

UNIVERSITY OF CAPE TOWN

*THE DEPARTMENT OF ANATOMICAL PATHOLOGY, FACULTY OF HEALTH
SCIENCES,
UNIVERSITY OF CAPE TOWN*

**THE ROLE OF TISSUE INFLAMMATION IN KENYAN WOMEN WITH
BREAST CANCER**

By
SHAHIN SAYED ABDULREHMAN
(MMED Path, FCPath-ECSA, FRCPath.)

*A thesis submitted in partial fulfilment of the requirements for the degree of PhD in
Anatomical Pathology*

Cape Town/South Africa

December 2021

©copyright

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

DEDICATION

I dedicate this thesis to my late parents, Sayed Mohamed, and Mariam Sayed, for teaching me the value of hard work, patience, and perseverance.

“ACKNOWLEDGEMENT”

First, I am grateful to my supervisors Prof Dhirendra Govender and Prof Richard Naidoo, whose scholarly advice, help and constant encouragement have contributed significantly to the completion of this study.

I also wish to thank the management, staff, faculty members, and my fellow students at the Department of Anatomical Pathology, University of Cape Town for their invaluable input and for being a great source of support to me during my study.

My sincere thanks to my Departmental Chair at Aga Khan University, Prof. Zahir Moloo, for his invaluable advice and support during my study.

I would like to acknowledge the support and guidance received from my collaborators at the Department of Pathology at Baylor College, the Department of Digital and Computational Pathology, Columbia University, New York and the Division of Cancer Epidemiology and Genetics, NCI, USA.

My sincere thanks to the laboratory technical team both at University of Cape Town (Subash Govender and Raymond Kriel) and Aga Khan University Hospital (Johnstone Ngau and Beth Nduta) for their invaluable assistance.

My appreciation to Dr. Jasmit Shah-Aga Khan University, Sharon Fan- Division of Cancer Epidemiology and Genetics, NCI, and Hela Koka-National Cancer Institute for their statistical support.

I wish to appreciate Angela Mutuku, Aga Khan University and Thelma Galsworthy, University of Cape Town for their administrative and logistical support.

Sincere thanks to Dr. Pumza Magangane, University of Witwatersrand, South Africa, who assisted in the proofreading and editing of my paper and to Gaylord Mwangi and Veronica Ngundo, Aga Khan University Hospital, who assisted with formatting of this document.

My sincere gratitude to Viviane Oluoch for her invaluable assistance with abstraction of patient clinical and outcome data.

I thank the library staff with special appreciation to Nasra Gathoni as well as Irene Kimani, Human Resource department, Wanjiru Njagi, Housing and Travel department and Eunice Wamaitha, Maryann Kariuki, Simon Wanyoike and Thomas Ombech from the Finance Department at Aga Khan University for their timely support.

Finally, my sincere gratitude to my family for their unwavering support, encouragement, and sacrifice throughout my studies.

This study was supported through the Deans Fund, Aga University, East Africa

Thank you all.

DECLARATION

I declare this thesis does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any university and that to the best of my knowledge it does not contain any material previously published or written by another person except where due reference has been made in the text.

Signed by candidate

(Signature of candidate)

_____ 8th December 2021 _____
Date

LIST OF ABBREVIATIONS

Abbreviation	Description
5NP	5 Negative Profile African
AA	American
AiA	Autoimmune Arthritis
AKU	Aga Khan University
AMBER	African American Breast Cancer Epidemiology and Risk
APC	Antigen-presenting cell
ASR	Age-standardized rates
BC	Breast cancer
BMI	Body mass Index Breast Cancer
BRCA	gene
CI	Confidence interval
CK	Cytokeratin
CSC	Cancer stem cells
CTLA	Cytotoxic T-lymphocyte associated antigen
DCIS	Ductal carcinoma in-situ Distant disease-free survival
DDFS	
DNA	Deoxyribonucleic acid
EGFR	Epidermal growth factor receptor
ER	Oestrogen receptor Food and Drug Administration
FDA	
FFTP	First full-term pregnancy
FISH	Fluorescence in-situ hybridization Glaxo Smith Kline
GSK	
HDI	Human development index
HER2	Human epidermal growth factor receptor 2 Hormone receptor
HR	
HRT	Hormone replacement therapy
IBC	Invasive breast cancer
IDC	Invasive ductal carcinoma
IHC	Immunohistochemistry
IL	Interleukin Interquartile range
IQR	
LCIS	Lobular carcinoma in-situ Leukocyte telomere lengths
LTL	
LVI	Lymphovascular invasion

	Oral
OC	contraceptives
OR	Odds ratio
PBCR	Population-based cancer registry
PBCS	Polish Breast Cancer Study
PD	Programmed cell death
	Provincial General
PGH	Hospital
PR	Progesterone receptor
Q1	Quartile 1
Q2	Quartile 2
Q3	Quartile 3
Q4	Quartile 4
RR	Risk ratio
RS	Relative survival
	Statistical analysis
SAS	software
	Standard
SD	deviation
	Sub-Saharan
SSA	Africa
TDLU	Terminal duct lobular unit
TN	Triple Negative
	Triple Negative Breast
TNBC	Cancer

LIST OF PUBLICATIONS AND CONFERENCE PRESENTATIONS

I confirm that I have been granted permission by the University of Cape Town Doctoral Degrees board to include the following publications in my thesis and my co-authors have agreed that I may include the publications.

- Tumour Infiltrating Lymphocytes and Outcomes in Breast Cancer Subtypes-accepted for poster presentation at San Antonio Breast Cancer Symposium, Dec-7th -10th 2021. Poster number #1049
- Sayed S, Fan S, Moloo Z, Wasike R, Bird P, Saleh M, Shaikh AJ, Figueroa JD, Naidoo R, Makokha FW, Gardner K, Oigara R, Njoroge FW, Magangane P, Mutebi M, Chauhan R, Mwanzi S, Govender D, Yang XR. Breast cancer risk factors in relation to molecular subtypes in breast cancer patients from Kenya. *Breast Cancer Res.* 2021 Jun 26;23(1):68. doi: 10.1186/s13058-021-01446-3. PMID: 34174935; PMCID: PMC8235821.
- Protein Biomarker Profiling of Estrogen Receptor Positive and Triple Negative Breast Carcinomas from Kenya Using Matrix-Assisted Laser Desorption/ionization Imaging Mass Spectrometry (MALDI-IMS) –submitted for publication to *Journal of Molecular Histology*
- Tumour Infiltrating Lymphocytes (TILs) and immune composition in breast cancer patients from Kenya: Spatial distributions and associations with risk factors and tumour characteristics. submitted to *Cancers*

ABSTRACT

Background

The immune landscape of breast cancer (BC) molecular subtypes from Sub Saharan Africa is understudied and findings are mainly extrapolated from studies in Caucasians. The objective of this study was to describe the cellular and signalling milieu of BC tumour microenvironment and its associations with BC risk factors, tumour characteristics and molecular subtypes in Kenyan women with breast cancer.

Methods:

Molecular subtyping of 838 cases of BC was performed based on Immunohistochemistry (Figure 1). Risk factor data and tumour characteristics were retrieved from an existing database. Visual quantification of overall Tumour Infiltrating Lymphocytes (TILs) was performed on Haematoxylin and Eosin-stained whole slide sections of 226 BC cases based on the guidelines of the International TIL working group. Tissue Microarrays (TMAs) of tumour were constructed, stained with immunohistochemistry for CD3, CD4, CD8, CD68, CD20 and Fox P3, scanned and analysed on the Leica Aperio platform. Cytokine profiles of ER positive and Triple Negative BC were performed using microarray gene arrays and results validated with the ELISA platform. Multivariable polytomous logistic regression models were used to determine associations between BC risk factors and tumour molecular subtypes. Linear and logistic regression models were used to assess the associations between risk factors and tumour features with TILs and TIL subtypes.

Results:

Distribution of molecular subtypes for luminal A, luminal B, HER2-enriched, and triple negative (TN) breast cancers was 34.8%, 35.8%, 10.7%, and 18.6%, respectively. Higher TILs were associated with high KI67, high grade and HER2 status, luminal B subtype, and smaller tumour size. The TILs were predominantly composed of CD3, CD8 and CD68 with a much smaller contribution of CD4 and CD20 and FOxP3. Differential expression of inflammatory related cytokines (GM-CSF, IL8, TGF β 1 and MDC), at both the gene and the protein expression level were observed between Triple negative and ER positive BC.

Conclusions:

Our findings with regards to enrichment of TILs in more aggressive breast cancers are like what has been previously published, however the differential expression of inflammatory cytokines in breast cancer subtypes some of which has also been described in Caucasian and Asian populations warrants further studies as potential targets for therapeutic approaches and

biomarkers of diagnosis and prognosis.

Table of Contents

LIST OF ABBREVIATIONS.....	v
LIST OF PUBLICATIONS AND CONFERENCE PRESENTATIONS.....	vii
ABSTRACT	viii
CHAPTER 1.....	1
1.0 INTRODUCTION.....	1
CHAPTER 2.....	5
2.0 MOLECULAR SUBTYPES OF BREAST CANCER IN KENYAN WOMEN AND RISK FACTOR ASSOCIATION	5
2.1 Background	5
LITERATURE REVIEW	6
2.2 Breast Cancer	6
2.2.1 Epidemiology.....	6
2.2.2 Race and Breast cancer.....	7
2.2.3 Breast Cancer Risk Factors	8
2.2.4 Pathology	18
2.2.5 Traditional subtyping of IBC	18
2.2.6 Molecular classification of IBC	19
2.3 MATERIALS AND METHODS	23
2.3.1 Objective: Categorize breast cancers into molecular subtypes and determine association with breast cancer risk factors	23
2.3.1.1 <i>Sample Source</i>	23
2.3.1.2 <i>Molecular Subtypes</i>	23
2.3.2 STATISTICAL ANALYSIS.....	25
2.4 RESULTS	26
2.4.1 Distribution of breast cancer risk factors (sociodemographic and reproductive characteristics) in overall study population and by hospitals (Table 4).....	26
2.4.2 Distributions of tumour subtypes and pathologic characteristics in overall study population and by hospitals (Table 4)	33
2.4.3 Association between breast cancer risk factors and receptor status ER, PR, and HER2 (Table 6).	39
2.4.4 Associations between breast cancer risk factors and molecular subtypes (Table 7)	45
2.5. DISCUSSION	52
Chapter 3.....	58
3.0: TUMOUR INFILTRATING LYMPHOCYTES AND OUTCOMES IN BREAST CANCER SUBTYPES.....	58
3.1 Background	58
3.1.1 Normal Wound Healing	58
3.1.2 Wound healing as an example of physiological inflammation	58
3.1.3 Role of Inflammation in Tumours	58
3.1.3.1 <i>Inflammation and Breast Cancer</i>	60

3.1.4 Tumour Microenvironment	61
3.1.5 Tumour Infiltrating Lymphocytes in Breast Cancer	65
3.1.5.1 <i>Categories of TILs</i>	66
3.1.5.2 <i>Other Immune cells in breast cancer Tumour Microenvironment</i>	68
3.1.5.2.1 Macrophages	68
3.1.5.2.2 T Regulatory (Treg) cells	69
3.1.5.2.3 Myeloid-derived suppressor cells (MDSCs)	69
3.1.5.2.4 Th17 cells	69
3.1.5.2.5 Mast Cells	69
3.1.5.2.6 Dendritic cells	70
3.1.5.3 <i>Methods used to evaluate TILs</i>	70
3.1.6 Prognosis and predictive value of TILs in breast cancer	71
3.1.6.1 <i>Relationship between TIL concentration and prognosis</i>	71
3.1.6.2 <i>Relationship between TIL and breast cancer subtype</i>	72
3.1.6.3 <i>Racial differences in TILs in breast cancer</i>	74
3.2 OBJECTIVES	76
3.3 METHODS	76
3.3.1 Case Selection	76
3.3.2 Visual TIL Quantification	76
3.3.3 Automated TIL quantification	77
3.3.3.1 <i>Digital pathology analysis</i>	77
3.3.4 Subtyping of TILs	78
3.3.4.1 <i>Case selection for TMA for IHC Markers (CD3, CD4, CD8, CD20, CD68 and FoxP3)</i>	78
3.3.4.2 <i>Construction and reading of TMA</i>	78
3.3.4.3 <i>Quality Control</i>	78
3.4 Data Analysis	79
3.4.1 Interpretation of IHC stained TMA sections for inflammatory cell subtypes	79
3.4.2 Survival data	79
3.4.3 Statistical Analysis	80
3.5 RESULTS	80
3.5.1 Descriptive analysis of sociodemographic and reproductive characteristics	80
3.5.2 Distributions of tumour subtypes and pathologic characteristics	82
3.5.3 Correlation between visual (pathologists) and automated (Halo) TILs	82
3.5.4 In multivariable-adjusted linear regression models (10% significance level highlighted in Table 8)	83
3.5.5 In multivariable-adjusted logistic regression models (10% significance level highlighted in Table 9)	83
3.5.6 Associations between breast cancer risk factors, TILs, and immune cell subtypes	108
3.5.7 Associations between tumour characteristics, molecular subtypes and TILs and immune cell subtypes	108
3.5.8 Associations between molecular subtypes and TILs and outcome data	111

3.6 DISCUSSION	123
Chapter 4.....	126
4.0 CYTOKINES	126
4.1 INTRODUCTION.....	126
4.2 TYPES OF CYTOKINES	131
4.2.1 Interleukin 1 (IL-1)	133
4.2.2 Interleukin 2 (IL-2)	133
4.2.3 Interleukin 4 (IL-4), ...	133
4.2.4 Interleukin 6 (IL-6)	133
4.2.5 Interleukin 8 (IL-8)	134
4.2.6 Interleukin 10 (IL-10).....	134
4.2.7 Tumour Necrosis Factor α (TNF- α).....	134
4.2.8 Transforming Growth Factor- β (TGF- β).....	134
4.3 CYTOKINES IN OESTROGEN RECEPTOR POSITIVE AND TRIPLE NEGATIVE BREAST CANCER.	135
4.4 MEASUREMENT OF CYTOKINES	136
4.4.2 Flow cytometry:	137
4.4.2 Enzyme linked immunosorbent assay (ELISA):.....	137
4.4.3 Bead based Immunoassays:.....	137
4.4.4 Chip/Slide based protein arrays:.....	137
4.4.5 Rolling circle amplification:.....	137
4.4.6 Antibody Based Flow Assays.....	138
4.4.6.1	138
4.4.6.2	138
4.4.7 Aptamer Assays:	138
4.4.8 Mass spectrometry:	138
4.4.9 Electrochemical Methods:	138
4.5 OBJECTIVES	138
4.5.1 Strategy of Approach	139
4.6 METHODS	140
4.6.1 Objective 1:	140
4.6.2 Selection of Cases	140
4.6.3 Preparation of Cases	140
4.6.4 Extraction of RNA.....	141
4.6.5 RNA Integrity testing	141
4.6.5 Final selection of cases for downstream conversion of total RNA to cDNA:	141
4.7 FIRST STRAND cDNA SYNTHESIS FROM TOTAL RNA	143
4.7.1 Preparation of Genomic DNA elimination mix.....	143
4.7.2 Preparation of Reverse Transcription (RT) mix for cDNA Synthesis	143
4.7.3 Pre-amplification of cDNA Target templates.....	143
4.7.4 Application of Pre-Amplified cDNA to RT-PCR Profiler microarray plates	144
4.7.5 Data quality Control (QC).....	145

4.8 DATA MANAGEMENT AND ANALYSIS.....	145
4.9 OBJECTIVE 2.....	145
4.10 MATERIALS AND METHODS	146
4.10.1 Protein extraction from FFPE tissues:	146
4.10.3 Samples	147
4.10.4 Analysis Details	147
4.10.5 Cytokine Validation Experiment using the Multi-Analyte ELISA array	152
4.10.6 Data Analysis	152
4.11 RESULTS	152
4.11.1 Results Objective 1:	152
4.11.2 Results Objective 2:	159
4.12 DISCUSSION	162
5.0 CONCLUSION	165
6.0 RECOMMENDATIONS FOR FUTURE WORK.....	167
REFERENCES.....	168
APPENDICES	186
APPENDIX I: Hospital Facilities in study.	186
APPENDIX II: Associations of key risk factors with oestrogen-receptor negative (ER-) tumours, by hospitals (N=834)	187
APPENDIX III: TIL H &E-Stained Images (Leading Edge, Intratumoural)	193
APPENDIX IV: Tissue Array Map Template	200
APPENDIX V: Aperio-Image Scope Digital Analysis Process	202
APPENDIX VI: Correlation Graphs of Intratumoral and Leading-Edge Stroma	209
APPENDIX VII- Quartile Decoding for individual TILs (figure in Excel Sheet)	213
APPENDIX VIII: RNA Extraction Protocol.....	214
APPENDIX IX: RNA concentration and purity	216
APPENDIX X: Samples that passed RNA integrity test.....	218
APPENDIX XI: Components of the RT2 Profiler PCR array	219
APPENDIX XII: Quantification of RNA.....	222
APPENDIX XIII: Sample size calculations for Validation experiment	224
APPENDIX XIV: Multi-Analyte ELISArray Kits.....	225
APPENDIX XV: Cytokine Elisa Raw Data	232

List of Tables

TABLE 1: ADDITIONAL GENES INVOLVED IN INCREASED RISK OF BREAST CANCER (BC)	11
TABLE 2: MAJOR MOLECULAR SUBTYPES OF BREAST CANCER DETERMINED BY GENE EXPRESSION PROFILING	19
TABLE 3: IMMUNOPHENOTYPING TO APPROXIMATE MOLECULAR SUBTYPE USING SIX BIOMARKERS	20
TABLE 4: CHARACTERISTICS OF KENYAN BREAST CANCER PATIENTS, OVERALL AND BY HOSPITALS (N=838)	27
TABLE 5: DISTRIBUTION OF TUMOUR CHARACTERISTICS IN KENYAN BREAST CANCER PATIENTS, OVERALL AND BY FACILITY (N=838)	35
TABLE 6: ASSOCIATIONS BETWEEN BREAST CANCER RISK FACTORS AND RECEPTOR STATUS IN KENYAN BREAST CANCER PATIENTS (N=838)	41
TABLE 7: ASSOCIATIONS BETWEEN BREAST CANCER RISK FACTORS AND TUMOUR MOLECULAR SUBTYPES IN KENYAN BREAST CANCER PATIENTS (N=821*	46
TABLE 8: COMPARISONS OF VISUAL TILS AND HALO TILS BY TUMOUR CHARACTERISTICS	85
TABLE 9: COMPARISON OF VISUAL TILS AND HALO TILS(N=226)	87
TABLE 10: ASSOCIATIONS OF RISK FACTORS WITH PERCENT POSITIVE CELLS FOR EACH IHC MARKER IN 226 KENYAN BREAST CANCER WOMEN	90
TABLE 11: ASSOCIATIONS OF TUMOR CHARACTERISTICS WITH PERCENT POSITIVE NUCLEI FOR EACH IHC MARKER IN 226 KENYAN BREAST CANCER WOMEN	98
TABLE 12: HAZARD RATIO (HR) AND 95% CONFIDENCE INTERVAL (CI) FOR THE ASSOCIATIONS BETWEEN IHC MARKERS AND TIL AND OVERALL SURVIVAL (OS) AND RECURRENCE-FREE SURVIVAL (RFS)	106
TABLE 13: CYTOKINE EFFECTS ON THE BREAST TUMOR MICROENVIRONMENT	129
TABLE 14: CYTOKINES' UTILITY FOR BREAST CANCER PROGNOSIS AND SURVIVAL	131
TABLE 15: SELECTED LIST OF ER POSITIVE CASES	141
TABLE 16: SELECTED LIST OF TRIPLE NEGATIVE CASES	141
TABLE 17: REVERSE TRANSCRIPTION (RT) MIX COMPONENTS	143
TABLE 18: PRE-AMPLIFICATION MIX COMPONENTS	144
TABLE 19: PCR CONDITIONS	144
TABLE 20: PCR COMPONENT MIX	144
TABLE 21: LIGHT CYCLER 480 CONDITIONS	145
TABLE 22: MULTI-ANALYTE CYTOKINE ELISA ARRAY PLATES	146
TABLE 23: CONCENTRATION OF PROTEINS IN THE PROTEIN EXTRACTS	146
TABLE 24: FOLD REGULATION COMPARISON BETWEEN GROUP 1 (TNBC) AND CONTROL GROUP (ER POSITIVE) WITH CORRESPONDING P-VALUE	153
TABLE 25: GENES OVEREXPRESSED IN GROUP 1 (TNBC) VS. CONTROL GROUP (ER POSITIVE)	156
TABLE 26: GENES UNDER EXPRESSED IN GROUP 1(TNBC) VS. CONTROL GROUP (ER POSITIVE)	157
TABLE 27: DIFFERENCES IN MEDIAN VALUES OF CYTOKINES BETWEEN TNBC AND ER POSITIVE BREAST CANCERS	161

LIST OF FIGURES

FIGURE 1: BREAST CANCER	7
FIGURE 2: BREAST TUMOUR SUBTYPES DEFINED BY EXPRESSION OF ER, PR, HER2, CK5/6, AND EGFR IN KENYAN BREAST CANCER PATIENTS (N=838)	24
FIGURE 3: BREAST TUMOR SUBTYPE DEFINITION IN KENYAN BREAST CANCER PATIENTS (N=838)	34
FIGURE 4: TUMOUR MICROENVIRONMENT	63
FIGURE 5: TUMOUR ELIMINATION PHASE	64
FIGURE 6: TUMOUR EQUILIBRIUM PHASE	64
FIGURE 7: TUMOUR ESCAPE PHASE	65
FIGURE 8: THE CELLULAR CROSS-TALK BETWEEN DIFFERENT LEUKOCYTE SUBSETS	68
FIGURE 9: CASE SELECTION PROCESS AND FINAL INCLUSION	81
FIGURE 10: BREAST TUMOR SUBTYPE DEFINITION IN KENYAN BREAST CANCER PATIENTS (N=226)	82
FIGURE 11: VISUAL LEADING EDGE/PERITUMOURAL TIL COUNTS COMPARED WITH HALO TILS	83
FIGURE 12: VISUAL LEADING INTRATUMOURAL STROMAL TIL COUNTS COMPARED WITH HALO TIL	83
FIGURE 13: DISTRIBUTION OF TILS AND TIL SUBTYPES BY BREAST CANCER MOLECULAR SUBTYPES	110
FIGURE 14: IMMUNE CELL SUBTYPES BY BREAST CANCER MOLECULAR SUBTYPES	110
FIGURE 15: OVERALL SURVIVAL AND TOTAL TILS	112
FIGURE 16: OVERALL SURVIVAL AND CD3	113
FIGURE 17: OVERALL SURVIVAL AND CD4	114
FIGURE 18: OVERALL SURVIVAL AND CD8	115
FIGURE 19: OVERALL SURVIVAL AND CD20	116
FIGURE 20: OVERALL SURVIVAL AND CD68	117
FIGURE 21: OVERALL SURVIVAL AND FOX P3	118
FIGURE 22 : RECURRENCE FREE-SURVIVAL RATE AND CD3	119
FIGURE 23: RECURRENCE FREE-SURVIVAL RATE AND CD4	120
FIGURE 24: RECURRENCE FREE-SURVIVAL RATE AND CD8	121
FIGURE 25: RECURRENCE FREE-SURVIVAL RATE AND CD68	122
FIGURE 26: SCHEMATICS OF MUTUAL COMMUNICATION BETWEEN CANCER CELLS, TUMOR ASSOCIATED HOST CELLS AND ECOSYSTEMS(259)	127
FIGURE 27: ROLE OF CYTOKINES IN THE DIFFERENT STAGES OF BREAST CANCER(260)	128
FIGURE 28: TUMOUR AREA	140
FIGURE 29: TUMOUR AREA SUPERIMPOSED ON UNSTAINED SLIDE	140
FIGURE 30: TICS OF THE SYSTEM BLANK AND SUITABILITY MIXTURE INJECTION	149
FIGURE 31: TICS OF THE SAMPLES IN ASCENDING NUMERICAL ORDER WITH SAMPLE 53 AT THE TOP. .	150
FIGURE 32: TICS OF THE SAMPLES IN ASCENDING NUMERICAL ORDER WITH SAMPLE 53 AT THE TOP. .	151
FIGURE 33: SCATTER PLOT NORMALIZED TO DEMONSTRATE FOLD DIFFERENCE BETWEEN GROUP 1 (TNBC) AGAINST CONTROL (ER POSITIVE)	155
FIGURE 34: VOLCANO PLOT	158
FIGURE 35: HEAT MAP	159
FIGURE 36: CYTOKINE EXPRESSION BETWEEN ER POSITIVE AND TRIPLE NEGATIVE BC.... ERROR! BOOKMARK NOT DEFINED.	
FIGURE 37: GENE EXPRESSION LEVELS (TNBC TEST GROUP/ ER POSITIVE CONTROL GROUP) ERROR! BOOKMARK NOT DEFINED.	

