *Plk1 overexpression induces chromosomal instability and suppresses tumor development.* Nat Commun, 2018. **9**(1): p. 3012.
581. de Cárcer, G., et al., *Plk1 overexpression induces chromosomal instability and suppresses tumor development.* Nature Communications, 2018. **9**(1): p. 3012.
582. Su, S., et al., *PLK1 inhibition-based combination therapies for cancer management.* Translational Oncology, 2022. **16**: p. 101332.
583. Bruinsma, W., et al., *Spatial Separation of Plk1 Phosphorylation and Activity.* Front Oncol, 2015. **5**: p. 132.
584. Ruf, S., et al., *PLK1 (polo like kinase 1) inhibits MTOR complex 1 and promotes autophagy.* Autophagy, 2017. **13**(3): p. 486-505.
585. Ueda, A., et al., *Therapeutic potential of PLK1 inhibition in triple-negative breast cancer.* Laboratory Investigation, 2019. **99**(9): p. 1275-1286.
586. Maire, V., et al., *Polo-like kinase 1: a potential therapeutic option in combination with conventional chemotherapy for the management of patients with triple-negative breast cancer.* Cancer Res, 2013. **73**(2): p. 813-23.
587. Patel, J.R., et al., *A Novel Allosteric Inhibitor Targets PLK1 in Triple Negative Breast Cancer Cells.* Biomolecules, 2022. **12**(4).
588. Zhao, S., et al., *Deciphering the performance of polo-like kinase 1 in triple-negative breast cancer progression according to the centromere protein U-phosphorylation pathway.* Am J Cancer Res, 2021. **11**(5): p. 2142-2158.
589. Ren, Y., et al., *Bioinformatics analysis of key genes in triple negative breast cancer and validation of oncogene PLK1.* Ann Transl Med, 2020. **8**(24): p. 1637.
590. Cebrián, A., et al., *Decreased PLK1 expression denotes therapy resistance and unfavourable disease-free survival in rectal cancer*

patients receiving neoadjuvant chemoradiotherapy. Pathol Res Pract, 2016. **212**(12): p. 1133-1137.
591. Barr, F.A., H.H. Silljé, and E.A. Nigg, *Polo-like kinases and the orchestration of cell division.* Nat Rev Mol Cell Biol, 2004. **5**(6): p. 429-40.
592. Wierer, M., et al., *PLK1 signaling in breast cancer cells cooperates with estrogen receptor-dependent gene transcription.* Cell Rep, 2013. **3**(6): p. 2021-32.
593. Giordano, A., et al., *Polo-like kinase 1 (Plk1) inhibition synergizes with taxanes in triple negative breast cancer.* PLoS One, 2019. **14**(11): p. e0224420.
594. Ha, G.-H., et al., *Combination Treatment of Polo-Like Kinase 1 and Tankyrase-1 Inhibitors Enhances Anticancer Effect in Triple-negative Breast Cancer Cells.* Anticancer Research, 2018. **38**(3): p. 1303.
595. Zhao, M., et al., *miR-107 regulates the effect of MCM7 on the proliferation and apoptosis of colorectal cancer via the PAK2 pathway.* Biochemical Pharmacology, 2021. **190**: p. 114610.
596. Kim, S.H., et al., *Upregulated expression of BCL2, MCM7, and CCNE1 indicate cisplatin-resistance in the set of two human bladder cancer cell lines: T24 cisplatin sensitive and T24R2 cisplatin resistant bladder cancer cell lines.* Investig Clin Urol, 2016. **57**(1): p. 63-72.
597. Suzuki, Y., et al., *Changes in MCM2–7 proteins at senescence.* Genes & Genetic Systems, 2019. **94**(3): p. 123-132.
598. Ota, T., et al., *Minichromosome maintenance protein 7 as a potential prognostic factor for progression-free survival in high-grade serous carcinomas of the ovary.* Mod Pathol, 2011. **24**(2): p. 277-87.
599. Tamura, T., et al., *Minichromosome maintenance-7 and geminin are reliable prognostic markers in patients with oral squamous cell carcinoma: immunohistochemical study.* J Oral Pathol Med, 2010. **39**(4): p. 328-34.
600. Nowinska, K., et al., *Correlation between levels of expression of minichromosome maintenance proteins, Ki-67 proliferation antigen and metallothionein I/II in laryngeal squamous cell cancer.* Int J Oncol, 2016. **48**(2): p. 635-45.
601. Wang, H., et al., *MCM7 expression is correlated with histological subtypes of lung adenocarcinoma and predictive of poor prognosis.* Int J Clin Exp Pathol, 2017. **10**(12): p. 11747-11753.
602. Fristrup, N., et al., *Multicenter Validation of Cyclin D1, MCM7, TRIM29, and UBE2C as Prognostic Protein Markers in Non-Muscle–Invasive Bladder Cancer.* The American Journal of Pathology, 2013. **182**(2): p. 339-349.
603. Liu, X., et al., *The Alterations and Potential Roles of MCMs in Breast Cancer.* J Oncol, 2021. **2021**: p. 7928937.
604. Asghari, A., et al., *A novel group of genes that cause endocrine resistance in breast cancer identified by dynamic gene expression analysis.* Oncotarget, 2022. **13**: p. 600-613.