CHAPTER 1

1.0 INTRODUCTION

Estimates from the GLOBOCAN 2020 data covering 185 countries and 36 cancers show that with an incidence of 2,26 million female breast cancers accounted for 11.7% of all cancer sites and 6.9% of all new cancer related mortality(1). In Kenya, breast cancer is the most common cancer diagnosed(2) and the GLOBOCAN estimates show that the incidence of BC in 2020 was 16.1% while mortality was 11.5% out of the 26550 new cases diagnosed in females(1). The disparate mortality from BC compared to the incidence of BC for women in Africa may be contributed partly by lack of awareness of the signs and symptoms of BC (3) however this clearly does not explain why larger more aggressive tumours are diagnosed (4) and why BC presents at least a decade earlier compared with Caucasians(5).

In 1863 Rudolf Virchow first proposed the linkage between inflammation and cancer with the concept of cancers being regarded as “wounds that do not heal”, due to the shared activated inflammatory pathways between the two processes(6).

There has also been considerable research interest recently, on the effect of the tumour microenvironment and inflammatory milieu surrounding tumour biology. The role of persistent chronic inflammation and host immunity in solid malignancies that tip the balance from apoptosis to survival are some of the current research directions. Indeed, immunologic subtyping of breast cancers has recently been proposed for stratifying patients and their tumour into prognostic and predictive groups and for establishing the basis for novel immunotherapies(7, 8).

Some of the long standing chronic inflammatory diseases shown to be associated with the development of cancer include ulcerative colitis with colonic carcinoma, hepatitis C with hepatocellular carcinoma, and chronic pancreatitis with pancreatic carcinoma(9). Although BC is traditionally not associated with an underlying chronic inflammatory disease state(10), it has been reported that elevated serum levels of markers associated with low grade chronic inflammatory states; C-reactive protein or β -Amyloid are associated with risk of BC recurrence post therapy even after adjustment for patient age, tumour stage, race or body mass index(11). The presumed role of the cytokines, IL-1 β , IL6 and IL8 in mediating this effect is supported by the reported predisposition to BC in individuals with genetic polymorphisms in these cytokines. The self-renewal of Cancer Stem cells (CSC) by these cytokines may promote tumour growth and metastases(12).

Cytokines primarily interferons (IFNs), interleukins (ILs), tumor necrosis factor super family (TNFs), transforming growth factor super family (TGFs), and colony stimulating factors (CSFs), provide a crucial molecular link between inflammation, tumor promotion, and development. The recent discovery of numerous single nucleotide polymorphisms (SNPs) within cytokine gene sequences and the differences in SNPs may potentially contribute to ethnic disparities in cancer(13). In breast cancer, polymorphisms in cytokine genes specifically in Interleukin 6(IL6), Interleukin 10 (IL10), Interferon- γ and Tumour Necrosis factor α (TNF α) have been associated with high risk of breast cancer, more aggressive phenotype of breast cancer and worse prognosis(13).

As an example, the low expression of 21082 AA genotype in IL10 has been associated with a high risk of breast cancer. Compared to AG heterozygotes and GG homozygotes which showed a decreased risk for breast cancer (OR:0.58 (95% confidence interval, 0.32–1.07) and 0.38 (95% confidence interval, 0.14–0.99), respectively, 48% of breast cancer cases and 33% of controls were AA homozygotes which showed a significant association for risk of breast cancer. The specific AA polymorphism possibly leads to low production of IL10 the function of which is to prevent cancer(14).Regarding IFN- γ , the IFN- γ TA+AA genotypes were estimated to increase breast cancer risk significantly in Turkish breast cancer patients compared to controls with significant differences between the genotype frequencies of the allele and genotype frequencies for the IFN- γ polymorphism ($p = 0.0312$ and $p = 0.001$, respectively)(15).

Although, considerable work has been published on the role of chronic inflammation in cancer and in breast cancer, many questions remain unanswered. The role and subtypes of various immune cells within the tumour microenvironment (16-24)in the initiation and progression of BC, the types of bioactive mediators that promote tumour growth, the cellular and molecular components that are common across all breast cancer molecular subtypes and the pro-tumour and anti-tumour role of the adaptive immune system are some of the areas that still need further research(8).

Women of African descent present with breast cancers a decade earlier than their Caucasian counterparts(5), however, when compared to their Caucasian counterparts they tend to have tumours with poor prognostic factors(25). As an example, in a study from Tanzania, more than 70% of the patients presented at stage III or IV(26); similarly, in studies from Libya and Nigeria, more than half of the patients presented at stage III or IV (27, 28). A recent systematic review and meta-analysis showed that out of a total of 83 eligible studies from 17 countries in SSA that included 26788 women with breast cancer, more than 70% of patients in 59 of these studies presented with late stage disease and this did not vary by region(29). Persistent antigenic exposure (e.g., through infectious agents), and the barrage of antigenic challenge probably results in an immune system that is under continuous stress and at risk of being constitutively activated. This in turn may cause a shift in the balance from apoptosis towards survival.

The subsequent inflammatory/immune response causes cytokine release, and the reactive oxygen and nitrogen species generated by leukocytes recruited to the inflammatory foci associated with antigenic and/or polyclonal stimulation of the inflammatory cells, epigenetic and genetic instability(12), may all thus play an important role in the pathogenesis of breast cancer in women of African descent. Cytokine polymorphisms are now believed to alter these critical molecules that have a dual face in carcinogenesis as, when implicated in the activation of the immune response, these molecules check the cancer development while their persistent inflammatory reaction can envisage the development of malignancy and tumor growth. Furthermore, multiple SNPs within cytokine gene may potentially drive ethnic differences observed in cancer and breast cancer(13).

There is increasing evidence suggesting that chronic inflammation, driven primarily by innate immunity, may contribute to critical pathways in cancer progression. Studies have found that germline variants in innate immune genes are unique in African populations may result from the adaptation to their unique exposures through

natural selection, suggesting that African breast cancer may have distinct immune profiles. In fact, a recent study investigating breast tumor microenvironment from 1315 patients from the Women's Circle of Health Study found that Black breast cancer patients displayed distinct exhausted CD8+ T-cell signatures indicating an immunobiological basis for their more aggressive disease and a rationale for the use of immune checkpoint inhibitors targeting the exhaustion phenotype(30). In addition, some African women present with locally advanced large breast tumours(31). In a subset of these patients, it should be possible to characterize key prognostic factors specifically, in those patients who have large tumours but no evidence of metastatic disease, which would imply that these tumours are interacting with the microenvironment and are kept in check thus resulting in large masses without systemic dissemination.

More recently, the concept of "Immunoscore" as a prognostic marker in different cancers has gained momentum (32). Common immune signatures comprising density of CD3+, CD8+, and CD45RO+ T cells and their spatial distribution in tissues have been established for tumours that share similar outcomes and therapy response(33, 34). Furthermore, in certain cancers these signatures have outperformed pathological parameters as prognostic markers of staging(7).

Increasingly, immunotherapy is rapidly becoming the next frontier in cancer treatment. The development of drugs that target, as an example, the tumour suppressive, apoptotic, and pro-inflammatory signalling pathways of the host immune system rather than the cancer cell itself, have seen a remarkable paradigm shift in the evolution of cancer specific targeted therapy. Sipuleucel-T, an antigen-presenting cell (APC) activation therapy for castrate resistant metastatic prostate cancer, Ipilimumab, an anti-cytotoxic T-lymphocyte associated antigen-4 (CTLA-4) monoclonal antibody therapy for advanced melanoma, Pembrolizumab, another anti-Programmed cell death Protein 1 (anti-PD-1) antibody for patients with advanced melanoma, which has also demonstrated efficacy in non-small cell lung cancer, are all examples of recently approved immunotherapeutic agents(35).

Furthermore, genomic studies that predict treatment response to these immunotherapeutic targets in tumours and the host have helped to stratify patients and their tumours for individualized treatment decisions(7). A more recent development has been the FDA approval of the first immune checkpoint inhibitor, Atezolizumab, used in the treatment of previously treated locally advanced or metastatic TNBC in the adjuvant setting(36). Importantly, findings from the KEYNOTE-522 trial showed a significant benefit of adding another checkpoint inhibitor, Pembrolizumab, to chemotherapy in the neoadjuvant setting in early stage TNBC with a 13.6% difference in pathologic complete response between the test group and the placebo arm(37).

This study was based on the hypothesis that exposure to multiple, recurrent, and persistent viral and parasitic infections, resulting in a heightened immune surveillance, would mean that the immune and inflammatory response to breast cancer in women of African ancestry within the continent would be very different from women with BC in the West. This unique immune response within the tumour tissue microenvironment would thus predict for risk factor, pathologic and clinical outcomes.

The research strategy, therefore, was to investigate the underlying and unique interaction of the tumour microenvironment and the associations between clinical characteristics, breast cancer risk factors, pathologic features, molecular breast cancer subtypes in Kenyan women with breast cancer.

CHAPTER 2

2.0 MOLECULAR SUBTYPES OF BREAST CANCER IN KENYAN WOMEN AND RISK FACTOR ASSOCIATION

2.1 Background

The classification of breast cancers into four major molecular subtypes is based on immunohistochemistry (IHC) surrogates for Oestrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2) status. Further categorisation of these subtypes is based on IHC determination of cytokeratin (CK) 5/6, epidermal growth factor receptor (EGFR) and the proliferation marker Ki67. The molecular subtypes can be further approximated to intrinsic subtypes using surrogate IHC markers. More recently, Gene Expression Profiling (GEP) has been used more frequently in clinical studies as an additional prognostic sign in clinical studies. Although, the comparative outcomes based on the intrinsic and clinical subtyping classification systems have shown a sizable discrepancy, current clinical practice continue the use of IHC to establish breast cancer subtypes based on ER, PR, HER2, and Ki67 staining(38).

Epidemiologic studies have demonstrated that established BC risk factors, including genetic, environment/lifestyle, and morphologic features such as terminal duct lobular unit (TDLU) involution variables, varies across different breast cancer subtypes(39) suggesting that different subtypes may develop through etiologically different pathways. Furthermore, the association of some risk factors with subtypes has been shown to vary with ethnicity. As an example, there is increased risk associated with obesity for triple negative breast cancer (TNBC) among premenopausal women while breastfeeding is protective for TNBC in African Americans but the association is not as strong in Caucasians(40). Few studies have included risk factor heterogeneity among indigenous African populations in whom common breast cancer risk factors such as nulliparity and exogenous hormone use have low prevalence and exposures to infectious agents such as malaria are more common, perhaps contributing to the observed differences in BC subtypes in this population.

Interestingly, regional variations in BC subtypes has been observed within the African continent despite sharing of similar risk factors(41) As an example, Ghanaian and Kenyan women with BC share reproductive factors such as prolonged breastfeeding duration and high parity but the frequency of receptor subtypes varies among the two populations(41, 42); the previous reported frequency of TNBC is 40% in Ghana(25), whereas it is 20% in Kenya(43).

LITERATURE REVIEW

2.2 Breast Cancer

2.2.1 Epidemiology

Breast cancer incidence and mortality demonstrate considerable variations in different populations, with low-resource countries showing lower incidence but higher mortality rates compared to high-resource countries. Even within the United States, African American (AA) women are 40% more likely to die as a result of breast cancer than white women(44). In addition, African American women suffer disproportionately from more aggressive forms of breast cancer such as TNBC(45). Women of African descent have a disproportionate mortality rate to incidence of breast cancer when compared with countries in the West. (**Figure 1**). Also, in the last decade twice as many cases of breast cancer occurred in individuals aged 15 to 49 years in developing countries than occurred in the same age group in the developed world(46).

The Kenya country figures indicate that, breast cancer is the most frequently diagnosed cancer among women representing 20.8% of all cancer cases, and the second most common of cancer mortality(2); it is estimated that the burden of breast cancer in Kenya will double by 2030(47). One of the likely contributing factors is the advanced stage at diagnosis. Local data show that over 60% of Kenyan women are diagnosed with Stage 3-4 disease(43).

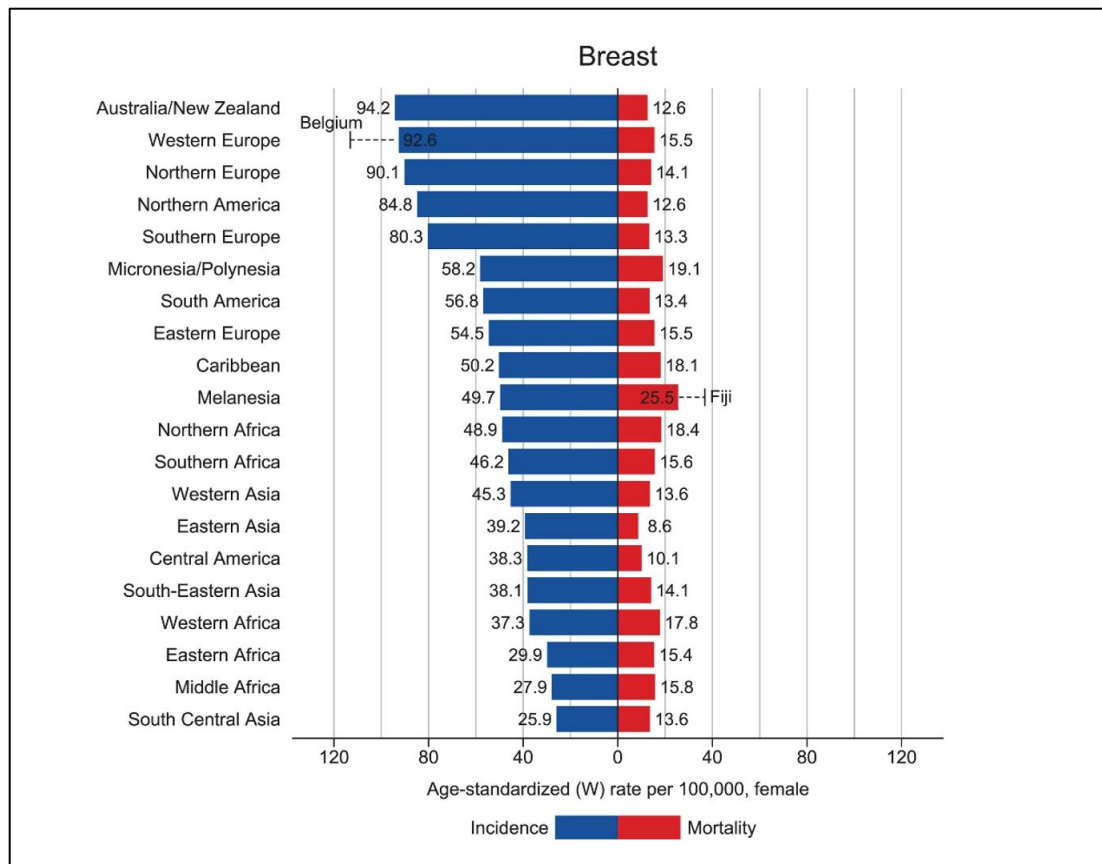


FIGURE 1: BREAST CANCER

BAR CHART OF REGION-SPECIFIC INCIDENCE AND MORTALITY AGE-STANDARDIZED RATES FOR CANCERS OF THE FEMALE BREAST IN 2018. RATES ARE SHOWN IN DESCENDING ORDER OF THE WORLD (W) AGE-STANDARDIZED RATE, AND THE HIGHEST NATIONAL AGE-STANDARDIZED RATES FOR INCIDENCE AND MORTALITY ARE SUPERIMPOSED(48). ©WILEY, 2018.

2.2.2 Race and Breast cancer

There is clear evidence that non-biological factors such as socioeconomic status, access to health care, living environments, and nutrition have a strong influence on racial disparities. However, even after controlling for these factors, overall phenotypic differences in breast cancer pathology and clinical outcomes exist among different ethnic groups. For example, several studies have reported worse survival for African-American women with TNBC after controlling for socioeconomic and treatment factors(49, 50) Further, a recent study comparing borderline oestrogen receptor (ER)-positive breast tumours (breast tumours expressing 1-10% ER) between white and African-American women suggested that the black race predicted high risk for recurrence among ER-negative borderlines and may be associated with poorer endocrine therapy response(51).

Recent work using molecular profiling and sequencing analyses suggest that differences in biological characteristics driven by distinct genetic factors and environmental exposures may contribute to racial heterogeneity in addition to disparities in health-care access and social-economic status. A somatic genomic analysis found that ER +/HER2- breast tumours in Nigerian women demonstrated more aggressive genomic features such as increased prevalence of the homologous recombination deficiency signature, increased frequency of TP53 mutations, and greater structural variations(52).

In addition, distinct survival-associated gene expression and methylation patterns have been found in breast tumours among African American compared to white women(53-56)

supporting the link between African ancestry and aggressive tumour biology. Indeed, the percentage of African ancestry estimated from germline genetic markers was significantly associated with TNBC tumours(57), suggesting that germline genetic factors may partly contribute to racial differences. In addition to genetic factors, obesity, and lack of breastfeeding, which showed higher frequency among African American women, have also been associated with the increased risk of developing TNBC in African Americans(58, 59). Studies of the association between breastfeeding and breast cancer risk by receptor status have yielded conflicting results, with some demonstrating similar protective effects for all subtypes and others for ER-negative only, particularly TNBC. In a study from Ghana(41), the incidence of ER-negative breast cancer, specifically TNBC, was higher than in the Kenyan study (28% vs. 18% of tumours. . In the Ghana study, breast feeding for more than 13 months per pregnancy was found to be protective for triple negative breast cancer in women less than 50 years of age whereas in older women extended breast feeding showed an inverse association, regardless of ER status although, the association was stronger for ER positive breast cancers unlike in our study, where regardless of age, there was no association between extended breastfeeding and breast cancer subtype (42). These results highlight the differences in risk associations with BC subtypes despite similar parity and breastfeeding characteristics among African populations emphasising the significant need for large-scale molecular epidemiological studies among various populations and regions from sub-Saharan Africa(42).

2.2.3 Breast Cancer Risk Factors

Using epidemiologic data and tissue samples collected in a Polish Breast Cancer Study (PBCS) Yang et al(60) demonstrated that distribution of established breast cancer risk/protective factors, including genetic, non-genetic, and terminal duct lobular unit (TDLU) involution variables, varied in different breast cancer subtypes, suggesting that different subtypes may develop through etiologically different pathways. Known risks associated with development of BC include non-modifiable factors such as female gender, age over 40 years, early menarche, late menopause, inherited mutation in the BRCA1 or BRCA2 gene, family history of BC and previous history of proliferative lesions of the breast(61, 62). Modifiable risk factors include obesity, nulliparity, lack of physical activity, limited duration of breastfeeding and use of combined hormone therapy after menopause(62).

2.2.3.1 Genetic susceptibility

The Breast Cancer genes 1 and 2 (*BRCA1* and *BRCA2*), are located on chromosome 17q and 13q respectively, and inherited as an autosomal dominant pattern. Mutations in these two tumour suppressor genes confers an increased risk of breast ovarian and prostate cancers(63).). The two major genes *BRCA1* and *BRCA2* as well as the minor genes *PALB2*, *P53*, *PTEN*, *CDH1*, *RAD51*, *MLH1*, *MSH2*, *PMS2*, and *EPCAM* have all been implicated as risk factors for these cancers. For *BRCA1* carriers and *BRCA2* carriers, the lifetime risk of developing breast cancer is projected to be 72% and 69%, respectively, whereas the lifetime risk of developing ovarian cancer is predicted to be 44% and 17%, respectively(64).

Racial/Ethnic variations have been observed in the mutational prevalence of these two genes and the variants for the same. About two percent of Ashkenazi Jews have a harmful variant in one of these two genes, more commonly *BRCA1*, although the prevalence in the general population is one in 400, and the mutational variants observed are frequently one of three distinct types called founder mutations since they have common ancestral haplotypes.

African Americans in the United States have *BRCA1* mutations that are not present in other racial/ethnic groupings(65-67). In the Icelandic community, those who inherit a *BRCA1* mutation frequently have a distinct founder mutational form of the gene. There are also

founder mutations in other populations, including the Norwegian and Dutch(68). Among Italian women BRCA2 is more dominant while in women younger than 50 years of age with breast cancer, the highest frequency of pathogenic/likely pathogenic variants have been observed in non- Hispanic blacks and whites while variants of unknown significance were more frequent in Asians and

Hispanics(69). In a recent study germline screen testing for BRCA1/BRCA2 was performed using NGS on peripheral blood sampling from 2168 individuals (high risk, clinically unaffected and patients with cancers of breast, endometrium, ovary, and prostate) across 7 regions of Turkey over a 7-year period. More than one genetic change in either of the two BRCA genes was seen in thirteen cases. The 225 distinct genetic abnormalities among the 471 discovered variants included 39.7% (n = 187) in the BRCA1 gene and 284 (60.3%) in the BRCA2 gene. Only 28 of the Variants of Unknown Significance VUS (n = 116) were found in clinically unaffected persons compared cancer patients who had the majority of VUS identified resulting in a comparable ratio(70).

Founder Pathogenic variants in BRCA1/ BRCA2 have also been described in racial/ ethnic groups from South Africa. Among the Afrikaners with familial breast cancer, BRCA1 c.2641G>T, BRCA1 c.1374del and BRCA2 c.7934del, the latter being the most common, have been assigned founder status while the BRCA2 c.5771_5774del and BRCA2 c.582G>A, have been identified as Black South African founder mutations(71).

A study from Tunisia reported that BRCA1-c.211dupA seems to be the most frequent BRCA mutation in Tunisia, accounting for approximately 41% of all identified BRCA1 mutations with a founder effect of the c.211dupA mutation originating from North-East of Tunisia(72). In a study from Senegal, which included 20 unrelated index cases of breast cancer, haplotype analysis revealed that specific alleles of three microsatellite markers—D17S855, D17S1323, and D17S1325—with sizes of 144, 156, and 173 bp, respectively—segregated together in the index cases and formed a common haplotype of about 400 kb. A founder effect of the pathogenic variation in the study population is suggested by the fact that this haplotype was not observed in any of the healthy controls investigated(73) .

In BRCA1 cancers, the two primary evolutionary paths were determined by the expression or loss of PTEN expression. The most likely initial event was PTEN loss, which was followed by a TP53 mutation or a BRCA1 LOH mutation with roughly similar likelihood. Triple negative cancers almost always showed PTEN loss as the first event, whereas luminal tumors displayed mutant TP53 or BRCA1 LOH as the first event, thus evolutionary routes to carcinogenesis were significantly related with the tumor subtype. Sporadic triple-negative breast cancers showed a reduced frequency of PTEN deletion and allelic imbalance of the BRCA1 locus was also infrequently (5%) seen. Even in cancers that exhibit apparent loss of wild-type BRCA1, loss of BRCA1 may not be the initial event in the majority of BRCA1-associated breast malignancies(74). The synergy between mutations in p53 and the BRCA genes in tumours indicates a role for loss of checkpoint control in tumour progression(63).

Although more than 95% of BC are sporadic which lack mutations in the *BRCA1* gene, absent or significantly reduced expression levels of *BRCA1* observed in these sporadic tumours may be a result of promoter hyper methylation leading to epigenetic silencing of *BRCA1* expression thus favouring proliferation of aggressive BC phenotypes (75). A study from Iceland reported BRCA1 gene's promoter is hypermethylated in a significant proportion (9.1%) of 143 primary sporadic breast tumors. The breast tumor tissue was analysed for TP53 (also known as p53) mutations by denaturing gel electrophoresis, BRCA1 promoter methylation by methylation-specific PCR, allelic imbalance (AI) at the BRCA1 and BRCA2

loci, and microsatellite analysis. Fluorescence in situ hybridization was used to check for BRCA1 copy number variations in the BRCA1 methylated tumors, and immunostaining was used to check for BRCA1 expression. It was discovered that the majority of the BRCA1 methylation tumors had no or significantly reduced BRCA1 expression, indicating transcriptional suppression by epigenetic changes. Additionally, the authors discovered evidence that AI at the BRCA1 locus was linked to BRCA1 methylation but AI at the BRCA2 locus was not. This suggested that BRCA1 methylation is uniquely linked to AI at the BRCA1 gene. FISH was used to further analyze the genetic changes at the BRCA1 locus, which indicated chromosome 17 deletions and heterogeneity in terms of chromosomal abnormalities. These findings imply that the BRCA1 gene is methylated in association with genomic rearrangements at the BRCA1 locus, which causes the genetic material containing non-methylated BRCA1 alleles to be lost and the methylated BRCA1 alleles to be retained. Additionally, they found that the subset of BRCA1 methylated tumors had a considerably higher frequency of TP53 mutations than had been previously known, suggesting that BRCA1 methylation may affect the same molecular pathways that are frequently changed in family BRCA1 malignancies. Collectively, these data implied that the BRCA1 gene's epigenetic silencing is a potential cause for sporadic breast cancer(76).

Mutations in *BRCA1* have been shown experimentally to result in accumulation of undifferentiated luminal progenitor cells, the suspected cell-of-origin of basal-like breast cancer. A greater number of terminal ductal structures and a reduced number of lobules have been linked to increased risk of mammary tumours suggesting that environmental insults that induce hyper methylation of the *BRCA1* gene and a less differentiated mammary gland may increase the risk of breast tumorigenesis. Although rare, mutations of *BRCA1* and *BRCA2* have high penetrance and increase the risk of BC by 10-20 times. Specifically, in *BRCA1* mutation carriers, the risk of BC is 60-85% and significantly increased in younger women(63, 77). Both germline *BRCA* mutations and Polygenic risk scores (PRS) are differentially associated with subtypes, suggesting that in addition to the previously observed heterogeneity for *BRCA1* mutations, inherited low-risk variants could also differentially increase risk of specific breast cancer subtypes(78). Also, *BRCA1* mutated tumours are associated with the triple negative phenotype, whereas *BRCA2* mutated tumours are associated with the luminal B subtype(77). More recent developments in molecular techniques, particularly Next-Generation Sequencing, have enabled discovery of multiple novel genes with varying penetrance estimations that are involved in the risk for breast and/or ovarian cancer. High penetrance genes for the risk of breast/ovarian malignancies include TP53, PTEN, STK11, and CDH1 whereas PALB2, BRIP1, ATM, CHEK2, BARD1, NBN, NF1, RAD51C, RAD51D, and mismatch repair genes have been identified as moderate and low penetrance genes for breast/ovarian cancer risk (Table 1)(79).

TABLE 1: ADDITIONAL GENES INVOLVED IN INCREASED RISK OF BREAST CANCER (BC)

NAME OF GENE	LOCATION	BC RISK	FUNCTION	MECHANISM OF ACTION IN BC
STK11(LKB1)	19p13.3	32–54%	Tumor suppressor	Encodes a serine/threonine kinase that regulates energy metabolism and cell polarity.
PTEN	10q23.31	25–85%	Tumor suppressor	Encodes a phosphatidylinositol-3,4,5-trisphosphate 3-phosphatase, which antagonizes the PI3K signalling pathway through its lipid phosphatase activity and negatively regulates the MAPK (mitogen-activated protein kinase) pathway through its protein phosphatase activity.
TP53	17p13.1	25–79%	-Tumor suppressor -Apoptosis -DNA repair -Genomic stability, -	Encodes the p53 protein, tumor suppressor. Responds to different cellular stresses to regulate expression of target genes, thereby inducing cell cycle arrest, apoptosis, senescence, DNA repair, or metabolism changes. Restoration of p53 activity has shown to lead to tumor regression hence it's a potential target for therapies. A biomarker for the diagnosis, tumor progression, poor prognosis, and reduced sensitivity for anticancer drugs.
CDH1	16q22.1	39–52% (by age 80)	-Tumor suppressor	Encodes the E-cadherin, a transmembrane calcium-dependent protein involved in cell-cell adhesion. Somatic genetic and epigenetic alterations of the CDH1 gene associated with poor survival. - CDH1 hypermethylation is a potential novel drug target for developing personalized therapy.
ATM	11q22.3	17–52%	DNA repair	Encodes a phosphatidylinositol 3-kinase protein that acts as a tumor suppressor and responds to DNA damage by phosphorylating key

				substrates involved in DNA repair and cell cycle control.
CHEK2	22q12.1	25–39%	DNA repair Induces apoptosis	Encodes a nuclear Ser/Thr kinase that acts as a tumor suppressor in different cellular processes. In response to DSBs, CHEK2 protein is phosphorylated by ATM and catalyzes the phosphorylation of CDC25C, down-regulating it and preventing entry into mitosis. After DNA damage, CHEK2 phosphorylates the p53 tumor suppressor protein and prevents its degradation, leading to cell cycle arrest in G1. Under gamma irradiation, CHEK2 also phosphorylates BRCA1 on Ser-988, activating the DNA repair Process. CHEK2 induces apoptosis independently from p53, via phosphorylation of the PML (promyelocytic leukemia protein) tumor suppressor.
BARD1	2q35	NA	-Tumor suppressor -Involvement in the (Homologous recombination repair) HR process	Encodes a protein that interacts with the N-terminal region of BRCA1 and acts as a tumor suppressor creating a RCA1/BARD1 heterodimer with ubiquitin E3 ligase activity.
NF1	17q11.2	~8% (by age 50)	-Tumor suppressor -Increases the rate of guanosine triphosphate (GTP) hydrolysis.	Encodes Neurofibromin, a cytoplasmic protein that regulates several intracellular processes, including the Ras-cAMP (cyclic adenosine monophosphate) pathway, the ERK (extracellular signal-regulated kinase)/MAPK cascade, adenylyl cyclase, and cytoskeletal assembly. Neurofibromin has been shown to increase the rate of guanosine triphosphate (GTP) hydrolysis of Ras and act as a tumor suppressor by reducing

				Ras activity. NF1 alterations increase cancer risk and drug resistance.
NBN	8q21.3	12–30%	-Repair of Double Stranded Breaks (DSBs) -Telomere maintenance -Tumor suppressor	The MRN complex (MRE11-RAD50-NBS1) is a complex of three proteins (encoded by MRE11, RAD50, and NBN genes). These proteins are known tumor suppressors because loss of function of any of them results in genome instability.
PALB2	16p12.2	44–63% (by age 80)	Recombinational repair and checkpoint functions.	Encodes a protein that colocalizes with BRCA2 in nuclear foci, promotes its localization and stability in nuclear structures, and enables its recombinational repair and checkpoint functions. Binds the single-strand DNA and directly interacts with the recombinase RAD51 to stimulate strand invasion during the HR process.
BRIP1	17q23.2	NA	DSB repair	Encodes a DNA helicase that binds directly to the BRCT repeats of BRCA1 and is necessary for efficient DSB repair in the HR process.
RAD51C	17q22	NA	Involvement in the Homologous recombination (HR) process	Encodes a protein that acts in HR creating complexes with other RAD51 paralogs.
RAD51D	17q22	NA	Involvement in the HR process	Encodes a protein that interacts with other RAD51 paralogs in the process of HR
MLH1	3p22.2	NA	Repair of DNA mismatches	PVs in the MLH1, MSH2, MSH6, and PMS2 genes, coding the proteins of the DNA mismatch repair (MMR) or by large deletions of the EPCAM gene, located upstream of the MSH2 gene. MMR proteins work coordinately in sequential steps to initiate repair of DNA

				mismatches, constituted by erroneous insertions, deletions, and substitutions of bases that can arise during DNA replication and recombination.
MSH2	2p21-p16	NA	Involvement in the HR process	As above
MSH6	2p16.3	NA	Involvement in the HR process	As above
PMS2	7p22.1	NA	Involvement in the HR process	As above
EPCAM	2p21	NA	Involvement in the HR process	As above

Key

Adapted from: Angeli D et.al(79).

Key

MAPK mitogen-activated protein kinase

PML promyelocytic leukemia protein

HR Homologous recombination

DSBs double-strand breaks

MRN MRE11-RAD50-NBS1

MMR Mismatch repair

2.2.3.2 Age

The incidence of breast cancer is highly related to increasing age probably related to the effect of increased lifetime duration of oestrogen and progesterone effect on the breast(80). The median age of White Americans with breast cancer is 63 years whereas that of African Americans is 60 years(81) with only 4% of American women developing breast cancer under the age of 40 years(82). In contrast, in Kenya the median age appears to be a decade earlier at 49 years(83).

The incidence of breast cancer increases exponentially until menopause and thereafter more slowly. This reflects rates of early and late onset disease, and suggests the persistence of replicating cell populations with increase in age(81, 84). Less than twenty-five per cent of breast cancer cases are diagnosed below the age of 50 years. During the peri- and early postmenopausal years, most involution of the breast lobules occurs. If this process is slowed down or stalls it is a significant predictor of increased breast cancer risk. (85).

It has been shown that irradiation of the breast after the age of 35 years fails to increase subsequent breast cancer risk. This suggests that BC initiating events must occur earlier in life for breast epithelium susceptibility to full malignant transformation(84). In radiation as a mechanism, apoptosis, senescence and tumour suppressing genes, specifically p53 and

p16INK4a regulatory proteins, ensure that cells with damaged DNA but uncontrolled replication competent cells are eliminated prior to malignant transformation(84). Just like other epithelial malignancies, breast cancer shares characteristics associated with the normal ageing process. This includes telomere shortening and increased genomic instability, together with promoter and global specific epigenetic changes, altered gene expression in normal cell cycle and lastly, extracellular matrix remodelling(84).

2.2.3.3 Family history

In a cohort of over 113,000 women, from the general UK population included in The Generations Study(86); it was shown that women with one first-degree relative with breast cancer have a 1.75-fold increased risk of developing the disease compared to women without any affected female relatives. The study also showed that the risk rose 2.5-fold or greater in women with two or more first-degree female relatives with breast cancer(87). The observed inherited susceptibility to breast cancer in part was attributable to mutations of breast cancer related genes for example *BRCA1* and *BRCA2*(80), both these genes of are importance in the DNA repair pathway in the cell cycle.

2.2.3.4 Reproductive factors

Early menarche, late age at first pregnancy, low parity and late menopause are reproductive factors strongly associated with ER status. All of these have been shown to increase risk of development of breast cancer (80, 88). Compared to multiparous women: nulliparous women have 30% higher breast cancer risk. Nulliparous women have a higher number of ovulatory menstrual cycles than do parous women because of the absence of pregnancy and lactation, and studies have shown that an increased number of cycles leads to an increase in breast, ovarian and uterine cancers. Women experiencing menarche before age 12 years had a slightly higher risk of breast cancer than did those who were older than 15 at menarche. Similarly, the risk of breast cancer is increased by 17% for every 5-year delay in menopause. Bilateral oophorectomy before age 40 years led to a 45% reduced risk of breast cancer compared with women with a natural menopause at ages 50–54 years(89).

2.2.3.5 Breast feeding and Parity

Breastfeeding is assumed to protect a woman from breast cancer by increasing breast differentiation, postponing the return of the ovulatory menstrual cycle post-pregnancy, and decreasing breast fluid oestrogen levels, all of which decrease exposure to total oestrogens that has been related to proliferative and carcinogenic effects(41). Compared with nulliparous women, never breastfeeding was associated with an increased risk of breast cancer among parous women, whereas breastfeeding was protective for breast cancer(78). In a hospital-based case-control study of 7,663 Malaysian women, a higher breastfeeding rate and duration, soy intake and level of physical activity were associated with a reduced risk of breast cancer among Southeast Asian women(88). In a population-based case-control study from Ghana (2013 to 2015), data from 1,126 breast cancer cases (aged 18-74 years) and 2,106 age matched population controls were analysed. The cases more often than the controls reported late age at menarche, fewer births, late age at first birth and low median breastfeeding months(41). Increased parity and breastfeeding were associated with reduced risks for both ER negative and ER positive tumours, with the lowest risks observed among women with ≥ 3 births who breastfed for ≥ 13 months per pregnancy compared to nulliparous women and among women aged 50 years and above.

On the contrary, among women aged < 50 years with ER negative tumours, compared to nulliparous women, the highest risk was for those with ≥ 3 births who breastfed < 13 months per pregnancy(41). Women with ≥ 3 births who breastfed, on average, ≥ 13 months per

pregnancy were not at increased risk (OR 1.09, 95% CI 0.56–2.10), due to the multiplicative joint association of two factors associated with risk in opposite directions(41).

2.2.3.6 Hormone replacement therapy (HRT) and Oral contraceptives

The use of hormone replacement therapy and preventive pregnancy hormones are among risk factors for breast cancer(90). The association of oral contraceptives with breast cancer has been linked with duration, dose, usage pattern, type of oral contraceptive and age of first use(87). While some studies have found a moderately increased risk of developing breast cancer with recent use of OC, especially with long-term use, recent reviews and case-control studies have suggested a protective effect of OC use in some populations and especially when using progesterone only oral contraceptives(91). Use of HRT has been associated with risk for the luminal A subtype with no effect among the other subtypes(78).

2.2.3.7 Oestrogens

Both endogenous (usually produced by the ovary in premenopausal women) and exogenous oestrogens (whose main sources are the oral contraceptives and the hormone replacement therapy) are associated with the risk of breast cancer(80).

Ovarian hormones, oestrogen and progesterone play a key role in the development and changes seen in normal breast tissue during puberty, pregnancy, lactation, and involution. Although growth hormone on its own can induce mammary duct development, the synergistic effect of oestrogen and growth hormone is greater. Furthermore, it has been demonstrated that oestrogen replacement on its own is not effective in restoring the development of mammary glands in ovariectomized and hypophysectomised animals(92).

On binding to the oestrogen receptor in breast epithelium, oestrogen has a proliferative effect borne by direct in vitro evidence in breast cancer cells showing that oestrogen inhibits apoptosis by up-regulation of Bcl-2, an anti-apoptotic proto-oncogene(92).

In a study involving 767 women with breast cancer aimed to characterise associations between circulating concentrations of oestrogens, progesterone, and androgens with breast cancer, and related risk factors, the levels of oestradiol, calculated free oestradiol, oestrone, androstenedione, DHEAS and testosterone were positively associated with breast cancer risk. Using a doubling scale, the ORs were 1.19 (1.06–1.35) for oestradiol, 1.17 (1.03–1.33) for calculated free oestradiol, 1.27 (1.05–1.54) for oestrone, 1.30 (1.10–1.55) for androstenedione, 1.17 (1.04–1.32) for DHEAS, and 1.18 (1.03–1.35) for testosterone(93).

2.2.3.8 Body mass index

Body mass index (BMI) has been positively correlated with breast cancer. BMI (overweight category) has the second highest odd ratio (OR = 1.61 [95% CI 1.43–1.80]) after nulliparity(90). Obesity is a known risk factor for breast cancer in both pre- and post-menopausal women. Obesity can disrupt some biological pathways, resulting in, insulin resistance, and synthesis of endogenous sex hormones(87, 90). Surprisingly, increased BMI was significantly associated with lower breast cancer risk in both pre- and post-menopausal women in a hospital-based case-control study of 7,663 Malaysian women(88).

Relatively high body fat levels associated with an elevated risk of invasive breast cancer was the finding of a study of post-menopausal women. The study also demonstrated a 56% increase in the risk of development of ER positive breast cancer per a 5-kg increase in truncal fat, despite normal BMI(94). Study women who remained in the normal BMI range during the follow-up period showed a persistent association between higher trunk fat and increased risk

of ER positive breast cancer was demonstrated(94).

A pooled analysis of 7 prospective cohort studies, which included 337,819 women and 4385 incident invasive breast cancers, reported an inverse association between BMI and premenopausal breast cancer risk when comparing women who had a BMI >33 kg/m² versus those who had a BMI ≤21 kg/m² (risk ratio [RR], 0.58; 95% confidence interval [95% CI]. 0.34-0.85). On the other hand, in postmenopausal women the relative risk of breast cancer was increased with increasing BMI (RR,1.27; 95% CI: 1.03,1.55)(95).

2.2.3.9 Physical activity and breast cancer

Regular physical activity of various intensity decreases the risk of breast cancer among postmenopausal (RR = 0.79, 95% CI = 0.76–0.84) and pre-menopausal women (RR =0.80, 95% CI = 0.74–0.87)(88). Pooled relative risks (RRs, 95% CI) for women with higher versus lower levels of moderate-vigorous recreational activity were RR = 0.80 (0.74-0.87) and RR = 0.79 (0.74-0.84) for premenopausal (43 studies) and postmenopausal (58 studies) breast cancer, respectively, with high heterogeneity. Inverse associations were weaker among postmenopausal cohort studies (RR = 0.90 [0.85-0.95]) and studies that statistically adjusted for non-recreational (e.g., occupational, household) activity (RR = 0.91 [0.77-1.06] premenopausal, RR = 0.96 [0.86-1.08] postmenopausal)(96). A suggested positive association between physical activity and premenopausal breast cancer risk, which contrasts with a meta-analysis that reported a RR of 0.80 (95% CI: 0.74–0.87) for premenopausal breast cancer was observed(97).

2.2.3.10 Alcohol and Smoking

The risk of breast cancer is associated excessive alcohol consumption increased dietary fat intake. With. Alcohol consumption can elevate the level of oestrogen-related hormones in the blood and trigger the oestrogen receptor pathways(80). In one study conducted by Jones.et.al(69)it was shown there is no raised risk of development of breast cancer with smoking in non-drinkers. The same study demonstrated that a significantly raised breast cancer risk in those who had ever been drinkers.

The association between smoking and breast cancer risk remains controversial. Mutagens from cigarette smoke have been detected in the breast fluid from non-lactating women(98).

In a study involving 102,927 women aged 16 years or older in the Generations Study cohort recruited from 2003 to 2013, Jones et.al(69) made the following observations. There was an increased risk of invasive breast cancer in ever and former smokers. This also applied to women who smoked more than five cigarettes daily, had 10 plus pack-years of use and had stopped for less than 20 years.

Another finding was that there was an increased risk with 10 plus years' duration of smoking, but no appreciating trend beyond 10 plus years. Risk of development of BC was significantly and similarly raised in former and current smokers.

The findings of another study were that smoking was associated with a modest but significantly increased risk of developing breast cancer. This was observed among those who started smoking at adolescent or peri-menarche ages. In addition, it was reported that the relative risk of breast cancer associated with smoking was significantly greater for women with a positive family history of the disease(69).

2.2.4 Pathology

Breast carcinomas are adenocarcinomas arising from the terminal duct lobular unit(99). Invasive ductal carcinoma (IDC), no special type is the most common histological type diagnosed accounting for up to 75% of all breast cancers whereas the special histologic types account for the rest(99) .

In addition to the invasive cancers, in situ carcinomas of the breast are also described including ductal and lobular carcinoma in situ (DCIS and LCIS). Models have been used to describe the progression from normal healthy breast tissue to atypical ductal hyperplasia to in situ carcinoma to breast carcinoma(100). Multiple linear pathways for the development of breast cancer have been described based on the Estrogen Receptor (ER) status, the Cell of Origin pathway, and the molecular pathways. Flat epithelial atypia (FEA), Atypical ductal hyperplasia (ADH) and ER positive Ductal carcinoma in situ (DCIS), are considered non-obligatory precursors of ER positive invasive breast cancer whereas micro glandular adenosis and ER negative DCIS are considered as precursors of ER negative breast cancer. Two distinct pathways in breast carcinogenesis at the cell-of-origin level have been described: the sporadic clonal evolution model and the cancer stem cell model. At the molecular level, gains of 1q, loss of 16q, sporadic amplification of 17q12, and a gene expression pattern that is predominately made up of genes linked to the ER-positive phenotype define the ER-positive pathway. These lesions have diploid or near-diploid karyotypes, express hormone receptors, lack HER2 overexpression, and express basal markers such as CK5/6. Most of the neoplastic lesions are low- to intermediate-grade phenotype. The second molecular pathway, known as the ER-negative pathway, is primarily distinguished by loss of 13q, gain of chromosomal region 11q13, amplification of 17q12, and a gene expression pattern made up of genes involved in cell-cycle and proliferation-related events. Many of the tumors along this pathway are intermediate- to high-grade tumors and include both the HER2 positive and HER negative tumours in the latter subtype of which TP53 mutations are commonly seen while PIK3 mutations are associated with ER negative HER2 positive tumours. High proliferative activity, an enhanced immune infiltration, a basal-like and a mesenchymal phenotype, and a lack of homologous recombination are significant features observed in ER-negative, HER2-negative breast cancer(101)

Other morphologic subtypes of Invasive breast carcinomas include tubular carcinoma, cribriform carcinoma mucinous carcinoma, mucinous cystadenocarcinoma, micropapillary carcinoma and metaplastic carcinoma and carcinoma with apocrine differentiation. More than 90% of the tumour must show a special histopathology pattern for it to be designated as a special type of breast cancer. The morphologic subtype, tumour grade, lymphovascular invasion and associated DCIS are important features reported that characterise biologic behavior. Invasive breast carcinomas are graded based on the Nottingham grading system(102). Three features evaluated tubule formation, nuclear pleomorphism and mitotic activity with each feature being scored independently from 1-3 giving aggregate total of 9. Grade 1 tumours have an aggregate of 3-5 points out of 9, grade 2 tumours 6-7 points out of 9 and Grade 3 tumours are defined as those with 8 or 9 points out of 9.

2.2.5 Traditional subtyping of IBC

The time-honoured procedure of classifying and staging of breast cancer is based on clinical and morphologic features. The latter involves evaluates tumour type, size, grade, lymphovascular invasion (LVI), the number and extent of lymph node involvement. Also the expression of hormone receptors; oestrogen (ER) and progesterone receptors (PR) with the Human Epidermal Growth Factor Receptor 2 (HER2)- a member of the tyrosine kinase

receptor family are all required (103). A combination of latter factors is used by various classification systems categorize patients into risk factor and prognostic groups(104, 105). Individual patients in the same risk category may have different clinical outcomes which can be attributable to the heterogeneity of breast cancer. The refinement of the traditional classification system has been provided for by access to molecular classification to better evaluate histology specimens to enhance the ability to assess the prognosis and predict response to therapy of BC at an individual level(106).

2.2.6 Molecular classification of IBC

More than 50% of ER positive BCs do not respond to hormone therapy either due to intrinsic or acquired resistance(106)therefore there is need for more individualized classification systems. Molecular methods especially gene expression profiling, are emerging as useful tools for defining molecular subtypes of breast cancers to assess prognosis and predict response to therapy (**Table 2**). In 2000 and 2001, Perou and Sorlie et al(107, 108)described the results of 65 and 85 invasive breast carcinoma specimens from 42 and an additional 44 individuals respectively. Using 8,102 genes on a complementary DNA microarray platform, they were able to demonstrate that there were great variations in IBC at the genetic level and through hierarchical clustering IBC could be classified into 4 distinct molecular “intrinsic variants” (luminal A, luminal B, HER2-enriched, and basal-like) and a Normal Breast-like group each with different clinical outcomes(109).

TABLE 2: MAJOR MOLECULAR SUBTYPES OF BREAST CANCER DETERMINED BY GENE EXPRESSION PROFILING

	MOLECULAR SUBTYPE		
	Luminal	HER2	Basal
Gene expression pattern	High expression of hormone receptors and associated genes (luminal A>luminal B)	High expression of HER2 and other genes in amplicon. Low expression of ER and associated genes	High expression of basal epithelial genes, basal cytokeratins. Low expression of ER and associated genes. Low expression of HER2
Clinical features	~70% of invasive breast cancers. ER/PR positive. Luminal B tends to be higher histological grade than luminal A. Some overexpress HER2 (luminal B)	~15% of invasive breast cancers. ER/PR negative. More likely to be high grade and node positive	~15% of invasive breast cancers. Most ER/PR/HER2 negative ('triple negative'). BRCA1 dysfunction (germline, sporadic). Particularly common in African American women.
Treatment response and outcome	Responds to endocrine therapy (but response to tamoxifen and aromatase inhibitors may be different for luminal A and luminal B). Response to chemotherapy	Respond to Trastuzumab (Herceptin). Respond to anthracycline-based chemotherapy. Generally poor prognosis.	No response to endocrine therapy or Trastuzumab (Herceptin). Appear to be sensitive to platinum-based chemotherapy and PARP inhibitors. Generally poor prognosis

variable (greater in luminal B than in luminal A).
Prognosis better for luminal A than luminal B.

(But not uniformly poor).

(103) ©Modern Pathology, 2010

Further gene expression profiling work has led to the discovery of additional new molecular subtypes of breast cancer. Gene expression analyses by Herschkowitz et al(110) described a new intrinsic subtype of IBC in human and mouse BC(110)and in BC cell lines(109)and called it the claudin-low subtype. Subsequently, Prat et al(109)Claudin-low tumours segregated close to the Basal subtype of BC when hierarchical clustering analysis of 320 human breast tumours and 17 normal breast samples using a 1900 gene intrinsic list was used(111).

Claudin-low BCs tend to be triple negative (ER/PR/HER negative) tumours, show low expression of the tight junction genes, and are highly enriched for immune response genes which are expressed by T- (CD4 T cells) and B-lymphocytes. In addition, Claudin-low tumours share biologic features with cancer stem cells(112). Clinically, they have a poor outcome although there are no differences in overall survival between these tumours and other poor prognosis BC subtypes(111). Immunohistochemistry has been used as a surrogate marker for the intrinsic BC subtypes as defined by gene expression profiling. Although, three markers ER, PR and HER2 were initially used to define the luminal A, luminal B, HER2-enriched and the basal subtypes of BC(103), this approach was further modified to include three more antibodies that better predict the molecular subtypes since not all basal like cancers are triple negative and only about 30% of luminal B subtype are HER2 positive (**Table 3**) (74).

Using two additional antibodies CK5/6 and EGFR (the five marker Core Basal antibodies), basal cell subtypes of BC can be identified with a sensitivity of 100% and a specificity of 76%. Interestingly, in a series of 4046 breast tumours(113), 17% (639 of 3744) were defined as the triple-negative, whereas 9.0% were Basal-like by the five-marker Core Basal definition. When the triple-negative group was segregated into Core Basal and the '5 Negative Profile' (5NP), the Core Basal group showed a significantly worse outcome compared to the 5NP group. The use of the proliferation marker Ki-67 additionally, helps to separate the luminal B from the luminal A molecular subtype. Thus a panel of six antibodies is more precise in predicating the molecular subtype at the gene expression level(103).

TABLE 3: IMMUNOPHENOTYPING TO APPROXIMATE MOLECULAR SUBTYPE USING SIX BIOMARKERS

Molecular subtype	Biomarker profile
Luminal A	ER positive and/or PR+, HER2–, and low Ki-67 (<14%)

Luminal B	ER positive and/or PR+ and HER2+ (luminal-HER2 group) ER positive and/or PR+, HER2–, and high Ki-67 (>14%)
HER2-enriched	ER–, PR–, and HER2+
Basal-like	ER–, PR–, HER2–, and CK5/6 and/or EGFR+

Abbreviations: CK, cytokeratin; EGFR, epidermal growth factor receptor; ER, oestrogen receptor; PR, progesterone receptor(103).