605. Liang, Z., et al., *Simvastatin suppresses the DNA replication licensing factor MCM7 and inhibits the growth of tamoxifen-resistant breast cancer cells.* Sci Rep, 2017. **7**: p. 41776.
606. Hagenaars, S.C., et al., *Standardization of the tumor-stroma ratio scoring method for breast cancer research.* Breast Cancer Res Treat, 2022. **193**(3): p. 545-553.
607. Wong, A.N.N., et al., *Current Developments of Artificial Intelligence in Digital Pathology and Its Future Clinical Applications in Gastrointestinal Cancers.* Cancers (Basel), 2022. **14**(15).
608. Shafi, S. and A.V. Parwani, *Artificial intelligence in diagnostic pathology.* Diagn Pathol, 2023. **18**(1): p. 109.
609. Eliyatkın, N., et al., *Molecular Classification of Breast Carcinoma: From Traditional, Old-Fashioned Way to A New Age, and A New Way.* J Breast Health, 2015. **11**(2): p. 59-66.
610. Testa, U., G. Castelli, and E. Pelosi, *Breast Cancer: A Molecularly Heterogenous Disease Needing Subtype-Specific Treatments.* Med Sci (Basel), 2020. **8**(1).
611. Bissell, M.J., H.G. Hall, and G. Parry, *How does the extracellular matrix direct gene expression?* J Theor Biol, 1982. **99**(1): p. 31-68.
612. Galea, M.H., et al., *The Nottingham Prognostic Index in primary breast cancer.* Breast Cancer Res Treat, 1992. **22**(3): p. 207-19.
613. Atallah, N.M., et al., *Deciphering the Morphology of Tumor-Stromal Features in Invasive Breast Cancer Using Artificial Intelligence.* Mod Pathol, 2023. **36**(10): p. 100254.
614. Makhlouf, S., et al., *Evaluation of tumour infiltrating lymphocytes in luminal breast cancer using artificial intelligence.* Br J Cancer, 2023.
615. Ehteshami Bejnordi, B., et al., *Using deep convolutional neural networks to identify and classify tumor-associated stroma in diagnostic breast biopsies.* Mod Pathol, 2018. **31**(10): p. 1502-1512.
616. Vo, V.T. and S.-H. Kim *Mulvernet: Nucleus Segmentation and Classification of Pathology Images Using the HoVer-Net and Multiple Filter Units*. Electronics, 2023. **12**, DOI: 10.3390/electronics12020355.
617. Zhao, T., et al., *RGSB-UNet: Hybrid Deep Learning Framework for Tumour Segmentation in Digital Pathology Images.* Bioengineering (Basel), 2023. **10**(8).
618. Wang, C. and H. Zhao *Spatial Heterogeneity Analysis: Introducing a New Form of Spatial Entropy*. Entropy, 2018. **20**, DOI: 10.3390/e20060398.
619. van der Walt, S., et al., *scikit-image: image processing in Python.* PeerJ, 2014. **2**: p. e453.
620. Davnall, F., et al., *Assessment of tumor heterogeneity: an emerging imaging tool for clinical practice?* Insights Imaging, 2012. **3**(6): p. 573-89.