©Modern Pathology, 2010

A threshold of $\geq 14\%$ for Ki-67 was defined in the 2011 St. Gallen meeting of breast cancer experts to distinguish luminal B from luminal A breast cancer subtypes(114); these subtypes may be approximated using clinicopathological rather than gene expression array criteria. In general, systemic therapy recommendations follow the subtype classification. Thus, ‘Luminal A’ disease generally requires only endocrine therapy, which also forms part of the treatment of the ‘Luminal B’ subtype. Chemotherapy is considered for most patients with ‘Luminal B’. The cut off for Ki-67 index was derived from comparison with gene array data as a prognostic factor. However, the Ki-67 index was later revised upwards through consensus by the expert panel to $\geq 20\%$ at the 2013 St. Gallen meeting(94).

In a retrospective study, over a period of 13 years (1994-2006), 9,415 women in Milan Italy, who had surgery for ER positive, HER2-negative breast cancer, were evaluated for distant disease-free survival (DDFS) after a median follow-up of 8.1 years. The study found that using the Ki-67 and PR threshold of 20% improved the prognostic discriminatory power of the subtypes. Using the 2013 proposed revision of the St. Gallen criteria, the study found that the number of luminal A increased by 9.2% (from 3169 to 4036 cases) and conversely the number of luminal B also decreased by 9.2% (from 6246 to 5,379) when compared with the 2011 St. Gallen definition of the IHC subtypes. The authors further grouped the Ki-67 into “low” ($<14\%$), “intermediate” (14% to 19%) or “high” ($\geq 20\%$) stratified by PR expression (negative or low versus high). On multivariate analysis adjusted for pathologic tumour stage, pathologic node stage, tumour grade, peritumoural vascular invasion and menopausal status, the study reported that the DDFS for women with low Ki-67 and PR less than 20% (new luminal B) and low Ki-67 and PR $>20\%$ (luminal A) were similar. On the other hand, the women with intermediate Ki-67 and high PR (new luminal A) had superior DDFS than those women with intermediate Ki-67 expression and low PR levels (HR = 1.86, 95% CI = 1.37 to 2.52; $P < 0.001$ by Gray test).

The authors therefore proposed a 20% cut off for Ki-67 for differentiation of luminal A from luminal B and that the PR cut off of 20% for designation of luminal A only be considered for cases for the intermediate category of Ki-67(115). In a population-based case-control study from Ghana (2013 to 2015)(41), data from 1,126 breast cancer cases (aged 18-74 years) and 2,106 age matched population controls were analysed. A total of 50%, 52% and 23% of cases were ER positive, PR positive and HER2-positive, respectively. Luminal A breast cancer was the most common subtype (49%) followed by triple-negative/basal-like (28%), HER2-enriched (15%) and luminal B-like breast cancers (8%). There were no significant differences in distributions of molecular subtypes between women <50 years compared to ≥ 50 years of age.

Using annotated risk factor and pathology data collected from 838 BC patients enrolled from multiple hospitals across Kenya, evaluated the distributions of established BC risk factors across BC molecular subtypes.

2.2.7 Race and subtypes of Breast Cancer

Whether there are unique genetic and environmental exposures that predispose African women to more aggressive subtypes of breast cancer remain unclear. Many genetic variants that are common in African populations may result from the adaptation to these unique exposures through natural selection. For example, Africans have longer leukocyte telomere lengths (LTLs) compared to Caucasians (70) which may make them more susceptible for specific types of cancers, including TNBC. In addition, a recent study found that pathways that are enriched for signals of adaptation tend to be involved in immune response and wound healing. Cytokine polymorphisms are now believed to alter these critical molecules that have a dual face in carcinogenesis as, when implicated in the activation of the immune response, these molecules check the cancer development while their persistent inflammatory reaction can envisage the development of malignancy and tumor growth. Furthermore, multiple SNPs within cytokine gene may potentially drive ethnic differences observed in cancer and breast cancer (13).

There is increasing evidence suggesting that chronic inflammation, driven primarily by innate immunity, may contribute to critical pathways in cancer progression. Studies have found that germline variants in innate immune genes are unique in African populations may result from the adaptation to their unique exposures through natural selection, suggesting that African breast cancer may have distinct immune profiles. In fact, a recent study investigating breast tumor microenvironment from 1315 patients from the Women's Circle of Health Study found that Black breast cancer patients displayed distinct exhausted CD8+ T-cell signatures indicating an immunobiological basis for their more aggressive disease and a rationale for the use of immune checkpoint inhibitors targeting the exhaustion phenotype (31).

In addition, a recent study found that pathways that are enriched for signals of adaptation tend to be involved in immune response and wound healing (71) which may partially explain the high prevalence in women of African ancestry of TNBC, especially basal-like subtypes, and suggest that African breast cancer may have distinct immune profiles. Although the degree to which adaptation to infectious agents or to other environmental factors has shaped the distribution of complex-disease alleles in humans is still debatable, infectious diseases represent one of the major causes of death for human populations, especially in developing countries. Not surprisingly, genes responsible for heritable variation in the response to pathogens are likely targets of natural selection while some of these may create an overactive/ over responsive immune environment leading to increased risk of cancer and autoimmune conditions (117).

Germline variants in immune genes that offer enhanced protection against environmental pathogens are enriched among African populations who were heavily exposed to those pathogens. However, these variants may be associated with pro-tumor activity leading to more aggressive tumor types or subtypes. For example, genetic variants that provide protection against Malaria are associated with increased risk of lymph node and distant metastasis among breast cancer patients (118). Nevertheless, the distribution of molecular subtypes and immune phenotypes in African breast cancer remains largely unknown. Further, Africans are not a monolithic group; there are great variations in genetic makeup and exposures across different African populations. Recent studies also suggest there may be regional variation in molecular subtypes between East and West Africa (23, 24, 72) as well as trend towards ethnic variations in breast cancer subtypes within the same country (53).

2.3 MATERIALS AND METHODS

2.3.1 Objective: Categorize breast cancers into molecular subtypes and determine association with breast cancer risk factors

2.3.1.1 *Sample Source*

Previous pathologically confirmed IBC cases consisting of definitive surgical specimens (403) and core breast core needle biopsies (435) were selected from the Glaxo Smith Kline – Triple Negative Breast Cancer (GSK-TNBC) study archive. The GSK-TNBC was a Kenya wide prospective, funded (GSK-Ethnic Research Initiative) study involving 15 public, faith-based and private institutions, which enrolled a total of 867 female patients with pathologically confirmed breast carcinoma between March 2012 and May 2015. The main objective of the TNBC study was to determine the prevalence of TNBC in the major ethnic groups within Kenya(116). Relevant socio-demographic, clinical, reproductive, and known breast cancer risk factor data were collected using a standardized questionnaire.

For this current work, data for pathologic characteristics of prognostic and predictive significance which included histologic grade, histologic tumour type, tumour size, lymph node stage, lymphovascular invasion and ER/PR/HER2 status, and FISH for HER2 status where applicable were extracted from the TNBC database. Additionally, the patients' socio-demographic and risk factor data; age, family history of BC, the age at menarche and menopause, contraception history and parity were also documented.

Where available, tissue blocks of confirmed tumour were retrieved and analysed for further molecular characterisation using immunohistochemical antibodies as surrogates.

2.3.1.2 *Molecular Subtypes*

The central pathology review and immunohistochemistry (IHC) for ER/PR/HER2 of all breast carcinoma tissue were done at Aga Khan University Hospital, Nairobi and interpreted by SS. For ER/PR/HER2 the, ER- EPI, PR-PGR 636 and HER polyclonal Rabbit Antihuman C-erbB2oneoprotein clones were used. Additional slides were cut at 5 µm and subjected to immunohistochemistry (IHC) stains for EGFR, CK5/6 and Ki-67 (Dako Monoclonal mouse anti-human antibodies were used; EGFR Wild Type EGFR polyclonal antibody in a dilution of 1:200in a dilution of 1:200, CK5/6 clone D5/16 B4 ready to use, Ki-67 Clone MIB-1, ready to use) according to the manufacturer specifications as previously described(83, 117). On slide cervical tissue was used as a positive control and tonsillar tissue as a negative control. The slides were stained on the DAKO Autostainer link instrument.

ER and PR tumour expression were considered positive for IHC with $\geq 1\%$ nuclear staining. HER2 expression was determined by IHC and fluorescence in situ hybridization (FISH) in case of an equivocal HER2 result.

The latter was outsourced to a CAP accredited laboratory. An IHC score of 3+ or a FISH-positive test result was defined as HER2+. The FISH data for HER2 where available was extracted from the patient medical charts. This test is outsourced by the Department of Pathology at AKU to a College of American Pathologists (CAP) accredited laboratory outside the country. Ki-67 was considered high if 20% or more of the cells showed nuclear staining based on St Gallen recommendation(117). We used Ki-67 status (low/high) to

discriminate luminal A and B and used tumour grade as a surrogate for patients with missing Ki-67(117). If reliable Ki-67 measurement is not available, some alternative assessment of tumor proliferation such as grade may be used to distinguish between ‘Luminal A’ and ‘Luminal B (HER2 negative)’(114). The tumour grading was based on the Elston and Ellis modified Scarff-Bloom-Richardson (SBR) grading system which separates breast cancers into three groups (Grades 1, 2, 3)(118). This grading system is based on morphologic assessment of tubule formation, nuclear atypia/pleomorphism and mitotic count /10HPF. Each element is scored from 1-3 with a total maximum possible score of 9 which is equivalent to a Grade 3 tumour. Grading which included the mitotic count was performed on the Olympus microscope with a high-power field diameter of 0.5mm which is equivalent to 0.196mm² as recommended by the recent WHO classification of breast tumours(119, 120).

For EGFR1 and CK5/6; a result was considered positive for any amount of cytoplasmic or membranous staining in any percentage of cells as per the recommendations from the British Columbia study for defining the basal subtype of breast cancer(114).

Molecular subtypes were defined based per previous clinically validated guidelines(115): luminal A: ER⁺ and/or PR⁺, HER2⁻, and low Ki-67/histologic grade (I or II); luminal B-HER2⁺: ER⁺ and/or PR⁺, and HER2⁺; luminal B-high proliferative: ER⁺ and/or PR⁺, HER2⁻, and high Ki-67/histologic grade (III); HER2-enriched: ER⁻, PR⁻, and HER2⁺; and triple-negative (TN): ER⁻, PR⁻, HER2⁻ (**Figure 2**).

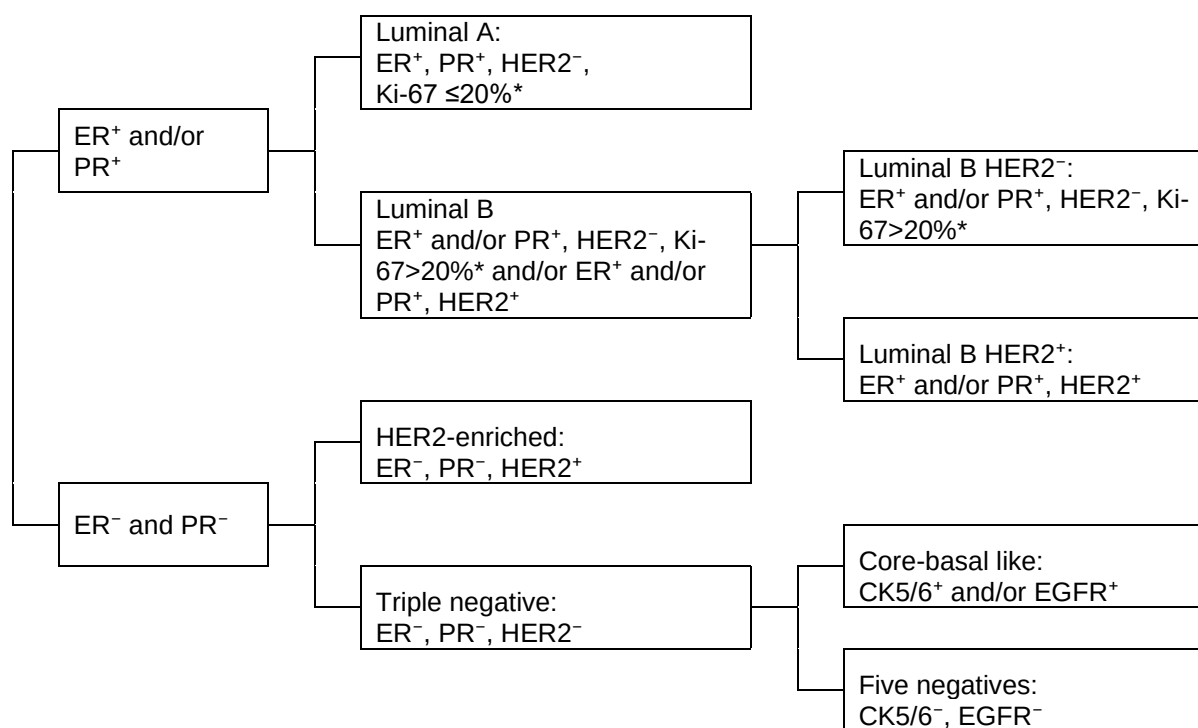


FIGURE 2: BREAST TUMOUR SUBTYPES DEFINED BY EXPRESSION OF ER, PR, HER2, CK5/6, AND EGFR IN KENYAN BREAST CANCER PATIENTS (N=838).

*Tumour grade was used to determine tumour subtypes in the absence of Ki-67: if HER2 is negative and the tumour grade was low or intermediate, define tumour subtype as “luminal A”; if tumour grade is high, define tumour subtype as “luminal B HER2⁻”. CK5/6, cytokeratin 5/6; EGFR, epidermal growth factor receptor; ER, oestrogen receptor; HER2, human epidermal growth factor receptor 2; PR, progesterone receptor.

2.3.2 STATISTICAL ANALYSIS

The 15 hospital/health facilities were classified into 5 hospital groups: Aga Khan University (AKU) hospitals (including AKU hospitals at Kisumu, Mombasa, and Nairobi), Kijabe Mission Hospital, Nyeri Provincial General Hospital (PGH), St Mary's Mission Hospital (Nairobi), and Others. Details of classifying hospital group are shown in **APPENDIX I**. Distributions of breast cancer risk factors, including sociodemographic, reproductive, and tumour pathologic characteristics in the overall study population and by hospital groups were assessed using the Chi-squared Test or Fisher's exact Test. Multivariable logistic regression models were used to estimate relationships of breast cancer risk factors and three common tumour subtypes defined by ER, PR, and HER2, respectively. Multivariable polytomous logistic regression models were used to determine associations between BC risk factors and tumour molecular subtypes (Luminal A as the reference).

Heterogeneity tests were performed to detect any significant differences in strength of associations among the non-Luminal A subtype tumours i.e., Luminal-B, HER2-enriched, and TNBC with each risk factor. All regression models were fully adjusted for the same covariates (except for where noticed): age at diagnosis, BMI, age at menarche, age at first pregnancy, number of children, averaged breastfeeding duration, age at menopause, family history of breast cancer in 1st degree female relatives, highest education level, and occupation. A two-tailed P-value less than 0.05 was considered statistically significant. All analyses were performed with SAS v9.4 statistical software (SAS Institute Inc.).

2.4 RESULTS

2.4.1 Distribution of breast cancer risk factors (sociodemographic and reproductive characteristics) in overall study population and by hospitals (Table 4).

Details of classification of the five health facility groups were presented in the Supplementary Table 1. The classification was made based on the facilities' geographic locations and the frequencies of admitted patients included in our study population. In our hypotheses, the location of health facility might function as a confounder and has an impact on the associations between breast cancer risk factors and breast cancer subtypes. In our Table 1, we observed the distributions of breast cancer risk factor are differed by health facility groups using Chi-square tests, for example BMI, family history of breast cancer in 1st degree female relatives, and age at first pregnancy. Therefore, we adjust our multivariable modeling analyses for the health facility group in addition to other covariates. Unfortunately, all participants in our study are confirmed breast cancer cases. For future study, we may expand our study population by enrolling controls of representativeness.

The details for the population have been previously published (83) All patients were Black Kenyans with the majority being from the Bantu ethnicity. Also in Table 1, we showed significantly different distributions of race, education, occupation, residence area, smoking, and alcohol use across groups of hospital, which support our decision that groups of hospital should be considered as a confounder in analyzing associations with breast cancer risk.

In total, our analysis included 838 invasive breast cancer cases with complete data on ER and PR status. Overall, most patients were diagnosed with breast cancer after 40 years (75%), had BMI ≥ 25 kg/m² (69%), had no family history of breast cancer in their first-degree female relatives (92%), worked as farmers or employed workers/traders (72%), received primary or further education (70%), lived in rural areas (61%), had never consumed alcohol (91%). Our study population was also characteristic for late age at menarche (≥ 13 years, 92%), young age at first pregnancy (< 25 years, 70%), having 3 or more children (68%), high prevalence in breastfeeding (95%) and long breastfeeding duration (≥ 1 year per child, 80%).

Most of the sociodemographic and reproductive characteristics distributed significantly different across five hospitals groups, except for age at diagnosis, age at menarche, menopausal status, and menopausal age. Compared to patients admitted to the other 4 hospital groups, AKU patients were more likely to be overweight or obese (79%), work as employed workers/traders (58%), have tertiary education level (45%), have consumed alcohol (13%), first pregnancy ≥ 25 years (35%), have < 3 children (39% compared to 30%, 25%, and 30% for Kijabe, Nyeri and St Mary's hospitals respectively); and less likely to get exposed to smoke (from cigarette or house cooking), and also to have shorter breastfeeding duration per child. Patients from Kijabe and Nyeri hospitals were more likely to work as farmers (41% and 75%, respectively) and live-in rural areas (75% and 90%, respectively); however, Kijabe patients had shorter breastfeeding duration per child.

TABLE 4: CHARACTERISTICS OF KENYAN BREAST CANCER PATIENTS, OVERALL AND BY HOSPITALS (N=838)

Hospitals													
	Overall (n=838)		AKU (n=350, 42%)		Kijabe (n=105, 13%)		Nyeri (n=110, 13%)		St Mary's (n=122, 15%)		Others (n=151, 18%)		
	N	%	N	%	N	%	N	%	N	%	N	%	P*
Demographic													
Age at diagnosis/year													
20-29	32	3.8	17	4.9	2	1.9	2	1.9	5	4.1	6	4.0	0.11
30-39	177	21.2	64	18.3	23	21.9	22	20.4	34	28.1	34	22.5	
40-49	242	29.0	100	28.7	37	35.2	34	31.5	35	28.9	36	23.8	
50-59	211	25.3	105	30.1	26	24.8	23	21.3	19	15.7	38	25.2	
≥60	172	20.6	63	18.1	17	16.2	27	25.0	28	23.1	37	24.5	
Mean (SD)	49.2 (12.8)		49.0 (11.2)		48.4 (11.3)		50.5 (13.5)		48.7 (15.1)		49.7 (12.9)		
Median (IQR)	48 (39, 57)		49 (40, 56)		47 (41, 57)		49 (41, 60)		45 (37, 58)		49 (39, 59)		
Missing	4		1		0		2		1		0		
BMI/ kg/m2													
Normal (<25.0)	210	31.3	67	21.3	35	40.2	25	39.7	36	40.4	47	40.2	<0.0001
Overweight (25.0 - 29.99)	264	39.3	126	40.0	33	37.9	28	44.4	37	41.6	40	34.2	
Obese (≥30.0)	197	29.4	122	38.7	19	21.8	10	15.9	16	18.0	30	25.6	
Missing	167		35		18		47		33		34		
Family history of breast cancer in first degree female relatives													
No	773	92.2	314	89.7	100	95.2	99	90.0	119	97.5	141	93.4	0.036

Hospitals													
	Overall (n=838)		AKU (n=350, 42%)		Kijabe (n=105, 13%)		Nyeri (n=110, 13%)		St Mary's (n=122, 15%)		Others (n=151, 18%)		
	N	%	N	%	N	%	N	%	N	%	N	%	P*
Yes	65	7.8	36	10.3	5	4.8	11	10.0	3	2.5	10	6.6	
Occupation													
Farmer	253	30.3	47	13.4	43	41.0	82	74.5	39	32.5	42	27.8	<0.0001
Employed worker	195	23.3	138	39.4	16	15.2	5	4.5	10	8.3	26	17.2	
Trader	155	18.5	66	18.9	24	22.9	2	1.8	35	29.2	28	18.5	
Housewife	156	18.7	68	19.4	19	18.1	8	7.3	27	22.5	34	22.5	
Casual worker	32	3.8	11	3.1	3	2.9	4	3.6	8	6.7	6	4.0	
Other	45	5.4	20	5.7	0	0.0	9	8.2	1	0.8	15	9.9	
Missing	2		0		0		0		2		0		
Highest education level													
None	257	30.7	54	15.4	41	39.0	44	40.0	57	46.7	61	40.4	<0.0001
Primary	163	19.5	40	11.4	22	21.0	29	26.4	36	29.5	36	23.8	
Secondary	209	24.9	99	28.3	24	22.9	31	28.2	21	17.2	34	22.5	
Tertiary	209	24.9	157	44.9	18	17.1	6	5.5	8	6.6	20	13.2	
Place of residence													
Rural	511	61.0	173	49.4	79	75.2	99	90.0	65	53.3	95	62.9	<0.0001
Urban	327	39.0	177	50.6	26	24.8	11	10.0	57	46.7	56	37.1	
Exposure to smoking†													
Never exposed	477	56.9	248	70.9	63	60.0	39	35.5	63	51.6	64	42.4	<0.0001
Exposed	361	43.1	102	29.1	42	40.0	71	64.5	59	48.4	87	57.6	
Alcohol consumption													

Hospitals													
	Overall (n=838)		AKU (n=350, 42%)		Kijabe (n=105, 13%)		Nyeri (n=110, 13%)		St Mary's (n=122, 15%)		Others (n=151, 18%)		
	N	%	N	%	N	%	N	%	N	%	N	%	P*
No	760	90.7	304	86.9	98	93.3	109	99.1	115	94.3	134	88.7	0.0009
Yes	78	9.3	46	13.1	7	6.7	1	0.9	7	5.7	17	11.3	
Ethnicity													
Bantu	656	78.3	266	76.0	92	87.6	109	99.1	106	86.9	83	55.0	<0.0001
Nilote	141	16.8	65	18.6	5	4.8	0	0.0	10	8.2	61	40.4	
Cushite/Mixed	41	4.9	19	5.4	8	7.6	1	0.9	6	4.9	7	4.6	
Reproductive													
Age at menarche/year													
9-12	68	8.5	30	8.6	10	11.1	7	6.9	7	5.8	14	10.0	0.40
13-14	340	42.4	150	42.9	31	34.4	39	38.2	61	50.8	59	42.1	
15-20	394	49.1	170	48.6	49	54.4	56	54.9	52	43.3	67	47.9	
Missing	36		0		15		8		2		11		
Age at first pregnancy/year													
Nulliparous‡	37	4.5	23	6.6	4	3.9	2	1.9	2	1.8	6	4.0	<0.0001
<20	216	26.3	62	17.7	29	28.4	34	31.8	38	34.2	53	35.3	
20-24	361	44.0	142	40.6	41	40.2	54	50.5	52	46.8	72	48.0	
25-29	150	18.3	90	25.7	21	20.6	14	13.1	14	12.6	11	7.3	
≥30	56	6.8	33	9.4	7	6.9	3	2.8	5	4.5	8	5.3	
Missing	18		0		3		3		11		1		
Number of children													

Hospitals													
	Overall (n=838)		AKU (n=350, 42%)		Kijabe (n=105, 13%)		Nyeri (n=110, 13%)		St Mary's (n=122, 15%)		Others (n=151, 18%)		
	N	%	N	%	N	%	N	%	N	%	N	%	P*
Nulliparous‡	44	5.3	26	7.4	6	5.7	2	1.8	2	1.6	8	5.3	0.0001
1 or 2	226	27.0	110	31.4	25	23.8	25	22.7	34	27.9	32	21.2	
3 or 4	307	36.6	135	38.6	43	41.0	47	42.7	34	27.9	48	31.8	
≥ 5	261	31.2	79	22.6	31	29.5	36	32.7	52	42.6	63	41.7	
Cumulative breastfeeding duration /month													
Nulliparous‡	37	4.6	23	6.6	4	4.0	2	1.9	2	1.7	6	4.2	0.013 ^a
Never breastfed	7	0.9	3	0.9	2	2.0	0	0.0	0	0.0	2	1.4	
Q1: 1 - <39	191	23.5	82	23.7	27	27.0	23	21.3	32	27.8	27	18.9	
Q2: 39 - <62	192	23.6	100	28.9	28	28.0	24	22.2	16	13.9	24	16.8	
Q3: 62 - <96	182	22.4	80	23.1	17	17.0	28	25.9	24	20.9	33	23.1	
Q4: ≥96	203	25.0	58	16.8	22	22.0	31	28.7	41	35.7	51	35.7	
Missing	26		4		5		2		7		8		
Averaged breastfeeding duration per child/month													
Nulliparous‡	37	4.6	23	6.6	4	4.0	2	1.9	2	1.7	6	4.2	0.0036 ^a
Never breastfed	7	0.9	3	0.9	2	2.0	0	0.0	0	0.0	2	1.4	
<12	120	14.8	53	15.3	23	23.0	8	7.4	18	15.7	18	12.6	
12-23	407	50.1	153	44.2	46	46.0	71	65.7	64	55.7	73	51.0	
≥ 24	241	29.7	114	32.9	25	25.0	27	25.0	31	27.0	44	30.8	
Missing	26		4		5		2		7		8		