621. Azimi, V., et al., *BREAST CANCER HISTOPATHOLOGY IMAGE ANALYSIS PIPELINE FOR TUMOR PURITY ESTIMATION.* Proc IEEE Int Symp Biomed Imaging, 2017. **2017**: p. 1137-1140.
622. Lawson, D.J., et al., *CLARITY: comparing heterogeneous data using dissimilarity.* R Soc Open Sci, 2021. **8**(12): p. 202182.
623. Higgins, J.P., et al., *Measuring inconsistency in meta-analyses.* Bmj, 2003. **327**(7414): p. 557-60.
624. Nunes, A., T. Trappenberg, and M. Alda, *The definition and measurement of heterogeneity.* Transl Psychiatry, 2020. **10**(1): p. 299.
625. Lüönd, F., S. Tiede, and G. Christofori, *Breast cancer as an example of tumour heterogeneity and tumour cell plasticity during malignant progression.* British Journal of Cancer, 2021. **125**(2): p. 164-175.
626. Marusyk, A., M. Janiszewska, and K. Polyak, *Intratumor Heterogeneity: The Rosetta Stone of Therapy Resistance.* Cancer Cell, 2020. **37**(4): p. 471-484.
627. Luque-Bolivar, A., et al., *Resistance and Overcoming Resistance in Breast Cancer.* Breast Cancer (Dove Med Press), 2020. **12**: p. 211-229.
628. Fisher, R., L. Pusztai, and C. Swanton, *Cancer heterogeneity: implications for targeted therapeutics.* Br J Cancer, 2013. **108**(3): p. 479-85.
629. Sarker, I.H., *AI-Based Modeling: Techniques, Applications and Research Issues Towards Automation, Intelligent and Smart Systems.* SN Computer Science, 2022. **3**(2): p. 158.
630. Laurinavicius, A., et al., *Bimodality of intratumor Ki67 expression is an independent prognostic factor of overall survival in patients with invasive breast carcinoma.* Virchows Arch, 2016. **468**(4): p. 493-502.
631. Heindl, A., S. Nawaz, and Y. Yuan, *Mapping spatial heterogeneity in the tumor microenvironment: a new era for digital pathology.* Laboratory Investigation, 2015. **95**(4): p. 377-384.
632. Greaves, M. and C.C. Maley, *Clonal evolution in cancer.* Nature, 2012. **481**(7381): p. 306-13.
633. Archetti, M. and K.J. Pienta, *Cooperation among cancer cells: applying game theory to cancer.* Nat Rev Cancer, 2019. **19**(2): p. 110-117.
634. He, L., et al., *Histology image analysis for carcinoma detection and grading.* Comput Methods Programs Biomed, 2012. **107**(3): p. 538-56.
635. Ramón, Y.C.S., et al., *Clinical implications of intratumor heterogeneity: challenges and opportunities.* J Mol Med (Berl), 2020. **98**(2): p. 161-177.
636. Ma, D., et al., *Clinical and molecular relevance of mutant-allele tumor heterogeneity in breast cancer.* Breast Cancer Research and Treatment, 2017. **162**(1): p. 39-48.
637. Turashvili, G. and E. Brogi, *Tumor Heterogeneity in Breast Cancer.* Frontiers in Medicine, 2017. **4**.