Hospitals													
	Overall (n=838)		AKU (n=350, 42%)		Kijabe (n=105, 13%)		Nyeri (n=110, 13%)		St Mary's (n=122, 15%)		Others (n=151, 18%)		
	N	%	N	%	N	%	N	%	N	%	N	%	P*
Number of children & Cumulative breastfeeding duration													
Nulliparous or ≤3 children & <62 months	354	43.60	179	51.7	44	44.0	43	39.8	44	38.3	44	30.8	<0.0001
≤3 children & ≥62 months	79	9.73	43	12.4	5	5.0	10	9.3	8	7.0	13	9.1	
≥4 children & <62 months	73	8.99	29	8.4	17	17.0	6	5.6	6	5.2	15	10.5	
≥4 children & ≥62 months	306	37.69	95	27.5	34	34.0	49	45.4	57	49.6	71	49.7	
Missing	26		4		5		2		7		8		
Age at first pregnancy & Number of children													
Nulliparous	44	5.36	26	7.4	6	5.8	2	1.9	2	1.8	8	5.3	<0.0001
Age 25+ yr, 1-3 births	154	18.76	98	28.0	16	15.5	13	12.1	13	11.7	14	9.3	
Age <25 yr, 1-3 births	244	29.72	101	28.9	27	26.2	38	35.5	37	33.3	41	27.3	
Age 25+ yr, 4+ births	49	5.97	23	6.6	12	11.7	4	3.7	6	5.4	4	2.7	
Age <25 yr, 4+ births	330	40.20	102	29.1	42	40.8	50	46.7	53	47.7	83	55.3	
Missing	17		0		2		3		11		1		
Menopausal status													
Premenopausal	438	52.4	183	52.3	64	61.0	50	45.5	67	55.4	74	49.3	0.18
Postmenopausal	398	47.6	167	47.7	41	39.0	60	54.5	54	44.6	76	50.7	
Missing	2		0		0		0		1		1		
Age at menopause/year ^b													

Hospitals													
	Overall (n=838)		AKU (n=350, 42%)		Kijabe (n=105, 13%)		Nyeri (n=110, 13%)		St Mary's (n=122, 15%)		Others (n=151, 18%)		
	N	%	N	%	N	%	N	%	N	%	N	%	P*
<50	191	57.4	91	55.8	18	62.1	28	59.6	20	71.4	34	51.5	0.45 ^b
≥50	142	42.6	72	44.2	11	37.9	19	40.4	8	28.6	32	48.5	
Missing	65		4		12		13		26		10		
Cumulative hormonal contraception exposure/month													
<48	216	45.9	88	46.3	25	44.6	26	40.0	27	34.6	50	61.0	0.014
≥48	255	54.1	102	53.7	31	55.4	39	60.0	51	65.4	32	39.0	
Missing	367		160		49		45		44		69		

* *P* values were computed from Chi-Square Tests except where noted. *P* values less than 0.05 are shown in bold font. ^a Nulliparous women and parous women who never breastfed were grouped together in Chi-Square Test. ^b Chi-Square Test was performed restricted to postmenopausal women. [†] Only 3.58% (*n*=30) of study participants reported ever having smoked or used smokeless tobacco. Exposure to smoking is summarized here as exposed/never exposed, where exposed is defined as personal use of tobacco as well as exposure to smoke at the workplace or home during child or adulthood. [‡] Nulliparous cases were women who reported never pregnant, never given birth, and had no children (*N*=37, 4.4%). AKU, Aga Khan University; BMI, body mass index; IQR, interquartile range; Q, quartile; SD, standar

Compared to patients admitted to the other 4 hospital groups, AKU patients were more likely to be overweight or obese (79%), have tertiary education level (45%), start the first pregnancy ≥ 25 years (35%), have < 3 children (39%), and have shorter breastfeeding duration per child, which is as expected given that AKU is a private health facility, and compared to the others, patients are generally from a higher socioeconomic status.

2.4.2 Distributions of tumour subtypes and pathologic characteristics in overall study population and by hospitals (Table 4)

The distribution of tumour subtypes defined by IHC markers are summarized in **Figure 3** and **Table 5**. Overall, 69.5%, 59.4 %, and 27.4% of patients were ER positive, PR positive, and HER2 positive, respectively. After classifying patients into molecular subtypes, 34.8%, 35.8%, 10.7%, and 18.6% of patients had luminal A, luminal B, HER2-enriched, and TN breast cancers, respectively. More than 90% of patients had tumour size of 2 cm or larger (2<5 cm, 53.5%; ≥ 5 cm, 38.9%) and had intermediate-to-high tumour grade (intermediate, 45.9%; high, 49.1%). Sixty-one percent of tumours showed lymphovascular invasion (37% in definitive surgery specimens and 24% in core biopsies). Nearly half of patients received definitive surgery, either wide local excision or mastectomy, among which 91% had stage II or higher and for those cases with lymph node metastases (60%), 39.5% were positive for extra nodal extension. AKU patients were more likely to have small (≤ 2 cm) and early-stage tumours ($P < 0.01$). Patients admitted to Kijabe and Nyeri hospitals had higher proportions of tumours with lymphovascular invasion: 71.4% and 69.1%, respectively. In addition, Nyeri patients were less likely to have ER positive tumours. There was no statistical difference in distributions of patient molecular subtypes across hospitals ($P = 0.08$).

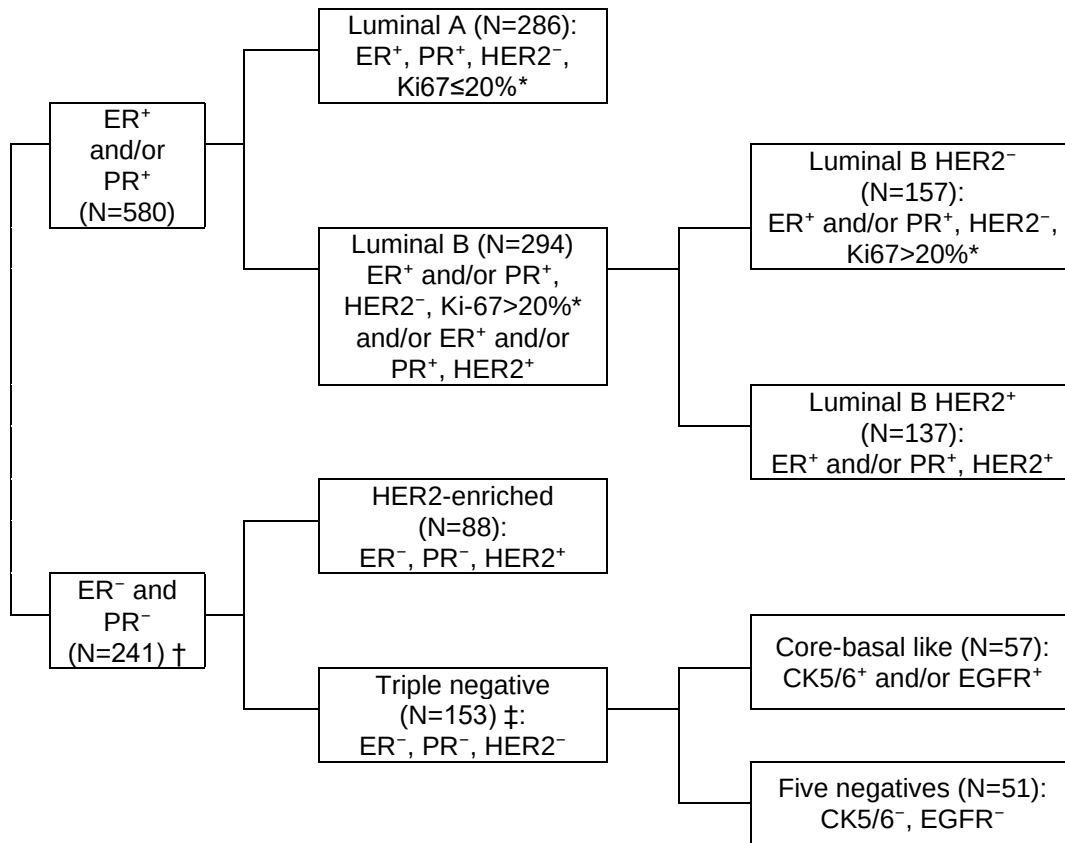


FIGURE 3: BREAST TUMOR SUBTYPE DEFINITION IN KENYAN BREAST CANCER PATIENTS (N=838).

*TUMOR GRADE WAS USED TO DETERMINE TUMOR SUBTYPES IN THE ABSENCE OF KI67: IF TUMOR GRADE IS LOW OR INTERMEDIATE, DEFINE TUMOR SUBTYPE AS "LUMINAL A"; IF TUMOR GRADE IS HIGH, DEFINE TUMOR SUBTYPE AS "LUMINAL B HER2-". †SEVENTEEN CASES ARE NOT INCLUDED DUE TO THEIR MISSING HER2 STATUS. ‡FORTY-FIVE CASES ARE NOT INCLUDED DUE TO THEIR MISSING CK5/6 AND EGFR STATUS. CK5/6: CYTOKERATIN 5/6; EGFR: EPIDERMAL GROWTH FACTOR RECEPTOR; ER: ESTROGEN RECEPTOR; HER2: HUMAN EPIDERMAL GROWTH FACTOR RECEPTOR-2; PR: PROGESTERONE RECEPTOR.

TABLE 5: DISTRIBUTION OF TUMOUR CHARACTERISTICS IN KENYAN BREAST CANCER PATIENTS, OVERALL AND BY FACILITY (N=838)

	Overall (n=838)		Hospitals										
			AKU (n=350, 42%)		Kijabe (n=105, 13%)		Nyeri (n=110, 13%)		St Mary's (n=122, 15%)		Others (n=151, 18%)		
Tumour characteristic	N	%	N	%	N	%	N	%	N	%	N	%	P*
Tumour subtypes													
ER status													
Negative	256	30.6	101	28.9	25	23.8	45	40.9	32	26.2	53	35.1	0.029
Positive	582	69.5	249	71.1	80	76.2	65	59.1	90	73.8	98	64.9	
PR status													
Negative	340	40.6	138	39.4	36	34.3	50	45.5	48	39.3	68	45.0	0.36
Positive	498	59.4	212	60.6	69	65.7	60	54.5	74	60.7	83	55.0	
HER2 status													
Negative	596	72.6	246	71.3	75	72.1	81	76.4	91	74.6	103	71.5	0.84
Positive	225	27.4	99	28.7	29	27.9	25	23.6	31	25.4	41	28.5	
Missing	17		5		1		4		0		7		
Tumour molecular subtype													
Luminal A ^a	286	34.8	121	35.1	46	44.2	37	34.9	40	32.8	42	29.2	0.080
Luminal B	294	35.8	127	36.8	33	31.7	26	24.5	52	42.6	56	38.9	
HER2-enriched	88	10.7	37	10.7	11	10.6	14	13.2	12	9.8	14	9.7	
Triple negative	153	18.6	60	17.4	14	13.5	29	27.4	18	14.8	32	22.2	
Missing	17		5		1		4		0		7		

	Overall (n=838)		Hospitals										
			AKU (n=350, 42%)		Kijabe (n=105, 13%)		Nyeri (n=110, 13%)		St Mary's (n=122, 15%)		Others (n=151, 18%)		
Tumour characteristic	N	%	N	%	N	%	N	%	N	%	N	%	P*
Luminal A ^a	286	36.9	121	36.7	46	48.9	37	37.4	40	33.9	42	31.1	0.094
Luminal B - HER2-	157	20.2	65	19.7	15	16.0	15	15.2	33	28.0	29	21.5	
Luminal B - HER2+	137	17.7	62	18.8	18	19.1	11	11.1	19	16.1	27	20.0	
HER2-enriched	88	11.3	37	11.2	11	11.7	14	14.1	12	10.2	14	10.4	
Core-basal like	57	7.3	24	7.3	2	2.1	13	13.1	7	5.9	11	8.1	
Five negatives	51	6.6	20	6.4	2	2.1	9	9.1	7	5.9	12	8.9	
Missing	62 ^b		20		11		11		4		16		
Tumour pathology													
Surgery													
Core biopsy only	435	51.9	157	44.9	26	24.8	46	41.8	73	59.8	133	88.1	<0.0001
Lumpectomy or mastectomy	403	48.1	193	55.1	79	75.2	64	58.2	49	40.2	18	11.9	
Tumour size (cm)													
<2	41	7.6	29	13.1	8	8.1	2	2.3	1	1.9	1	1.3	<0.0001
2-<5	287	53.5	134	60.4	53	53.5	39	44.8	26	50.0	35	45.5	
≥5	209	38.9	59	26.6	38	38.4	46	52.9	25	48.1	41	53.3	
Missing ^c	301		128		6		23		70		74		
Tumour overall grade													
Grade 1 (low)	35	5.0	20	7.3	4	4.4	6	5.7	2	1.7	3	2.8	0.17
Grade 2 (intermediate)	319	45.9	130	47.3	34	37.8	44	41.5	59	50.9	52	48.2	
Grade 3 (high)	341	49.1	125	45.5	52	57.8	56	52.8	55	47.4	53	49.1	

	Overall (n=838)		Hospitals										
			AKU (n=350, 42%)		Kijabe (n=105, 13%)		Nyeri (n=110, 13%)		St Mary's (n=122, 15%)		Others (n=151, 18%)		
Tumour characteristic	N	%	N	%	N	%	N	%	N	%	N	%	P*
Missing/Not applicable	143		75		15		4		6		43		
Lymphovascular invasion													
No	326	38.9	139	39.7	30	28.6	34	30.9	47	38.5	76	50.3	0.0029
Yes	512	61.1	211	60.3	75	71.4	76	69.1	75	61.5	75	49.7	
Among cases with wide local excision or mastectomy (n=403):													
Tumour stage													
Stage 0, i	30	8.2	21	12.1	2	2.6	6	10.3	1	2.4	0	0.0	0.0040
Stage ii	154	42.3	85	49.1	30	39.5	20	34.5	12	29.3	7	43.8	
Stage iii, iv	180	49.5	67	38.7	44	57.9	32	55.2	28	68.3	9	56.3	
Missing	39		20		3		6		8		2		
Lymph nodes with metastasis													
No	162	40.2	89	46.1	23	29.1	22	34.4	20	40.8	8	44.4	0.96
Yes	241	59.8	104	53.9	56	70.9	42	65.6	29	59.2	10	55.6	
Extranodal extension													
No	244	60.6	132	68.4	39	49.4	35	54.7	26	53.1	12	66.7	0.022
Yes	159	39.5	61	31.6	40	50.6	29	45.3	23	46.9	6	33.3	

** P values were computed from Chi-square test except where noted. P values less than 0.05 are shown in bold font. ^a Seventy-four cases, who had missing data for both ki67 and tumour grade, were grouped into the subcategory "Luminal A" in tumour molecular subtype. ^b Sixty-two cases whose tumour molecular subtype cannot be determined: 17 cases are due to their missing HER2 status; the other 45 cases are due to their missing CK5/6 and EGFR status. ^c Ninety-eight percent of missingness are from cases with core biopsy only. AKU, Aga Khan University; CK5/6, cytokeratin 5/6; EGFR, epidermal growth factor receptor; ER, oestrogen receptor; HER2, human epidermal growth factor receptor 2; PR, progesterone receptor*

2.4.3 Association between breast cancer risk factors and receptor status ER, PR, and HER2 (Table 6).

Results of adjusted associations between risk factors and ER, PR, and HER2 status are shown in **Table 6**. Compared to ER positive patients, ER negative patients were more likely to have higher parity (OR=2.03, 95% CI= (1.11, 3.72), $P_{trend}=0.021$, comparing ≥ 5 to < 2 children) and to have the first child at younger ages (OR trend=0.77, 95% CI trend= (0.61, 0.98), $P_{trend}=0.032$, comparing older (50 years of age and above) to younger age (less than 50 years). We categorised parity as nulliparous, 1-2, 3-4, and ≥ 5 children. We considered the characteristic of our study population as well as the cut points used in prior studies. More than 67% of our study population had ≥ 3 children, which enables us to have enough power to investigate associations between higher parity with breast cancer subtypes. And we did observe significantly higher likelihood with breast cancer of ER negative among women with ≥ 5 children but not among women with 3-4 children from the multivariable modelling analyses. With a median cut point, such information will be lost. In addition, the cut points also benefit us in comparing our results to those reported in another publication. Reproductive factors and risk of breast cancer by tumor subtypes among Ghanaian women: A population-based case-control study (41), which used the same cut points for parity in their study population. ER negative and PR negative patients were also more likely to have longer cumulative breastfeeding duration (ER negative: OR=2.38, 95% CI= (1.33, 4.24); PR negative: OR=1.95, 95% CI= (1.16, 3.29), comparing ≥ 62 to < 39 months); however, these positive associations became insignificant after adjusting for numbers of children. BMI, either overall or by menopausal status, did not significantly vary by ER or PR status. Compared to HER2- patients, HER2+ patients were less likely to be obese (OR=0.58, 95% CI= (0.34, 0.97), $P_{trend}=0.038$), especially among postmenopausal women (OR=0.26, 95% CI= (0.10, 0.62), $P_{trend}=0.0026$), and to have positive history of breast cancer in first-degree relatives (OR=0.39, 95% CI= (0.16, 0.93), $p=0.033$), but they were more likely to have consumed alcohol (OR=2.49, 95% CI= (1.35, 4.61), $p=0.0036$).

To determine whether patients admitted to different hospital could modify associations between BC risk factors and tumour subtypes, we selected four key risk factors (i.e., BMI, age at first pregnancy, number of children, and cumulative breastfeeding duration) and stratified their association analyses with tumour subtypes by five hospital groups (**APPENDIX II**). Analysis results suggest that BMI was negatively associated with HER2+ tumours among the AKU patients (ORtrend=0.56, 95% CItrend= (0.40, 0.78), $P_{trend}=0.0006$). HER2+ patients who were admitted to AKU hospitals were less likely to be obese (BMI ≥ 30 kg/m²): OR=0.31, 95% CI= (0.16, 0.60), $p=0.0006$. For parity, results suggest a positive trend between cumulative breastfeeding duration and ER negative tumours among Kijabe patients (ORtrend=2.19, 95% CItrend= (1.00, 4.78), $P_{trend}=0.049$) and patients admitted to non-mainstream hospitals in Kenya (ORtrend=1.76, 95% CItrend= (1.01, 3.04), $P_{trend}=0.044$).

Kenya is composed of 42 major tribes which can be broadly categorized into 43 major ethnic groups. Each group is unique in origin, culture, economic activities, and diet. We have previously demonstrated differences in some of the reproductive risk factors with breast cancer subtypes among the major ethnic groups in Kenya(83). Furthermore, Africans have the most genetically diverse gene pool. The extreme

diversity of climate and varying environmental pressures are some of the strong drivers of selection across Africa. Globally, African populations are understudied despite being the most genetically diverse (121).

TABLE 6: ASSOCIATIONS BETWEEN BREAST CANCER RISK FACTORS AND RECEPTOR STATUS IN KENYAN BREAST CANCER PATIENTS (N=838)

	ER⁺ N=582		ER⁻ N=256		ER⁻ vs. ER⁺	
	N	%	N	%	OR (95% CI) †	P†
Age at diagnosis/year						
20-39	155	26.8	54	21.1	1.00 (Ref)	
40-49	166	28.7	76	29.7	1.38 (0.82, 2.32)	0.23
50-59	138	23.9	73	28.5	1.20 (0.51, 2.83)	0.68
≥60	119	20.6	53	20.7	0.84 (0.30, 2.35)	0.74
Trend‡					0.995 (0.72, 1.39)	0.98
BMI/ kg/m2						
Normal (<25.0)	136	29.8	74	34.6	1.00 (Ref)	
Overweight (25.0 - 29.9)	185	40.5	79	36.9	0.77 (0.49, 1.21)	0.26
Obese (≥30.0)	136	29.8	61	28.5	0.84 (0.51, 1.38)	0.49
Trend‡					0.92 (0.71, 1.18)	0.49
Premenopausal: BMI ^a						
Normal (<25.0)	89	34.6	41	39.8	1.00 (Ref)	
Overweight (25.0 - 29.9)	102	39.7	30	29.1	0.58 (0.31, 1.10)	0.09
Obese (≥30.0)	66	25.7	32	31.1	0.91 (0.46, 1.77)	0.77
Trend‡					0.93 (0.66, 1.31)	0.69
Postmenopausal: BMI ^a						
Normal (<25.0)	47	23.6	33	29.7	1.00 (Ref)	
Overweight (25.0 - 29.9)	83	41.7	49	44.1	1.07 (0.55, 2.08)	0.85
Obese (≥30.0)	69	34.7	29	26.1	0.77 (0.37, 1.61)	0.49
Trend‡					0.87 (0.61, 1.26)	0.47

	ER ⁺ N=582		ER ⁻ N=256		ER ⁻ vs. ER ⁺	
	N	%	N	%	OR (95% CI) †	P†
Age at menarche/year						
≤13 (9-13)	141	25.4	57	23.2	1.00 (Ref)	
14	139	25.0	71	28.9	1.57 (0.94, 2.61)	0.083
≥15 (15-20)	276	49.6	118	48.0	1.29 (0.81, 2.04)	0.28
Trend‡					1.10 (0.88, 1.37)	0.41
Age at first pregnancy/year						
Nulliparous ^b or age ≥30	70	12.3	23	9.1	1.14 (0.48, 2.70)	0.76
<20	130	22.9	86	34.1	1.00 (Ref)	
20-24	253	44.5	108	42.9	0.71 (0.44, 1.15)	0.16
25-29	115	20.2	35	13.9	0.58 (0.30, 1.11)	0.10
Trend‡					0.77 (0.61, 0.98)	0.032
Parity						
Nulliparous ^b	37	6.4	7	2.7	0.36 (0.10, 1.32)	0.12
Parous	545	93.6	249	97.3	1.00 (Ref)	
Number of children						
1 - 2	171	31.4	55	22.1	1.00 (Ref)	
3 - 4	217	39.8	90	36.1	1.43 (0.86, 2.36)	0.16
≥ 5	157	28.8	104	41.8	2.03 (1.11, 3.72)	0.021
Trend‡					1.43 (1.05, 1.93)	0.021
Cumulative breastfeeding duration/month						
Q1: 1 - <39	152	28.8	39	16.2	1.00 (Ref)	
Q2: 39 - <62	127	24.1	65	27.0	2.38 (1.33, 4.24)	0.0033

	ER ⁺ N=582		ER ⁻ N=256		ER ⁻ vs. ER ⁺	
	N	%	N	%	OR (95% CI) †	P†
Q3: 62 - <96	124	23.5	58	24.1	1.44 (0.74, 2.80)	0.28
Q4: ≥96	124	23.5	79	32.8	1.58 (0.76, 3.30)	0.22
Trend‡					1.10 (0.87, 1.39)	0.43
Mean breastfeeding duration per child/month						
<12	84	15.9	36	14.9	1.00 (Ref)	
12- 23	277	52.6	130	53.9	1.10 (0.62, 1.94)	0.74
≥ 24	166	31.5	75	31.1	1.25 (0.68, 2.30)	0.48
Trend‡					1.12 (0.83, 1.51)	0.45
Age at first pregnancy & Number of children						
Age 25+ yr, 1-3 births	117	22.0	37	15.0	1.00 (Ref)	
Age <25 yr, 1-3 births	173	32.6	71	28.9	1.20 (0.68, 2.12)	0.53
Age 25+ yr, 4+ births	34	6.4	15	6.1	1.47 (0.60, 3.61)	0.40
Age <25 yr, 4+ births	207	39.0	123	50.0	1.69 (0.93, 3.05)	0.085
Trend‡					1.19 (0.99, 1.43)	0.063
Number of children & Cumulative breastfeeding duration						
Nulliparous or ≤3 children & <62 months	265	47.0	89	35.9	1.00 (Ref)	
≤3 children & ≥62 months	56	9.9	23	9.3	1.09 (0.56, 2.09)	0.81
≥4 children & <62 months	51	9.0	22	8.9	1.52 (0.74, 3.09)	0.25
≥4 children & ≥62 months	192	34.0	114	46.0	1.44 (0.89, 2.34)	0.14
Trend‡					1.14 (0.97, 1.33)	0.12
Menopausal status						

	ER ⁺ N=582		ER ⁻ N=256		ER ⁻ vs. ER ⁺	
	N	%	N	%	OR (95% CI) †	P†
Premenopausal	317	54.6	121	47.5	1.00 (Ref)	
Postmenopausal	264	45.4	134	52.5	1.44 (0.74, 2.81)	0.28
Age at menopause/years						
< 50	117	21.8	74	31.5	1.00 (Ref)	
> 50	102	19.0	40	17.0	0.73 (0.40, 1.36)	0.32

† Point estimates and 95% confidence intervals were from multivariable models, adjusting for the same series of covariates (except where noticed): age at diagnosis, BMI, age at menarche, age at first pregnancy, number of children, mean breastfeeding duration per child, age at menopause, family history of breast cancer in first degree female relative, occupation, education level, and location of facility. Estimates of numbers of children, cumulative and averaged breastfeeding duration, and combined age at first pregnancy and number of children were computed among parous women. ‡ Results were from the trend analysis using the categorical risk factor as a trend. ^aMultivariable modelling analysis without adjusting for age at menopause. ^b Women who reported never pregnant, never gave birth, and had no child were grouped as "Nulliparous" in modelling analyses. ^c Multivariable modelling analysis without adjusting for mean breastfeeding duration per child. ^d Multivariable Modelling analysis was restricted to postmenopausal women. BMI, body mass index; CI, confidence interval; ER, oestrogen receptor; OR odds ratio; Q, quartile.

2.4.4 Associations between breast cancer risk factors and molecular subtypes (Table 7)

Table 7 shows the associations between BC risk factors and molecular subtypes defined by joint receptor status. Compared to luminal A patients, luminal B patients were more likely to have lower parity (patients with 3 or 4 children, OR=0.47, 95% CI= (0.28, 0.79), $p=0.005$; with 5 or more children, OR=0.45, 95% CI= (0.23, 0.87), $p=0.018$, compared to patients with 1 or 2 children). HER2-enriched patients were less likely to be obese (OR=0.36, 95% CI= (0.16, 0.81), $p=0.013$) or to have older age at menopause (OR=0.38, 95% CI= (0.15, 0.997), $p=0.049$). The HER2-BMI association appeared to be stronger among postmenopausal women (OR=0.24, 95% CI= (0.07, 0.81), $p=0.022$) than among premenopausal women. TN patients seemed to have had longer cumulative breastfeeding duration compared to luminal A patients, although only patients who had breastfed for 39 - 62 months reached statistical significance (OR=2.98, 95% CI= (1.33, 6.71), $p=0.008$), and the averaged breastfeeding duration per child did not differ between the two subtypes.

TABLE 7: ASSOCIATIONS BETWEEN BREAST CANCER RISK FACTORS AND TUMOUR MOLECULAR SUBTYPES IN KENYAN BREAST CANCER PATIENTS (N=821*

	Luminal A n=286		Luminal B n=294		Luminal B vs. Luminal A		HER2- enriched n=88		HER2- enriched vs. Luminal A		Triple Negative n=153		Triple Negative vs. Luminal A		P- het‡	Global P-value†
Characteristic	N	%	N	%	OR (95% CI) †	P†	N	%	OR (95% CI) †	P†	N	%	OR (95% CI) †	P†		
Age at diagnosis/year																
< 50	142	50.2	180	61.4	1.00 (Ref)		44	50.0	1.00 (Ref)		77	50.3	1.00 (Ref)		0.24	0.27
≥ 50	141	49.8	113	38.6	0.87 (0.36, 2.06)	0.74	44	50.0	1.84 (0.57, 5.92)	0.30	76	49.7	0.52 (0.20, 1.40)	0.20		
BMI/ kg/m2																
Normal (<25.0)	60	27.3	71	30.5	1.00 (Ref)		29	41.4	1.00 (Ref)		43	32.6	1.00 (Ref)		0.16	0.17
Overweight (25.0 - 29.9)	84	38.2	99	42.5	1.09 (0.66, 1.82)	0.73	25	35.7	0.55 (0.28, 1.11)	0.10	49	37.1	0.91 (0.49, 1.68)	0.75		
Obese (≥30.0)	76	34.5	63	27.0	0.76 (0.43, 1.33)	0.33	16	22.9	0.36 (0.16, 0.81)	0.013	40	30.3	0.89 (0.46, 1.72)	0.73		
Trend									0.59 (0.39, 0.88)	0.011						
Premenopausal: BMI ^a																
Normal (<25.0)	36	32.4	51	34.7	1.00 (Ref)		16	45.7	1.00 (Ref)		23	38.3	1.00 (Ref)		0.35	0.40
Overweight (25.0 - 29.9)	43	38.7	59	40.1	0.98 (0.51, 1.88)	0.95	11	31.4	0.52 (0.19, 1.40)	0.20	16	26.7	0.49 (0.20, 1.21)	0.12		
Obese (≥30.0)	32	28.8	37	25.2	0.78 (0.37, 1.61)	0.49	8	22.9	0.44 (0.14, 1.35)	0.15	21	35.0	0.84 (0.34, 2.07)	0.71		

Postmenopausal: BMI^a

Normal (<25.0)	24	22.0	20	23.5	1.00 (Ref)		13	37.1	1.00 (Ref)		20	27.8	1.00 (Ref)		0.065	0.20
Overweight (25.0 - 29.9)	41	37.6	40	47.1	1.35 (0.57, 3.20)	0.50	14	40.0	0.48 (0.17, 1.40)	0.18	33	45.8	1.70 (0.68, 4.26)	0.25		
Obese (≥30.0)	44	40.4	25	29.4	1.10 (0.43, 2.82)	0.84	8	22.9	0.24 (0.07, 0.81)	0.022	19	26.4	1.28 (0.48, 3.46)	0.62		
Trend									0.47 (0.26, 0.88)	0.018						

Age at menarche/year

≤13 (9-13)	69	25.8	71	24.8	1.00 (Ref)		23	26.4	1.00 (Ref)		33	22.9	1.00 (Ref)		0.95	0.84
14	68	25.4	69	24.1	1.28 (0.73, 2.25)	0.39	26	29.9	1.65 (0.73, 3.74)	0.23	37	25.7	1.43 (0.73, 2.81)	0.30		
≥15 (15-20)	131	48.9	146	51.1	1.29 (0.79, 2.13)	0.31	38	43.7	1.32 (0.64, 2.73)	0.46	74	51.4	1.43 (0.78, 2.60)	0.24		

Age at first pregnancy/year

Nulliparous ^b or ≥30	31	11.3	39	13.4	1.06 (0.40, 2.82)	0.90	10	11.5	1.18 (0.31, 4.53)	0.81	11	7.3	0.66 (0.20, 2.18)	0.50	0.15	0.36
<20	72	26.2	58	19.9	1.00 (Ref)		31	35.6	1.00 (Ref)		51	34.0	1.00 (Ref)			
20-24	121	44.0	133	45.5	1.51 (0.84, 2.70)	0.17	30	34.5	0.64 (0.29, 1.37)	0.25	69	46.0	0.81 (0.43, 1.51)	0.51		
25-29	51	18.5	62	21.2	1.76 (0.85, 3.65)	0.13	16	18.4	0.92 (0.34, 2.51)	0.87	19	12.7	0.67 (0.28, 1.60)	0.37		

Parity

Nulliparous	14	4.9	20	6.8	1.09 (0.30, 3.88)	0.90	4	4.5	0.56 (0.07, 4.77)	0.60	3	2.0	0.67 (0.11, 4.12)	0.66	0.77	0.90
Parous	272	95.1	274	93.2	1.00 (Ref)		84	95.5	1.00 (Ref)		150	98.0	1.00 (Ref)			
Number of children																
1 or 2	73	26.8	101	36.9	1.00 (Ref)		19	22.6	1.00 (Ref)		31	20.7	1.00 (Ref)		0.012	0.0068
3 or 4	113	41.5	101	36.9	0.47 (0.28, 0.79)	0.005	26	31.0	0.85 (0.37, 1.98)	0.71	59	39.3	1.16 (0.59, 2.29)	0.66		
≥ 5	86	31.6	72	26.3	0.45 (0.23, 0.87)	0.018	39	46.4	1.71 (0.65, 4.53)	0.28	60	40.0	1.27 (0.56, 2.89)	0.56		
Trend					0.65 (0.47, 0.90)	0.0099										
Cumulative breastfeeding duration/monthc																
Q1: 1 - <39	64	24.3	88	33.2	1.00 (Ref)		15	18.3	1.00 (Ref)		20	13.9	1.00 (Ref)		0.090	0.069
Q2: 39 - <62	59	22.4	68	25.7	0.99 (0.54, 1.80)	0.97	22	26.8	2.48 (0.97, 6.33)	0.058	40	27.8	2.98 (1.33, 6.71)	0.008		
Q3: 62 - <96	71	27.0	53	20.0	0.66 (0.33, 1.30)	0.23	14	17.1	0.80 (0.27, 2.39)	0.69	40	27.8	1.80 (0.74, 4.38)	0.20		
Q4: 96 - 720	69	26.2	56	21.1	0.99 (0.44, 2.20)	0.97	31	37.8	1.30 (0.40, 4.16)	0.66	44	30.6	1.87 (0.69, 5.07)	0.22		
Trend					0.93 (0.71, 1.20)	0.56			0.97 (0.67, 1.41)	0.88			1.12 (0.82, 1.52)	0.47		

Continuous duration					0.99 (0.99, 1.00)	0.07						0.998 (0.99, 1.01)	0.56						0.999 (0.99, 1.00)	0.72		
Mean breastfeeding duration per child/monthd																						
<12	34	12.9	48	18.1	1.00 (Ref)		11	13.4	1.00 (Ref)		23	16.0	1.00 (Ref)		0.38	0.52						
12- 23	142	54.0	139	52.5	0.65 (0.35, 1.21)	0.18	49	59.8	1.24 (0.47, 3.29)	0.66	71	49.3	0.75 (0.36, 1.59)		0.46							
≥ 24	87	33.1	78	29.4	0.64 (0.32, 1.27)	0.20	22	26.8	1.21 (0.42, 3.52)	0.73	50	34.7	1.09 (0.49, 2.43)		0.83							
Trend					0.84 (0.60, 1.16)	0.28					1.07 (0.66, 1.72)	0.80					1.13 (0.76, 1.66)	0.55				
Continuous duration					0.98 (0.96, 1.01)	0.30						1.00 (0.96, 1.05)	0.92						1.01 (0.97, 1.05)	0.59		
Age at first pregnancy & Number of children																						
Age 25+ yr, 1-3 births	54	20.7	64	23.5	1.00 (Ref)		15	17.9	1.00 (Ref)		20	13.6	1.00 (Ref)		0.058	0.21						
Age <25 yr, 1-3 births	80	30.7	93	34.2	0.93 (0.52, 1.67)	0.81	19	22.6	0.82 (0.31, 2.15)	0.69	47	32.0	1.39 (0.65, 2.97)		0.40							
Age 25+ yr, 4+ births	15	5.7	18	6.6	0.77 (0.29, 2.05)	0.60	8	9.5	2.10 (0.58, 7.54)	0.26	7	4.8	1.12 (0.29, 4.27)		0.87							
Age <25 yr, 4+ births	112	42.9	97	35.7	0.55 (0.29, 1.05)	0.068	42	50.0	1.31 (0.51, 3.39)	0.58	73	49.7	1.38 (0.62, 3.08)		0.43							

Trend					0.81 (0.66, 0.99)	0.041					0.03					
Number of children & Cumulative breastfeeding duration																
Nulliparous or ≤3 children and breastfeeding <62 months	117	42.2	146	51.2	1.00 (Ref)		31	36.0	1.00 (Ref)		51	34.7	1.00 (Ref)		0.067	0.11
≤3 children and breastfeeding ≥62 months	30	10.8	25	8.8	0.61 (0.30, 1.22)	0.16	6	7.0	0.41 (0.11, 1.56)	0.19	17	11.6	1.19 (0.53, 2.66)		0.68	
≥4 children and breastfeeding <62 months	20	7.2	30	10.5	0.89 (0.38, 2.07)	0.79	10	11.6	2.34 (0.78, 6.99)	0.13	12	8.2	1.31 (0.49, 3.53)		0.59	
≥4 children and breastfeeding ≥62 months	110	39.7	84	29.5	0.52 (0.30, 0.89)	0.02	39	45.3	1.26 (0.57, 2.78)	0.56	67	45.6	0.96 (0.50, 1.84)		0.91	
Trend					0.81 (0.68, 0.97)	0.02										
Menopausal status ^a																
Premenopausal	144	50.3	174	59.4	1.00 (Ref)		44	50.0	1.00 (Ref)		68	44.7	1.00 (Ref)		0.088	0.14
Postmenopausal	142	49.7	119	40.6	0.69 (0.32, 1.48)	0.33	44	50.0	0.69 (0.22, 2.10)	0.51	84	55.3	1.82 (0.78, 4.28)		0.17	
Age at menopause/years																
Premenopausal	144	50.3	174	59.4	(Results not shown)		44	50.0	(Results not shown)		68	44.7	(Results not shown)			
< 50 years	58	22.5	58	20.9	1.00 (Ref)		28	33.7	1.00 (Ref)		44	32.1	1.00 (Ref)			

> 50 years	56	21.7	46	16.6	0.88 (0.45, 1.74)	0.72	11	13.3	0.38 (0.15, 0.997)	0.049	25	18.3	0.66 (0.31, 1.40)	0.28
Family history of breast cancer in first degree female relatives														
No	267	93.4	269	91.5	1.00 (Ref)		85	96.6	1.00 (Ref)		137	89.5	1.00 (Ref)	0.34 0.49
Yes	19	6.6	25	8.5	1.22 (0.55, 2.70)	0.62	3	3.4	0.51 (0.13, 1.96)	0.33	16	10.5	1.41 (0.61, 3.28)	0.42
Alcohol Use														
No	261	91.3	258	87.8	1.00 (Ref)		83	94.3	1.00 (Ref)		143	93.5	1.00 (Ref)	0.29 0.24
Yes	25	8.7	36	12.2	1.76 (0.89, 3.49)	0.10	5	5.7	1.37 (0.45, 4.19)	0.58	10	6.5	0.81 (0.31, 2.15)	0.68

*Seventeen cases were excluded from analyses because of their missing data for HER2 status.

† Point estimates and 95% confidence intervals were from multivariable models, adjusting for the same series of covariates (except where noticed): age at diagnosis, BMI, age at menarche, age at first pregnancy, number of children, mean breastfeeding duration per child, age at menopause, family history of breast cancer in first degree female relative, occupation, education level, and location of facility. Estimates of numbers of children, cumulative and mean breastfeeding duration, and combined age at first pregnancy and number of children were computed among parous women.

‡ P values from Heterogeneity Test comparing risks of tumour luminal subtypes other than Luminal A.

^a Multivariable modelling analysis without adjusting for age at menopause.

^b Women who reported never pregnant, never gave birth, and had no child were grouped as "Nulliparous" in modelling analyses.

^c Multivariable modelling analysis without adjusting for mean breastfeeding duration per child.

^d Multivariable modelling analysis without adjusting for cumulative breastfeeding duration.

^e Multivariable modelling analysis without adjusting for menopausal status.

BMI, body mass index; CI, confidence interval; HER2, human epidermal growth factor receptor 2; OR odds ratio; Q, quartile

2.5. DISCUSSION

Although there has been considerable research to identify risk factors and biomarkers with resultant development of some risk stratification models associated with breast cancer, there is limited data from Africa where some of the breast cancer associated risk/protective factors such as parity and breastfeeding have extremely different distributions.

The overall risk factor distribution for BC in our study is similar to a large case control study from Ghana(41).Among BC patients in Ghana and Kenya, >60% of women had ≥ 3 children, >80% women had the first child before age 25 years, and >90% women had breastfed with the average breastfeeding duration per child near two years however in contrast for the BC subtypes, the frequency of ER negative cancers was 30.5% for Kenya and 50% for Ghana while TNBC frequency was 18.6% for Kenya which is lower compared to 28% observed in BC patients from Ghana (ER negative: 50%; TNBC: 28%).suggesting that these differences may be driven by distinct genetic factors and environmental exposures which may contribute to ethnic heterogeneity.

In this case series, we studied the associations of established risk factors for breast cancer with molecular subtypes among 838 Kenyan women. Our findings show that BMI, breastfeeding duration, parity, and alcohol consumption varied by molecular subtypes as defined by IHC markers. Specifically, women with higher BMI were more likely to develop luminal subtype of breast cancer while increased duration of breast feeding and a higher parity were more likely observed with the triple negative subtype. Alcohol consumption was more likely to be associated with HER2 breast cancer subtype. Risk factors such as patient age at diagnosis, and other reproductive risk factors that included age at menarche, age at onset of menopause and first full-term pregnancy were not different among the molecular subtypes.

Distributions of sociodemographic and reproductive characteristics in this analysis are similar to what we have previously described (52) but strikingly different to those from other parts of Africa and among African-Americans. The median age at diagnosis was 48.9 years, which is similar to findings from Tanzania in East Africa (122) and Guinea (123) and Ghana (25) in West Africa. To investigate the role of TILs in the survival of African American TNBC patients, a study examined 60 patients belonging to this population subgroup and found a positive association between TILs and survival benefits [1-10% Hazard Ratio (HR)=0.32, 95% Confidence Interval (CI) 0.12-0.90, $p=0.031$; >10% HR=0.18, 95% CI 0.05-0.67, 9.9×10^{-3}] (84). An association was also identified between 524 transcripts (326 coding, 198 non-coding) and TIL levels, among which, 34 transcripts were correlated to both TILs and survival ($p<0.05$) (185). In contrast, an urban based clinicopathologic survey from South Africa, reported a higher median age of 56 years for women with breast cancer. However, it is to be noted that the South African study did not analyse the differences in age at presentation separately for the Black, White, Indian (Asian), and Coloured populations which may explain the difference as well as the adoption of a more westernized lifestyle in South Africa compared with the rest of Africa(124). Differences in reproductive factors between indigenous Africans and African Americans have also been described. As an example, the majority of the women in our study tended to be parous, have prolonged breastfeeding and cumulative breast feeding periods which contrasts with African-Americans who are more likely to have higher parity but less likely to breastfeed and have higher prevalence of obesity(125).

The distribution of molecular subtypes in our series is similar to Abubakar(126) et al involving 3012 breast cancer patients from Malaysia (Luminal A 34.8% and 34 %, Luminal B 35.8% and 33%, HER2 enriched 10.7% and 13%, TNBC 18.6% and 20% respectively) but different from that observed from a population-based case–control study in Ghana. The data from 1,126 cases (aged 18-74 years) with invasive breast cancer, showed luminal A breast cancer to be the most common subtype (49%) followed by triple-negative/basal-like (28%), HER2-enriched (15%) and luminal B breast cancers (8%)(41). In another study of 114 breast cancer patients from Ethiopia(127), a neighbouring country of Kenya, the distribution of molecular subtypes was: Luminal A: 40%; Luminal B: 26%; HER2-enriched: 10%; TNBC: 23%. Although in the latter study biomarker status was determined on digitised tissue microarray sections, the definition of the subtypes was similar to ours. However, in the Ghana study a cut off value of 10% was used to determine ER and PR positivity. Also, all HER2 negative and HER2 equivocal cases were negative for HER2. Notably, less than 2% of cases were scored 1-9% for ER. The heterogeneity in subtype distribution may also be attributable to geographic variations.

The median age at diagnosis did not differ significantly among the four molecular subtypes of breast cancer in our case series which is different from the study in Ethiopia(127) whose results showed an earlier onset of luminal B compared to the other molecular subtypes. Also, the study reported that frequencies of the different molecular subtypes were significantly different in different age ranges with an increase in the incidence of luminal A in women 50 years of age and above. The study, however, was small which could account for the difference seen.

With regards to age and molecular subtype distribution our findings are similar to the results of a case only analysis from Ghana (41) which did not show any difference in age between 390 and 386 ER positive and ER negative cases respectively, however these observations are different from the Africa American Breast Cancer Epidemiology and risk (AMBER) consortium study with significantly more frequent basal-like and HER2 enriched cancers in women less than 50 years across all participating study sites(125). These differences in frequency of subtypes related to age may be attributable to genetic heterogeneity, environmental exposures, and similarities in the prevalence of reproductive risk factors across sub-Saharan Africa.

Obesity is a known risk factor for breast cancer in post-menopausal women but protective for breast cancer in pre-menopausal women. Obesity can disrupt some biological pathways, including, insulin resistance, and synthesis of endogenous sex hormones(57), (58). The majority of the women with BC in our study were overweight or obese which contrasts with the Asian population; studies from both China(128) and Malaysia(126) show that in general, the prevalence of obesity among Asian women with breast cancer is less than that which has been reported from the West among African-American(129) and Caucasian (130, 131) patients with breast cancer. Surprisingly, increased BMI was significantly associated with lower breast cancer risk in both pre- and post-menopausal women in a hospital-based case-control study of 7,663 Malaysian women (58). The variations in obesity and breast cancer among different racial/ethnic populations may underscore the need for a combination of other obesity parameters that can serve as more robust markers of breast cancer risk.

In a secondary analysis of the Women's Health Initiative (WHI) clinical trial and observational study, postmenopausal women with normal BMI, relatively high body fat levels were associated with an elevated risk of invasive breast cancer (94).

The study found a 56% increase in the risk of developing ER positive breast cancer per 5-kg increase in trunk fat, despite a normal BMI. Furthermore, a persistent association between higher trunk fat and increased risk of ER positive breast cancer in women who remained in the normal BMI range during the follow-up period was also observed(94).

When we examined the association of obesity with molecular subtypes we found that HER2 enriched patients were less likely to have a high BMI similar to a study conducted in Malaysia, which showed that women with HER2-enriched and triple-negative tumours were significantly less likely to be obese than those with the luminal A subtype(126). In contrast, in the AMBER consortium study, it was noted that a BMI ≥ 30 was associated with reduced risk of ER positive breast cancer but not of ER negative or triple negative BC(129). The association we observed is unlikely to be entirely driven by weight loss due to cancer diagnosis since BMI did not vary by tumour stage in our study.

Albeit small, we showed some association with increased alcohol consumption and HER2-enriched breast cancers which warrants further investigation into possible unique risk factors associated with this subtype. Current evidence suggests that among other known pathogenetic pathways, including folate depletion, alcohol mediates its carcinogenic effect through DNA methylation resulting in DNA instability inducing carcinogenesis(132). Contrary to our findings, a large case-case study of breast cancer subtypes from the Surveillance, Epidemiology and End Results (SEER) cancer registries, showed that the risk of HER2 positive BC was decreased by over 25% relative to ER positive cancers in current alcohol users(133). Studies in mouse models however, indicate that the effect of alcohol induced HER2 breast cancers is dependent upon the presence of ovarian oestrogen and showed that oophorectomy blocked alcohol induced HER2 mammary gland tumour development(134).

Regarding reproductive factors, we did not find any association between age at menarche, age at menopause and age at first pregnancy with molecular subtypes. Studies around these risk factors have shown inconsistent results with some large case studies reporting the proportion of women ≤ 13 years at the age of menarche being significantly lower than those who were ≥ 16 years among luminal A cancers than the basal-like cases(135), while, similar to our findings, in the Ghana study (41) analyses for all cases combined did not show statistically significant differences for ER negative compared to ER positive cases by age at menarche, however there was a suggestion of a positive association with older ages at menarche for ER positive breast tumours.

Parity has been reported to have dual effect on breast cancer risk; it is protective for ER positive women while increases risk for ER negative women especially among younger women(125, 136). Despite the heterogeneity in parity related exposures, the differential effect of parity by ER has been consistently reported across different populations(41, 125, 126, 137).

Although we were not able to compare relative risks associated with parity in different molecular subtypes due to the case-only design, our result of higher number of children in ER negative than in ER positive patients is consistent with results from previous case-control studies. Taking advantage of the extremely high prevalence of high parity among patients in Kenya, we observed that the association of parity with ER followed a dose-dependent manner, with the highest variation by ER observed among women with five or more children. Similarly, in a population where the vast majority of women had their first

children before 30 years old, we found the similar association between younger age at first birth and ER negative breast cancer as reported in previous studies conducted in different populations (126, 138, 139) suggesting that the associations between parity/age at first birth and ER subtypes are quite robust. Interestingly, in our study, we also found that luminal B patients, both luminal B/HER2- and luminal B/HER2+, had less children as compared to luminal A patients. These results are in line with data from the Nurse Health Study reporting greater reduced risks associated with parity, especially among women who never breastfed, in luminal B than luminal A patients (137) suggesting that parity may have a stronger protective effect for luminal B as compared to luminal A patients. However, using data based on a Malaysian case-series, we found that luminal B patients were more likely to be parous and to have breastfed compared to luminal A patients (126). These inconsistent results warrant further investigations especially in diverse populations.

Investigations of associations between breastfeeding and breast cancer risk by receptor status have resulted in inconsistent findings, with some showing similar protective effect for all subtypes (140) and others showing a stronger protection against ER negative especially TNBC (141). In the Ghana study, the heterogeneity by ER was only evident in women <50 years, where the reduced risk associated with extensive breastfeeding was only seen in ER negative but not in ER positive patients; while in older women, extended breastfeeding showed an inverse association regardless of ER status yet a stronger association for ER positive patients (41). In contrast, we did not observe significant differences of breastfeeding by ER or by intrinsic subtype, either in all women or by age. TNBC patients were more likely to have longer cumulative breastfeeding duration compared to luminal A patients, however, this is more likely due to the higher parity among TNBC patients. The average duration of breastfeeding per child did not vary significantly by ER or subtype. The inconsistent findings even within African populations with similar extensive breastfeeding highlight the complexity of subtype-specific risk associated with breastfeeding and the importance of conducting large molecular epidemiologic studies in diverse African populations.