638. Sung, H., et al., *Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries.* CA Cancer J Clin, 2021. **71**(3): p. 209-249.
639. Shen, K., et al., *Impact of adjuvant chemotherapy on T1N0M0 breast cancer patients: a propensity score matching study based on SEER database and external cohort.* BMC Cancer, 2022. **22**(1): p. 863.
640. Diest, P.J.v., E.v.d. Wall, and J.P.A. Baak, *Prognostic value of proliferation in invasive breast cancer: a review.* Journal of Clinical Pathology, 2004. **57**(7): p. 675.
641. Gamble, P., et al., *Determining breast cancer biomarker status and associated morphological features using deep learning.* Commun Med (Lond), 2021. **1**: p. 14.
642. Miller, I., et al., *Ki67 is a Graded Rather than a Binary Marker of Proliferation versus Quiescence.* Cell Reports, 2018. **24**(5): p. 1105-1112.e5.
643. Davidson, T.M., et al., *Breast Cancer Prognostic Factors in the Digital Era: Comparison of Nottingham Grade using Whole Slide Images and Glass Slides.* J Pathol Inform, 2019. **10**: p. 11.
644. Polley, M.Y., et al., *An international Ki67 reproducibility study.* J Natl Cancer Inst, 2013. **105**(24): p. 1897-906.
645. Gam, L.H., *Breast cancer and protein biomarkers.* World J Exp Med, 2012. **2**(5): p. 86-91.
646. Jeong, S.B., et al., *Essential Role of Polo-like Kinase 1 (Plk1) Oncogene in Tumor Growth and Metastasis of Tamoxifen-Resistant Breast Cancer.* Molecular Cancer Therapeutics, 2018. **17**(4): p. 825-837.
647. Gutschner, T. and S. Diederichs, *The hallmarks of cancer: a long non-coding RNA point of view.* RNA Biol, 2012. **9**(6): p. 703-19.
648. Rivenbark, A.G., S.M. O'Connor, and W.B. Coleman, *Molecular and cellular heterogeneity in breast cancer: challenges for personalized medicine.* Am J Pathol, 2013. **183**(4): p. 1113-1124.
649. Sauerbrei, W., et al., *Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK): An Abridged Explanation and Elaboration.* J Natl Cancer Inst, 2018. **110**(8): p. 803-811.
650. Dennis, J., et al., *The type and quality of paraffin wax is important when constructing tissue microarrays.* Mol Pathol, 2003. **56**(5): p. 306.
651. Johnson, A., *The effect of five investing techniques on air bubble entrapment and casting nodules.* Int J Prosthodont, 1992. **5**(5): p. 424-33.

Appendix

1- Contribution of by collaborators on the study:

a- Bioinformatics analysis for atypical mitosis study and Ki67 study was performed by Prof Nigel Monga

b- Deep learning models in Ki67, PR and heterogeneity chapters that were developed by Dr Wahab Noorul, Dr Wenqi lu (computer scientists, Warwwivk University)

c- Double scoring for the immunohistochemical markers was performed by Dr Michael Toss and Dr Nabeela El-Malki

2- Contribution of the author to histological review, data collection and TMA construction

a- I reviewed 1070 cases histological between December 2020 to July 2022

b- I collected the clinical data for 540 patients.

c- I collected tumour blocs for 1050 luminal cancer patients and did matching the blocks with their corresponding the WSI then I annotated these digital slide for TMA construction by the technical team.

2- Steps of TMA construction

Preparation of the donor blocks

Sampling the representative site from the donor block is very important for proper construction tissue array. To identify suitable cases of BC that could be used to build the TMA, slides that were previously reviewed were selected and scanned. FFPE tumour block (donor) that was characteristic of the specimens from each patient was obtained from the tissue block archives. Those blocks that were ≥ 4 mm in depth bearing viable tumour tissue with an adequate burden of tumour and priority was given to blocks containing adequate tumour tissue that had not been frequently sectioned for previous studies.

Matching of donor blocks with their corresponding slides

To confirm the presence of tumour and to identify which areas would be most appropriate for the TMA sampling, the author, and other pathologists (Dr Shorouk Makhlouf, Dr Nehal Atalla, Dr Amera Sheha, Dr Suzan Ghanem) reviewed the digital slides and compared it with the retrieved blocks. Visual assessment of blocks against the slides was done to obtain matched blocks with slides. To avoid the potential risk of exhausting a patient's diagnostic material, tumours with one block or small tumour samples were excluded from the study to ensure that the diagnostic patient's materials were not depleted.

Digital images marking

TMA cores were marked using Panoramic Viewer Software v. 1.15.4 (3D HISTECH®, Budapest, Hungary). TMA cores were selected based on the higher density of invasive tumour. 0.6 mm tissue core diameter was annotated, and 4 to 12 cores were marked for each image (**Figure 1).**

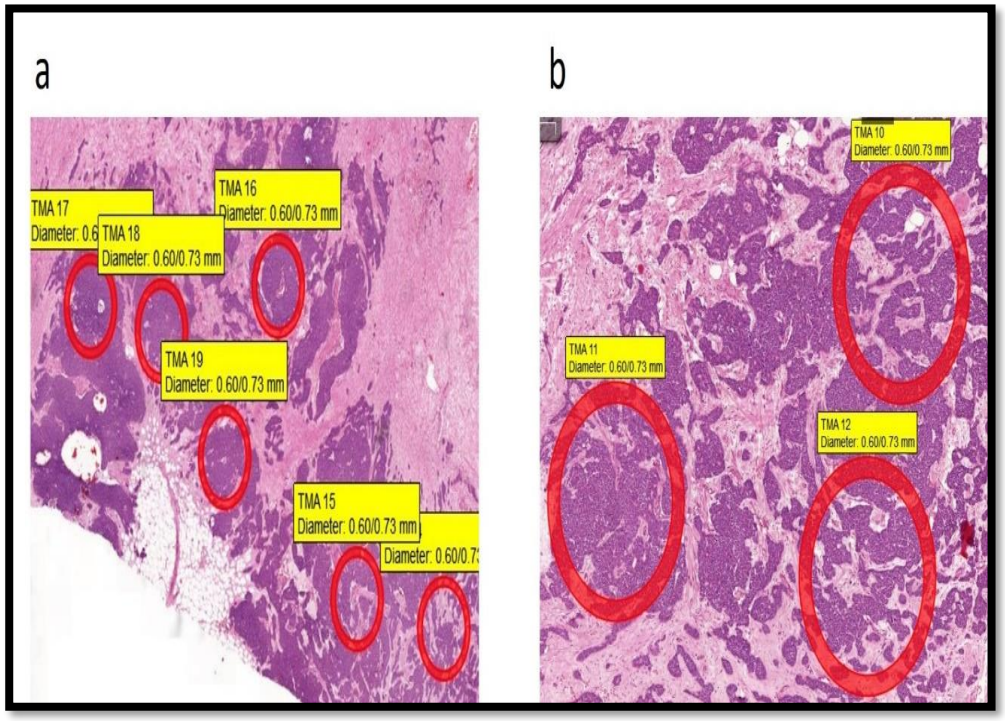


Figure 1*:* Annotation of scanned images for TMA construction

a- low power view shows 6 TMA cores, b- high power view of tumour cores showed sufficient tumour within the selected cores.

Recipient block construction

TMA building was achieved using ordinary paraffin wax-based blocks that had been prepared by pouring paraffin into blocks (5–10 mm deep) and then extruding the tissue array cores into the block. To prepare recipient blocks, plain molten paraffin wax (56°–58°C) was poured into a mould.

This was then topped with a histopathological tissue cassette [650].Then the blocks were cooled on ice until the wax had hardened and the wax was checked for presence of air bubbles, striations or fault line [651]. The blocks were also checked to ensure that they were covered completely but did not have excess wax at the edges or back. Then, the block was trimmed smooth to ensure a flat surface using a microtome.

TMA layout design

TMA Grand Master 2.4-UG-EN software (3D HISTECH®, Budapest, Hungary) was used to design the block layouts. Creation of a new layout followed by insertion of a new layout name was performed. Block size was inserted as a standard of 24x35 mm. Four replications were prepared for each block, and each was given a unique version number, which was added to the block.