There are conflicting reports regarding the association of parity and breast feeding with breast cancer molecular subtypes. Our finding of increased parity and younger age at first full term pregnancy (FFTP) in ER negative BC relative to ER positive BC is consistent with most studies. We observed that TNBC patients had longer cumulative breastfeeding duration compared to luminal A patients, although only patients who had breastfed for a cumulative 39 - 62 months reached statistical significance (OR=2.98, 95% CI= (1.33, 6.71), p=0.008), however, this is more likely due to the higher parity among TNBC patients. Of note, the mean breastfeeding duration per child did not differ between the two subtypes.

Results from the Breast Cancer Association Consortium comprising of 34 studies with over 11000 women with BC reproductive risk factor and ER/PR/HER2 data (142) showed that relative to ER positive patients, women with 4 or more children were more likely to develop TNBC (10.3% versus 11.4% respectively), and this did not vary by age although a slight increase of TNBC was observed at 55 years of age; a weak nonlinear relationship was also observed for parity with age and HER2 enriched BC with an increase in this subtype noted at age 35 and 55 years (OR 1.38 (0.86–2.20) and 1 (0.98–2.47) respectively). Furthermore, FFTP at 25 years versus 20 years was associated with a reduced likelihood of TNBC (OR 0.78 (0.70–0.88) <0.0001).

In the case control study from Ghana(41), comprising 1,126 BC cases and 2,106 controls recruited between 2013 and 2015, from the case only analysis among 378 women <50 years and 401 women ≥50 years of age, it was observed that increasing parity was associated with TNBC in women <50 years of age but not among those ≥50 years. However, the risk of luminal A subtype was inversely related to parity irrespective of age. Regarding breast feeding duration, the authors observed that in women <50 years with TNBC, median months of breastfeeding of ≥13 months per pregnancy was negatively associated with risk of TNBC, but this relationship was not observed among women <50 years of age with luminal A tumours. In contrast, in women ≥50 years of age, there was no association noted with TNBC and extended breast-feeding duration but a strong negative association with luminal A tumours (OR 0.52, 95% CI 0.33–0.82). Similarly, in our study where the vast majority of women had their first children before 30 years old, we found similar association between younger age at first birth and ER negative breast cancer as reported in previous studies conducted in different populations (126, 138, 139) suggesting that the associations between parity/age at first birth and ER subtypes are quite robust.

Interestingly, in our study, we also found that luminal B patients, both luminal B/HER2- and luminal B/HER2+, had less children as compared to luminal A patients. Our findings are similar to the Nurse Health Study(137) suggesting that parity may have a stronger protective effect for luminal B as compared to luminal A patients.

In the study by Abubakar et al(126), it was observed that patients with TNBC and HER2-enriched subtypes were more likely to have ever breastfed with a higher frequency of more than 10 months of breast feeding even though this was not significant. Our study findings and that of Abubakar et al (126) are supported in a study of 1041 breast cancer cases from Mexico which reported more than twice higher odds of TNBC than luminal A tumour in women who breastfed for more than 12 months(143). Palmer and colleagues(125), consolidated four studies enriched for African-American BC patients in the AMBER consortium, and investigated the relation of parity and lactation to risk of specific breast cancer subtypes. In contrast to our findings, they found no evidence of an association among women who had breastfed and TNBC whereas there were one and a half-increased odds of TNBC in parous women with four or more births who had never breastfed. The inconsistent findings even within African populations with similar extensive breastfeeding highlight the complexity of subtype-specific risk associated with breastfeeding and the importance of conducting large molecular epidemiologic studies in diverse African populations. The strength of our study includes representation of BC cases from multiple hospitals in Kenya, well annotated risk factor questionnaire and clinical data, and centralized high-quality biomarker assessment.

This study was limited by the retrospective collection of data and possible reverse causation, as well as the case-only design, which prohibited us from estimating relative risks associated with each risk factor. Further, despite being the largest BC study of this type conducted in Kenya, the sample size was still relatively small to evaluate risk factors in rare tumour subtypes, especially in age stratified analyses.

In summary, our study findings, based on data from an indigenous African population with unique risk factor profiles, add to the growing body of knowledge regarding the aetiological heterogeneity of breast cancer molecular subtypes among geographically diverse ethnic groups. Further investigations of genetic and environmental factors that modify breast cancer risk in African populations are recommended. Inclusion of diverse regional population groups from sub-Saharan Africa in global breast cancer studies may help provide

a better understanding of the subtype-specific breast cancer risk aetiology, which will be critical for the development of risk prediction models in African populations.

Chapter 3

3.0: TUMOUR INFILTRATING LYMPHOCYTES AND OUTCOMES IN BREAST CANCER SUBTYPES

3.1 Background

3.1.1 Normal Wound Healing

3.1.2 Wound healing as an example of physiological inflammation

Wound healing has three principal phases: the initial inflammatory phase of wound healing begins with haemostasis. Following the initial platelet haemostatic plug and release of TGF- β 1 and PDGF, formation of granulation tissue is facilitated by chemotaxis of neutrophils, monocytes, fibroblasts and myofibroblasts, as well as by synthesis of new extracellular matrix (ECM) and neoangiogenesis(144, 145).

Macrophages play a critical role in the inflammatory phase. Deployment of monocytes/macrophages to the site of injury peaks as the number of neutrophils decline. Once present, however, they differentiate into mature macrophages or immature dendritic cells. They help to debride the wound through phagocytosis, but, more importantly, they are the primary source of proinflammatory cytokines and growth factors such as the interleukins (IL-1, IL-6, IL-8), PDGF, TGF- β , epidermal growth factor (EGF), fibroblast growth factor (FGF), vascular endothelial growth factor (VEGF), and insulin-like growth factor (IGF)(144, 145). These humoral factors stimulate the recruitment, activation, and proliferation of additional macrophages, lymphocytes, fibroblasts, and endothelial cells. These cytokines also act in an autocrine fashion to tremendously amplify their expression(144, 145).

During the remodelling phase which begins at about 2 to 3 weeks, fibroblasts organize and cross-link the collagen, wound strength gradually increases, and wound contraction occurs. Immunosuppressive factors such as infection, malnutrition can influence the duration of each phase of the wound healing process(144, 145).

3.1.3 Role of Inflammation in Tumours

The immune system determines tumour immunogenicity as is evident by the fact that tumours arising in RAG2-deficient mice (lacking T, B and NKT cells) are, as a group, more immunogenic than those derived from immunocompetent hosts(12, 146). Furthermore, inflammatory cells and soluble factors are present in all tumours(147). The signs of “smouldering” inflammation which include tissue remodelling, angiogenesis, and other wound healing-like features, are commonly used by pathologists as morphological cues of invasive cancer(147). Recent evidence demonstrates that these stromal processes play a fundamental role in cancer development and progression, and, at least in some cases, may predict the clinical behaviour of a cancer better than the characteristics of the neoplastic cells themselves(148).

Furthermore, inflammation and inflammatory conditions cause inflammatory cells, cytokines, chemokines, and enzymes, to establish an inflammatory microenvironment which may initiate and promote tumour development(12).

Although the inflammatory host response may suppress tumour growth, it can also promote cancer development via multiple signalling pathways that include cell adhesion pathways(149). Inflammation provokes a marked increase in normal cell turnover and this effect coupled with the capacity of activated phagocytes to induce DNA damage might eventually culminate in neoplasia(150). Epidemiological studies delineate a close association between chronic inflammatory disease and cancer(151). The aberrant production of cytokines or defective cytokine signalling may lead to alterations in cell growth, differentiation and apoptosis. The dysregulated expression of chemokine receptors allows tumour cells to subvert the normal cues for cell migration, fostering metastasis(152). Immune cell populations may also facilitate tumour development. As an example, CD4⁺ and CD8⁺ T lymphocytes are required for the development of hepatocellular carcinoma in mice transgenic for hepatitis B virus(153). It has also been shown that nonsteroidal anti-inflammatory agents substantively reduce tumour formation in mice and patients with polyposis syndromes, although the precise mechanisms are unclear(154).

Although, it is well known that certain chronic inflammatory diseases are associated with cancers, and that chronic inflammation and infection contributes to about 25% of all human cancers(147), the critical role of inflammatory cells in cancers that cannot be linked to a pre-existing inflammatory condition has been recognized only recently(155).

Oncogenes target directly or indirectly pro-inflammatory pathways. For example, Ras activates the transcription of the inflammatory cytokine interleukin-8 (IL-8), whereas c-myc and bcl-2 inhibit apoptosis leading to necrotic tumour cell death and release of damage associated molecular pattern molecules or DAMPs(156).

In both circumstances, the resulting host response is inflammation that promotes tumour invasion and growth(157).

Chronic inflammation in the tumour microenvironment and the resulting tumour evasion of the immune system have recently been recognized as another hallmark of cancer(158).

As a matter of fact, the role of the innate and adaptive immune system in tumour development and progression has also been an area of recent research interest. In colorectal carcinoma for example, the presence (or absence) of CD8⁺ T lymphocytes within tumour-cell nests is equivalent to the widely used Duke's staging system for assessing prognosis(159). Similarly, enhanced proliferative activity of intratumoural CD8⁺ T cells in renal-cell carcinoma is associated with improved patient outcome(160).

Taken together, these correlative studies underscore the idea that anti-tumour T-cell responses may be clinically important(161).

A major component of the inflammatory milieu in solid tumours are macrophages, the biological role and possible clinical significance of which remains unclear. Macrophages have been shown to have both protumour (promotion of angiogenesis and secretion of growth factors) and antitumor effects (direct anti-tumour cytotoxicity and presentation of tumour associated antigens(162).

3.1.3.1 *Inflammation and Breast Cancer*

Although pre-existing inflammation and infection are not considered a risk factor for breast cancer development, it is generally accepted that infiltration of immunosuppressive leukocytes and accompanying chronic inflammation during tumour progression promote breast cancer growth(10, 163). As an example, the role of HPV infection as a risk factor for breast cancer is controversial with variable expression of the virus in breast tissue from none to 86% (164). A recent review describes intrinsic (though an endogenous primary source such as cervix) and extrinsic pathways (direct infection through skin to skin contact) by which the virus can infect breast tissue (165). Viral infection of breast tissue is thought to be mediated via genomic instability caused by interference with the expression of BRCA 1 and 2 genes, increase in highly reactive oxygen species, reduction in levels and activity of the tumour immunosuppressive p53 protein, increase in inflammatory cytokines within the tumour microenvironment specifically IL-6, IL-1, TGF- β , TNF- α and IL-17, and augmentation of the alternate metabolic pathway of breast cancer cells. Additionally, compared to wild-type breast cancer cells, HPV has been shown to dampen immunity as evident by an increase in CD4+CD25+FOXP3+ T-regulatory cells (Tregs) and a decrease in natural killer cells(165).

Most current understanding of breast cancer among African women derives from studies in AAs, however the genetic structure of AAs demonstrates a mixture of African and non-African ancestry(166). In addition, indigenous African populations live in a diverse range of environments, have high exposures to infectious agents (examples include malaria and other parasites), diverse diets, and a low prevalence of traditional breast cancer risk factors (including low or late parity, lack of breastfeeding, obesity, and exogenous hormone use)(43, 116, 167). The large number of genetic variants that are common in African populations may result from the adaptation to these unique exposures through natural selection. For example, a recent study found that pathways that are enriched for signals of adaptation tend to be involved in immune response and wound healing(168), suggesting that African breast cancer may have distinct immune profiles.

Recent TCGA (The Cancer Genome Atlas), Pan-Cancer studies identified substantial heterogeneity in immune profiles across and within cancer types as well as within cancer subtypes(169). The presence of infiltrating lymphocytes (TILs) has been associated with a favourable prognosis among patients with HER2-enriched and triple-negative tumours(170, 171). Among breast cancer molecular subtypes, luminal-A tumours showed the greatest heterogeneity in immune subtypes(169) and a subset of luminal tumours showed high level of TILs comparable to that of HER2-enriched and basal-like tumours(172).

In a study analysing the presence of B lymphocytes (CD20+) in 1470 breast cancers Mahmoud et al(173), showed that total B-cell counts correlated with a higher tumour grade and a basal phenotype, as well as oestrogen receptor (ER)/progesterone receptor (PR) negative breast cancers. Other studies have reported statistically significant risk ratios between Autoimmune Arthritis (AiA) and breast cancer(174). Mammary gland tumours were generated in SKG mice that were genetically prone to develop AA.

Two breast cancer cell lines, one highly metastatic (4T1) and the other non-metastatic, (TUBO), were used to generate the tumours in the mammary fat pad. A three-fold increase in lung metastasis and a significant increase in the incidence of bone metastasis was seen in the pro-arthritic and arthritic mice compared to non-arthritic control mice. An elevation of serum pro-inflammatory cytokines that included interleukin-17 (IL-17) was seen and treatment with

anti-IL17 + celecoxib, an anti-inflammatory drug completely abrogated the development of metastasis and significantly reduced the primary tumour burden(174). Given the prognostic value of TILs, characterizing immune subtypes in a subtype-specific manner in diverse populations will have important implications for immunotherapy.

3.1.4 Tumour Microenvironment

The tumour microenvironment (**Figure 4**)(175), is a dynamic network that includes the tumour cells, stromal tissue within which are the immune cells, fibroblasts, pericytes, cytokines, and tumour vasculature and the extracellular matrix(176). In addition to epithelial and stromal compartments, the tumour microenvironment contains several cell types of the innate and adaptive immune systems including B and T lymphocytes, dendritic cells, and macrophages(176).

The tumour microenvironment promotes tumour initiation, growth, migration, metastasis, and therapeutic resistance(177). The immune system is a versatile defence mechanism that can recognise and eliminating a limitless variety of foreign invaders. The basic steps involved in a successful immune response against a tumour include the following: First tumour antigens are successfully captured and displayed on Antigen Presenting Cells (APCs), which are then presented to lymphocytes(178). This then causes CD4+ and CD8+ lymphocyte expansion leading to cell-to-cell contacts between activated lymphocytes and tumour cells. The last step is the production of inflammatory cytokines (by the CD4+ cells) and killing of the tumour cells (by the CD8+ cells) and other mechanisms involving B cells, NK cells and macrophages(178).

The mediators and enzymes produced can induce important genetic changes in tumour-associated cells, such as fibroblasts, endothelial cells, adipocytes, and leukocytes. These cells are critical components of the tumour stroma and can damage the microenvironment, enabling malignancy(176, 179).

THE TUMORIGENIC AND TUMOUR SUPPRESSIVE BALANCE BETWEEN THE TUMOUR AND THE IMMUNE SYSTEM DURING TUMOUR PROGRESSION IS CALLED IMMUNOEDITING AND COMPRISES THREE PHASES: ELIMINATION, EQUILIBRIUM, AND ESCAPE(7, 179). IN THE ELIMINATION AND EQUILIBRIUM PHASES, TUMOURS ARE ELIMINATED OR HELD IN CHECK BY THE PRODUCTION OF TUMOUR-INHIBITING CYTOKINES AND THE INFILTRATION OF CELLS OF BOTH THE INNATE IMMUNE SYSTEM AND THE ADAPTIVE IMMUNE SYSTEM, SUCH AS Th1 CD4+ AND CD8+ T CELLS (FIGURE 5,6)(7, 179) . ON THE OTHER HAND, IN THE ESCAPE PHASE, TUMOURS OFTEN DEVELOP MULTIPLE EVASIVE MECHANISMS, THUS DIRECTLY PROMOTING CANCER CELL PROLIFERATION AND MIGRATION (FIGURE 7)(7, 179).

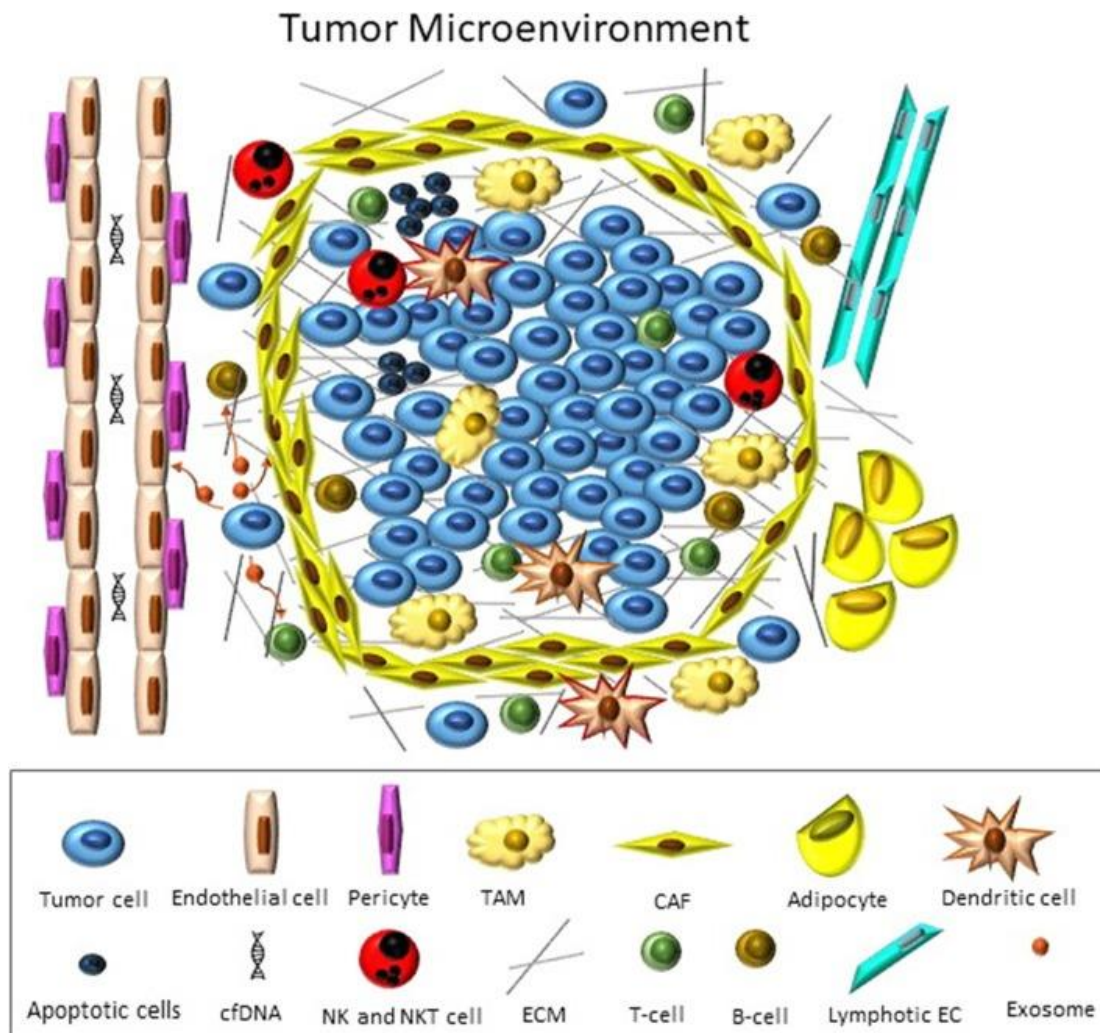


FIGURE 4: TUMOUR MICROENVIRONMENT

©Cell Communication Signal, 2020

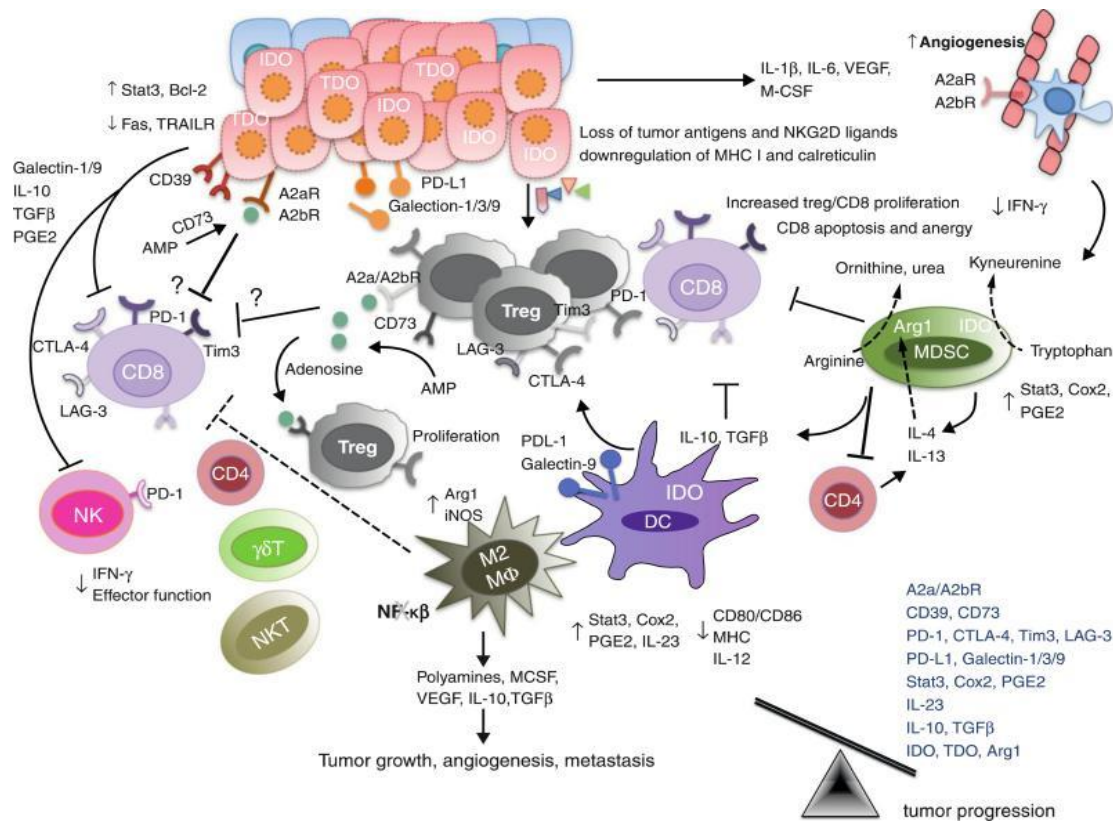


FIGURE 7: TUMOUR ESCAPE PHASE

©Current Opinion in Immunology,2014.

3.1.5 Tumour Infiltrating Lymphocytes in Breast Cancer

Tumour infiltrating lymphocytes (TILs) are defined as immune cells that are present in or have moved to the tumour itself and are around its microenvironment (**Figure 8**)(180, 181), and produced by the immune system in response to a neoplastic transformation(182). The role of TILs in BC has generated much interest in recent times. Denkert and colleagues(183) were the first to demonstrate that a high TIL count in BC tissues was associated with a significant pathologic complete remission (pCR) rate following Neoadjuvant therapy compared to TIL low tumours. Higher levels of TILs also predict for increased responsiveness to chemotherapy independent of oestrogen receptor (ER), progesterone receptor (PR), and Human Epidermal Growth Factor Receptor (HER2), with each 1 % increment associated with a further increase in the rate of pCR(170). On the other hand, humoral responses to cancer have been associated with a poor prognosis; B lymphocytes are the central component of Humoral Immunity and exert their effector function through antibody production, antigen presentation, and secretion of pro-inflammatory cytokines(184). B cells limit anti-tumour immunity through inhibition of Th1 and CTL responses while simultaneously bolstering Th2-effector cell protumour functions(184).

Various studies have determined that the TIL population in breast cancer is composed of cytotoxic CD8⁺ T cells, Helper CD4⁺ T cells, CD19⁺ B cells and rare NK cells(23).

In a study cohort of 1165 invasive breast cancers from the retrospective Nottingham Primary Breast Carcinoma Series recruited over a ten-year period (1988-1998), 74.7% of the tumours were reported to be ER positive, 59% were PgR positive, 16% were TNBC and 9.3% were HER2+. On immunohistochemical analysis, among the T-lymphocyte population, CD3+ immune cell subtype counts/core diameter assessed as absolute number of cells were the major component of all TILs in the whole series (n=1165), with a mean of 55 and a median value of 33 (Range 0-450) followed by CD68+ macrophages (mean value of 32, median of 24 with a range of 0-250 while the mean and median values for CD8 positive TILs was 25 and 11 respectively (range 0-242). T regulatory FOXP3 positive cells were the least number of T-cells with a mean value of 14, a median of 7 and a range of 0-153. Among all TILs the CD20+ B cells formed the most minor component (mean 13, median 1 with a range of 0-415). Within the molecular subtypes of breast cancer, the median CD3 values were highest for TNBC at 67 followed by HER2 enriched BC at 51, Lumina B at 25 with the least value for Luminal A at 22. Additionally, the TNBC molecular subtype of breast cancer had the highest mean and median absolute number of cells value for all the other subtypes of immune cells followed by HER2 enriched, Luminal B and Luminal A in that order(185).

A previous study examining the immune cell subtypes in thirteen “medullary” breast cancer in comparison to thirty-four ductal and four lobular cancers found that the former showed a higher tumour epithelial CD8/FOXP3 ratio at 2.6 versus a 1:1 ratio for the ductal and lobular variants of breast cancer. This predominance of CD8 + effector T cells could perhaps explain the better prognosis associated with the medullary pattern of breast cancer over the other variants(186).