Each array contained 120 samples arranged in three groups (subarrays), with a 1mm space between each core. The first group contained the 53 cores (6 rows, 10 columns) from the upper half of the block. The second group of the block contained 40 cores (4 rows, 10 columns), while the third group of the block was made up of 23 cores (3 rows, 10 columns). Cores in each subarray were arranged at a centred distance, and the subarrays were separated by a single gap for orientation and macroscopic identification of the TMA block. Positive internal controls were introduced by adding orientation cores from normal kidney and

tonsil tissues outside of the geometric margin of the array on the top and bottom left-hand sides.

Documentation of both the donor block and the TMA position of each individual core was done with great care. An array map grid was produced for each array block. Each grid contained a unique identifier and was arranged in the same sequence corresponding to its position in the array block.

Coring process

Each time, the automated TMA arrayer was loaded with 60 donor blocks and two recipient blocks (only one recipient block was loaded during construction cases with only one available core). Donor blocks were given pseudo-anonymised labels, and recipient blocks were accordingly labelled PS1, PS2, PS3, PS4, etc. Matching and overlying of the scanned and marked images with their related donor blocks was performed using the software before the cores from the donor blocks were automatically sampled and embedded into recipient blocks. Figure 2 shows steps of TMA construction.

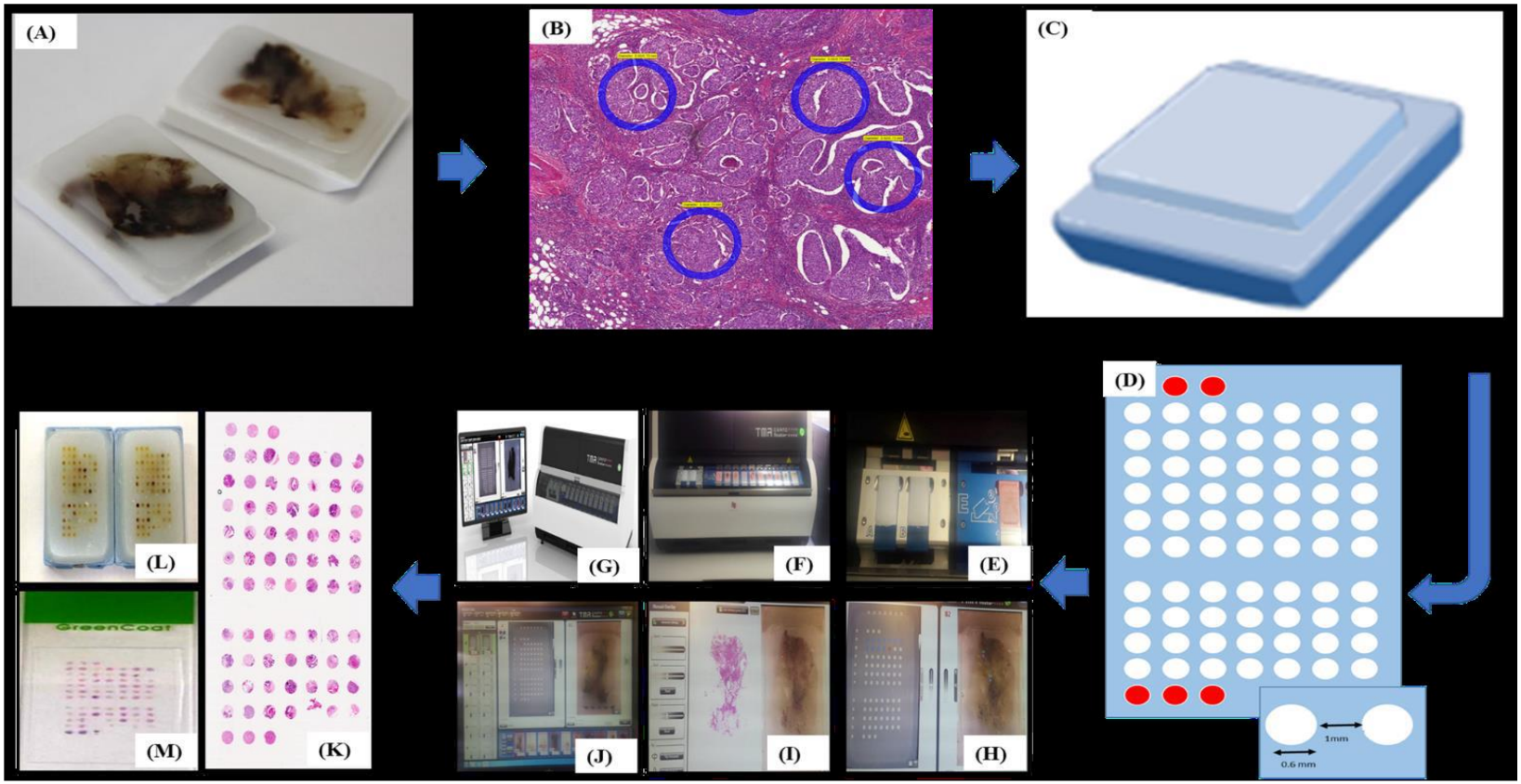


Figure 2 : Steps of TMA construction.

(A) Selection of tumour block, (B) Selection of 0.6 mm cores, (C) Tissue cassettes block, (D) Layout of recipient block, E) TMA machine and screen, F) Loading machine with donor and recipient blocks, G) Loading recipient blocks in TMA machine, H) Images of donor block mapped with selected layout , I) Matching the scanned images with the relevant donor block, J) Marked suitable sites for coring in donor block, K) sections are scanned, L) Constructed TMA blocks, and M) H&E-stained sections.

TMA block annealing and sectioning

After the TMA blocks were constructed, they were placed face-up in an oven for 3 h at 37°C to anneal them to help the tissue cores adhere to the walls of the holes in the array block and makes the wax flexible for handling. In addition, it maximises the number of sections containing all the samples in the array. A clean glass microscope slide was placed on the face surface of the block, then gentle pressure was applied to push all the cores on the array to the same level. The blocks were left at room temperature for 3 hours to cool. The annealing cycle was replicated, and the blocks were stored overnight at 4°C before sectioning.

Using a standard microtome sectioning technique, 4-µm sections were cut from TMA blocks and transferred into permanent positively charged glass slides (Surgipath, X-tra™ Adhesive, Peterborough, UK). To ensure the quality of the histological samples and the TMA construction, the first section of each block was stained with H&E and scanned histologically.

Table 1: Definitions and description of measures of intra-tumoural heterogeneity in breast cancer.

	Name of Heterogeneity measure	Short Name	Definition	Nature and Range	Mean ±SD
Grade heterogeneity	Atmospheric scattering model of grade	G (ASM)	ASM from the cooccurrence matrix of the grade predictions for patches of size 123 micro m.	Binary (0—1)	0.55±0.32
	Grade contrast	G contrast	Contrast from the cooccurrence matrix of the grade predictions for patches of size 123 micro m.	Continuous (0 and 2.9)	0.85±0.73
	Grade correlation	G correlation	Correlation from the cooccurrence matrix of the grade predictions for patches of size 123 micro m.	Binary (-1—1)	0.14±0.36
	Grade dissimilarity	G dissimilarity	Dissimilarity from the cooccurrence matrix of the grade predictions for patches of size 123 micro m.	Binary (0 and 1)	0.47±0.39
Pleomorphism heterogeneity	Grade energy	G energy	Energy from the cooccurrence matrix of the grade predictions for patches of size 123 micro m	Binary (0 and 1)	0.7±0.2

	Name of Heterogeneity measure	Short Name	Definition	Nature and Range	Mean ±SD
			staining can be computed.		
	Atmospheric scattering model of pleomorphism	pASM	ASM from the cooccurrence matrix of the pleomorphic predictions for patches of size 123 micro m.	Binary (0 and 1)	0.43±0.22
	Pleomorphism contrast	P contrast	Contrast from the cooccurrence matrix of the pleomorphic predictions for patches of size 123 micro m.	Continuous, (0—1.5)	0.40±0.21
	Pleomorphism Energy	P energy	Energy from the cooccurrence matrix of the pleomorphic predictions for patches of size 123 micro m.	Binary (0 and 1)	0.63±0.17