A more recent retrospective study from Leeds Teaching Hospitals of patients recruited between 2011 and January 2012 examined 62 breast cancer tissues (ER/PR positive 35 (56.4%) and 17.7 HER2+ cancers) and found that CD4 + TIL subsets were the most common infiltrates however the differential stromal infiltration by these cells was also most variable and ranged from 0.5% to 80%. The least number of TIL subsets were the CD20+ B lymphocytes whereas unlike the CD4+ cells which were located mainly in the stroma and the tumour edge, the CD8 + cells were the most prevalent in tumour nests. FoxP3+ lymphocytes were the most rare in all locations for all tumour subtypes. Additionally, all the TIL subsets (CD4+, CD20+, FoxP3+) were significantly higher in the stroma and the tumour edge of grade 3 tumours and in ER/PR negative tumours(187).

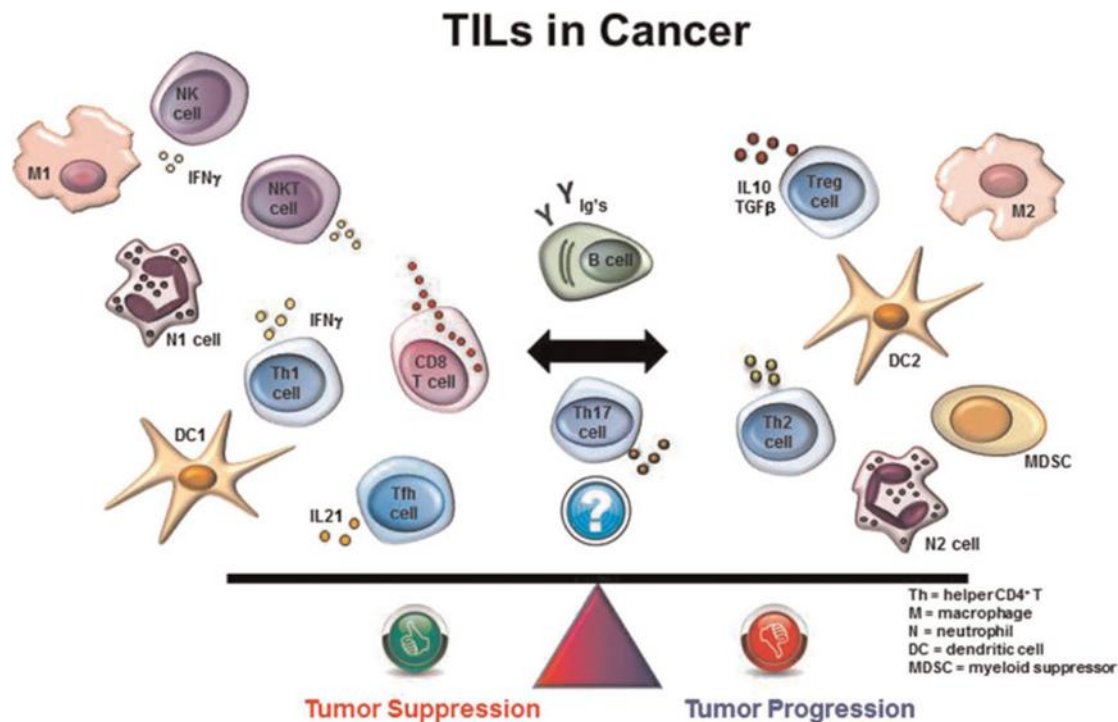
In addition to these, macrophages and dendritic cells also infiltrate the tumour in various ways(170, 188). CD4+ T helper 1 (Th1) cells, CD8+ are essential in preventing further tumour proliferation. On the other hand, CD4+ forkhead box P3 (FOXP3+), CD4+ Th2 cells, M2 macrophages and myeloid-derived suppressor cells (MDSCs) promote tumour growth(170, 188).

3.1.5.1 *Categories of TILs*

Tumour Infiltrating Lymphocytes are widely defined into two categories: Stromal and Intratumoural TILs. Stromal TILs are described as mononuclear immune cells which are not directly in contact with the tumour, while intratumoural TILs are those which have direct cell to cell contact with the carcinoma cells(170). Although it was previously hypothesized that intratumoural TILs were a better marker for prognosis in breast cancer(180), and considered more reproducible due to their location, current studies(181), have shown that stromal TILs are more clinically relevant due to their larger number (most of the TILs in breast cancer are

concentrated in the tumour stroma), and because they are easily observable in H&E stained slides(180). Not only does the amount of lymphocytic infiltration but also the phenotype of that infiltrate determine clinical outcome(173, 189).

The term Lymphocyte Predominant Breast Cancer (LPBC) describes cancers which contain more lymphocytes than tumour cells. These type of cancers represent only 10% of primary breast cancers(190). "Medullary carcinoma" was classified as a unique histologic form of breast cancer in the first through third editions of the WHO classification of breast cancers. It was characterized by Grade 3 nuclei with syncytial growth, a well- defined border, triple negative biomarker profile and a vigorous lymphoplasmacytic infiltration. Due to the immune rich stroma, these tumours were noted to have a better prognosis compared to grade-matched invasive ductal cancers(191). However, a significant inter observer variability in the reporting and use of diagnostic criteria was observed which consequently led to the terminology for this unique histologic type to be broadened in the 4th edition (2012) to "carcinoma with medullary characteristics," that incorporated invasive ductal carcinomas (no special type) possessing medullary features, atypical medullary carcinomas, and medullary carcinomas. In acknowledgement of the prognostic advantage of the TIL rich stroma, these tumours are no longer regarded as a distinct specific histologic type in the 5th edition (2019) of the WHO classification of breast tumors, but rather as a TILs rich pattern in the spectrum of high-grade invasive ductal carcinoma (no special type). According to the WHO, tumors having these characteristics should be labelled as traditional invasive ductal carcinoma with "medullary pattern" (192).



THE CELLULAR CROSS-TALK BETWEEN DIFFERENT LEUKOCYTE SUBSETS AND THEIR PREDOMINANT CONTRIBUTION TO EITHER PRO- OR ANTITUMOR ACTIVITIES, INCLUDING MYELOID LINEAGE LEUKOCYTES, TUMOUR-ASSOCIATED MACROPHAGES WITH EITHER PROTUMOURIGENIC (M2) OR ANTITUMOURIGENIC (M1) PROPERTIES, HELPER T-CELL SUBSETS, CYTOTOXIC T CELLS, REGULATORY T CELLS, B CELLS, DENDRITIC CELLS, AND MYELOID-DERIVED SUPPRESSOR CELLS ARE SHOWN. THESE CELLS PLAY CENTRAL ROLES IN SHAPING THE MICROENVIRONMENT VIA THE FACTORS THEY PRODUCE THEREBY DRIVING EITHER AN IMMUNE-MEDIATED ANTI- OR PROTUMOR ACTIVITIES IN THE MICROENVIRONMENT.

FIGURE 8: THE CELLULAR CROSS-TALK BETWEEN DIFFERENT LEUKOCYTE SUBSETS

©Annals of Oncology, 2015.

3.1.5.2 Other Immune cells in breast cancer Tumour Microenvironment

3.1.5.2.1 Macrophages

The role of macrophages in breast cancer can be linked to altered function since macrophages are thought to play an important physiological role in branching morphogenesis during puberty and pregnancy, and in post lactation involution(193).

Tumour associated macrophages (TAMs) correlate inversely with the expression of hormone receptors which are considered favourable prognostic factors(194).

Macrophages are often classified into either Th1-activated M1 or Th2-activated M2 subtypes and those within tumour sites are usually of the M2 subtype; Inhibition of cyclooxygenase (COX)-2 or enhanced expression of microRNA miR-19a-3p suppresses breast cancer metastasis by preventing M2 phenotype conversion(195).

Hypoxic conditions within tumour microenvironment promotes M2 differentiation which leads to an immunosuppressive environment through production of IL10, Arginase 1 and Vascular Endothelial Growth Factor (VEGF all of which suppress Th1 T cell function(196). Normalization of the tumour vascularization coupled with anti-angiogenic therapy has thus been suggested as a potential immunotherapeutic strategy(196).

3.1.5.2.2 T Regulatory (Treg) cells

Regulatory T cells (Tregs) are crucial to the maintenance of immune tolerance. Natural Tregs express cell surface CD4 and CD25, and develop in the thymus from where they emigrate into the circulation to suppress a wide range of immune cells, including CD4+ and CD8+ T cells, NK cells, NKT cells, B cells, and antigen presenting cells(197). Adaptive Tregs are non- regulatory CD4+ T-cells which acquire CD25 (IL-2R alpha) expression outside of the thymus, and are activated by diseases such as cancer(197). The Tregs also express Forkhead Box Protein 3 (Foxp3) which helps to differentiate them from CD4 /CD25 positive T effector cells. The proportion of Tregs are increased from 5- 10% of total CD4 positive T cells in circulation to up to 30% within the tumour mass(198). Infiltration of Treg cells into breast cancers has been shown to be associated with a worse prognosis(199).

Circulating Treg cells can be recruited to a breast cancer site through multiple signalling pathways, including CCL22/CCR4(200). In a study of 164 TNBC cases from Singapore, the abundance and localization of T cells, B cells and Tregs was analysed for associations with clinicopathologic features of the cancer or clinical outcomes(201). The study concluded that the combination of high densities of intra-tumoural but not stromal Tregs, CD8+ T cells and CD20+ B cells provided overall and disease free survival benefit in TNBCs highlighting how the interaction between the various TILs impacts clinical outcome(201).

3.1.5.2.3 Myeloid-derived suppressor cells (MDSCs)

The MDSCs comprise a heterogeneous population of immature cells of myeloid origin that expand during pathological conditions such as cancer, inflammation, and infection(202). These cells express CD33 and CD11b(203) . MDSCs are another major immunosuppressive cell type found in breast tumours(203). In breast cancers, MDSC inhibit T cells, NK cells, and dendritic cells (DCs) but stimulate immune regulators such as Th2 T cells, Tregs and TAMs. Some of these effects are mediated through secretion of IL-6 and IL-4(203).Higher levels of MDSC in circulation correspond to advanced stage and metastatic breast cancer while reduced levels have been observed in patients on neoadjuvant chemotherapy who achieve complete pathologic response(203) thus suggesting a potential role of MDSCs in breast cancer as prognostic and predictive biomarkers.

3.1.5.2.4 Th17 cells

The Th17 is a recently discovered type of CD4 T helper cell, characterized by the production of IL-17. Both antitumor and tumour-promoting functions have been identified in Th17 cells, depending on the tumour type(204). A high number of IL-17 producing cells in breast tumours have been shown to correlate with high histological grade, negative ER/PR status, and triple negative phenotype(205).

3.1.5.2.5 Mast Cells

Mast cells are commonly found in tumours. Pre-clinical data suggest that they contribute to tumour progression, possibly through promotion of angiogenesis (165). Mast cells have been shown to have both pro tumour and anti-tumour effects(206). The presence of stromal mast cells has been shown to have a good prognosis in node negative breast cancer(207).

3.1.5.2.6 Dendritic cells

Dendritic cells (DCs) are antigen-presenting cells with a unique ability for antigen cross-presentation for T helper lymphocytes (CD4+) and cytotoxic T lymphocytes (CD8+) in secondary lymphoid organs(208). Dendritic cells mature through interactions with Toll like Receptors (TLRs) or inflammatory cytokines and activate naïve T cells by presentation of a specific antigen(209). In BC, infiltration of mature DCs into primary tumour lesions is associated with prolonged patient survival and reduced risk of metastasis(210).

3.1.5.3 Methods used to evaluate TILs

Various methods have been employed to determine and properly evaluate TILs; Light Microscopy of H&E stained slides(211), IHC/Flow cytometry(211), Multiplex immunofluorescence (IF)(198)and Gene expression analysis(201). The main advantage of using H&E-stained slides is that the sample is easy to prepare and evaluate, and results and material/reagents are easily available. However, its limitation includes the fact that lymphocyte population cannot be accurately estimated, and that the exact functional status of the lymphocytic infiltrates cannot be determined. Immunohistochemistry(IHC) allows for better identification of TILs, however the requirements for analysis are a lot more demanding compared to a simple H&E stain(211).

Gene expression analysis together with quantitative polymerase chain reaction (PCR) and RNA sequencing is a procedure becoming increasingly popular to detect the various subpopulation of data at once(211). One of the main drawbacks is that it requires complex statistical and bioinformatics tools for analysis(190).

However, development of immunohistochemical markers that include FOXP3 for T-regulatory cells (Tregs) and Granzyme B for Cytotoxic CD8 cells have made it possible to identify individual functional subsets of TILs thus contributing to advances in this field and predicting tumour response to future targeted immunotherapy(212).

To standardize TIL reporting the International Immuno-Oncology Biomarker Working Group developed a reporting guideline in 2014 for scoring TILs in breast cancer(21).

With regards to TIL heterogeneity, a study of 230 H&E-stained sections was used to assess TIL infiltration and heterogeneity and found that there was no difference in overall TIL scores for H/E slides from different blocks of the same case(185).TIL heterogeneity was noted across a core rather than multiple cores from the same case. The authors thus recommended a single full face H/E section per tumour to assess TIL scores(185), which is the current guideline of the International Immuno-Oncology Biomarker Working Group(180, 213)

1.5.4 Relationship between TIL and tumour characteristics

A study by Wang et al (174), reported that TILs correlated with high tumour grade. Similarly, in a pooled analysis from 2,148 patients with early-stage triple negative breast cancer (TNBC), from nine studies, significantly lower stromal TIL levels were found to correspond to increased tumour size ($p = 0.01$), more nodal involvement ($p = 0.02$), and lower histologic grade ($p = 0.001$)(190).

3.1.6 Prognosis and predictive value of TILs in breast cancer

3.1.6.1 Relationship between TIL concentration and prognosis

A directly proportional relationship between total TILs and improved prognosis of breast cancer was reported by Yu et al's., 2016 systematic review(214). Whereas some subtypes predicted a worse prognosis, a larger value of total tumor-infiltrating lymphocytes (both intraepithelial and stromal) was linked with a better prognosis (pooled HR 0.88, 95 percent CI 0.83-0.94). According to the pooled data, both iTILs and sTILs were linked to better DFS and cancer-specific survival, with estimated HRs in iTILs and sTILs being 0.90 (95 percent CI 0.83-0.98) and 0.85 (95 percent CI 0.76-0.94), respectively. However, they were not significantly correlated with OS (iTILs: 0.90, 95 percent CI 0.76-1.06; sTILs).

In the Clinical Evaluation of Pertuzumab and Trastuzumab (CLEOPATRA) multicenter multicounty Trial, treatment naïve patients with either locally recurrent unresectable or Stages IV HER2 positive breast cancers were randomized either to pertuzumab or placebo with docetaxel and trastuzumab in the neoadjuvant setting. TILs were assessed on full face BC tissue sections in continuous increments of 10%. Outcomes of interest were progression free survival and overall survival. In the 678 patients with evaluable TILs, the mean stromal TIL value was 21.07% (SD 22.4), and the median stromal TIL value was 10% (IQR 5–30). Although there was no association between TILs and progression free survival, patients with TIL values greater than 20% had a better overall survival (HR 0.76, 95% CI 0.60–0.96, $p=0.021$). Furthermore, the three-year overall survival was higher in the pertuzumab group (78%) and in the placebo group (55%) when the TIL values were 20% or more as compared with TIL values < 20% (64% and 50% three year overall survival respectively)(213).

In the NeoALTTO trial (Neoadjuvant Lapatinib and/or Trastuzumab Treatment Optimization) trial conducted over 29 months, 387 patients with early-stage HER2-positive breast cancer with tumours greater than 2 cm in size were assigned to 1 of 3 neoadjuvant treatment arms: trastuzumab, lapatinib, or the combination for 6 weeks followed by the addition of weekly paclitaxel for 12 weeks, and by 3 cycles of fluorouracil, epirubicin and cyclophosphamide post- surgery. In this study, TILs were assessed on pre-treatment core biopsy sections. Results showed that the median (interquartile range [IQR]) level of TILs was 12.5% (5.0%-30.0%). Independent of the treatment group, TIL levels above 5% were associated with higher pCR rates (adjusted odds ratio, 2.60 [95% CI, 1.26-5.39]; $P=.01$). Furthermore, every 1% increase in TILs was linked to a 3% decrease in the rate of an event (adjusted hazard ratio, 0.97 [95% CI, 0.95-0.99]; $P=.002$) across all treatment groups with a median follow-up period of 3.77 years thus confirming that in early HER2 positive breast cancer TILs serve as an independent positive prognostic biomarker(215).

In the multicenter FinHER (Finland Herceptin) adjuvant clinical trial participating women with breast cancer were randomized to docetaxel or vinorelbine, followed by fluorouracil, epirubicin, and cyclophosphamide. Tumour infiltrating lymphocytes were assessed in continuous increments of 10% in 935 full face tissue sections of breast cancer tissue. Results showed that compared to luminal A tumours higher TIL levels were significantly correlated with Triple negative breast cancer, HER2 positive tumours, larger tumours, and high-grade tumours. Additionally, in early TNBC, the study confirmed that for each 10% increase in TILs there was a 13% reduction in the relative risk of distant recurrence (HR 0.77; 95% CI 0.61-0.98; $P=0.02$). Due to the few numbers of events, TIL correlation with overall survival was not possible. Furthermore, in patients who also got trastuzumab in addition to their chemotherapy, each 10% increase in TIL infiltration was associated with an 18% decrease in the relative risk of distant recurrence(171).

A follow up pooled analysis of nine studies that included 2148 patients which examined the role of stromal TILs in early-stage TNBC in the adjuvant setting. The Invasive disease-free survival hazard ratio was 0.87 (95% CI, 0.83 to 0.91), 0.83 (95% CI, 0.79 to 0.88) for Distant Disease-Free survival, and 0.84 (95% CI, 0.79 to 0.89) for Overall Survival for every 10% increase in sTILs. Interestingly, the three-year Invasive disease free survival was 92% (95% CI, 89% to 98%), Distant Disease Free survival was 97% (95% CI, 95% to 99%), and Overall Survival was 99% (95% CI, 97% to 100%) in node-negative patients with sTILs 30%(216). Seventeen studies involving 12,968 BC participants between January 1990 to July 2014 were subjected to assessment based on the Newcastle–Ottawa scale (NOS). Interestingly, the authors found that both intraepithelial iTILs and stromal sTILs played a part in better breast cancer prognosis. However, TIL subtype, as previously discussed, showed different results. While CD8+ TILs predicted improved prognosis, PD-1+ TILs or Foxp3+ TILs predicted otherwise. It is also worth noting that the study did not find a significant correlation between total TILs in breast cancer and clinicopathological characteristics, probably due to limited data and lack of large-scale studies.

In a different meta-analysis involving twenty-three studies with 13100 breast cancer patients, high TIL levels were significantly associated with elevated pathological complete response (pCR) as well as overall survival post neoadjuvant therapy, specifically in the triple-negative subtype(217).

Similarly, in a pooled analysis of 3771 breast cancer patients undergoing neoadjuvant therapy(218), higher TIL concentration was found in a proportional relationship with elevated survival rates in patients with HER2-positive breast cancer and TNBC. However, higher TIL levels predicted decreased survival in luminal-HER2-negative breast cancer patients, which Denkert et al(218) inferred as a different biological response to TIL levels in this subtype. Biopsies of 3771 breast cancer patients analysed in this study originated from six trials conducted by the German Breast Cancer Group. Assessment methods of stromal TILs in the study adhered to the International TIL working group's guideline(183, 188, 218). In study by Luen et al., (190, 213), a significant correlation between every 10% increase in stromal TILs and enhanced overall survival (adjusted HR 0.89, 95% CI 0.83-0.96, $p=0.0014$) was seen in HER2 positive BC.

The most consistent positive correlation between TILs level and improved breast cancer survival can be witnessed in the increase of CD8+ T cells. In a critical meta review in 2016, Stanton and Disis et al, reported the association between greater intratumoural CD8+ T infiltrate and higher specific survival rate in a study involving 1334 breast cancer patients(188).

3.1.6.2 Relationship between TIL and breast cancer subtype

A systematic review(188)involving fifteen studies ($n = 13\,914$) identified the variability of TIL levels within and across breast cancer subtypes. Specifically, 48% of all breast cancers showed CD8+ T-cell infiltrates (range, 32%-80%). While TN (60%; range, 40%-91%) and HER2+ disease (61%; range, 40%-83%) showed complementary numbers, CD8+ TIL was observed in fewer HR+ tumours (43%; range, 30%-73%)(47). Similarly, TNBC (70%; range, 65%-76%) and HER2+ disease (67%; range, 61%-74%) showed the highest FOXP3+ levels, whereas a small number of HR+ breast cancers possessed high levels of FOXP3+(38%; range, 35%-41%)(188).

According to various studies, there has been a significant correlation between increased TILs and improved pCR and recurrence free survival for TNBC and HER2+ breast cancer(171). The Finland Herceptin (FinHer) trial found that higher TIL levels led to decreased recurrence of disease in HER2+ early breast cancer (Each 10% increase in TILs showed an 18% decreased risk of distant recurrence). A review article by Wein et al(211), noted that in the BIG 02-98 trial, each 10% increase in stromal TILs resulted in a 15% increase in disease free survival and a 17% reduction in the number of deaths in TNBC irrespective of the chemotherapy regime. In a review conducted by Stanton et al(188), on the association of TILs in BC, it was concluded that TNBCs had a large number of lymphocytic infiltrates particularly in patients of a younger age. It was also noted that the expression of the PDL1 ligand was predominant in patients with a higher proportion of TILs (>50%) and when treated with neoadjuvant chemotherapy showed a 40% pCR. In contrast, those patients with tumours that had no lymphocytic infiltrate showed a pCR of 7% only.

Similarly, a retrospective analysis of a case-cohort series of 987 patients with early ER+/HER2- patients, observed a positive correlation between TILs and tumour grade ($p < 0.0001$)(219). Although under univariate Cox regression analysis, no relationship was found between TIL and distant disease-free survival (DDFS), a positive relationship with grade was observed in patients undergoing adjuvant chemotherapy, especially in the subgroup rich in Ki-67(219). The study concluded that higher immunogenicity might be present in high-risk groups.

A meta-analysis(220) on TIL levels in triple-negative breast cancer (TNBC) also observed findings similar to previous studies(221). In this analysis, 37 selected studies were compiled from MEDLINE, EMBASE, and Web of Science databases dated before August 2019. The results showed a directional relationship between TIL levels in TNBC and pathological complete response (pCR) to treatment (odds ratio [OR] 2.14, 95% confidence interval [CI] 1.43-3.19). Every 10% TIL increase corresponded to improved pCR in TNBC patients (OR 1.09, 95% CI 1.02-1.16). Increased TILs also corresponded to increasing OS (HR 0.90, 95% CI 0.87-0.93) and Disease Free survival (DFS) (HR 0.92, 95% CI 0.90-0.95)(222). With regards to TIL subtypes, CD4+ TIL level showed a positive correlation with both OS and DFS, whereas CD8+ TIL and FOXP3+ TIL level only showed a correlation with DFS(220). A study involving 1,463 patients with early stage TNBC also identified a correlation between TIL and decreased mortality(223). The study adopted multivariable regression models regarding clinicopathological aspects to investigate the relationship between absolute lymphocyte count (ALC), breast cancer-specific mortality (BCM), and overall mortality. TIL levels also demonstrated a significant relationship with increased peripheral lymphocyte counts during neoadjuvant chemotherapy as deduced from an exploratory analysis of 70 cases emphasising the potential role immune function plays in the efficacy level of early TNBC treatment.

Implementing the random or fixed-effects model to review 25 studies involving 22,964 patients, a meta-analysis by Mao et al(221), identified CD8+ as a reliable predictor for better prognosis in ER negative (HR = 0.73; 95% CI, 0.68–0.80) and HER2+ (HR = 0.71; 95% CI, 0.57–0.88), but not in ER positive patients. In contrast, FOXP3+ was identified in correlation with low OS (HR = 3.26; 95% CI, 1.51–7.04) and BCSS (HR = 1.26; 95% CI, 1.01–1.57) in ER positive patients, but this was not observed in ER negative patients(221).

3.1.6.3 Racial differences in TILs in breast cancer

Table available:

<http://www.plosone.org/article/fetchSingleRepresentation.action?uri=info:doi/10.1371/journal.pone.0152500.s006>

Natural selection of genetic variants common in populations of African ancestry are likely a result of adaptation to exposure to infective agents(224). For example, a recent study found that pathways that are enriched for signals of adaptation tend to be involved in immune response and wound healing(168). However, there is little data from Africa to support the notion that tumour biology, and by implication tumour immune repertoire, differs by ethnicity.

In women under the age of 50 years, the median mitotic rate 14.5 (9.0–29.0) was more than three times higher in the Tanzanian group than in the white American group 4.0 (1.0–13.0), and more than two times higher in the black American group 15.0 (4.0–29.0) than in the white American group. Regardless of the biomarker status, white American women had the lowest median mitotic rate and Tanzanian women had the highest. Also, the median TIL level in Tanzanian women under the age of 50 years was higher 21.5 (12.5–30.5) than the median TIL level in White Americans 16.0 (8.0–20.0). When categorizing into ER and HER2 groups, the differences amongst races were no longer statistically significant ($p = 0.78$) (88). In a study by Sawe et al(225), tissue microarrays (TMAs) were constructed from 68 cancerous and 89 non-cancerous Kenyan breast tissue samples and stained for TIL subtypes. In comparison to benign tissue, analysed breast cancer tissue samples showed increased T cell and macrophage immune cell markers (CD4, CD25, CD63, CD163) expression(225). A follow-up study in 2017, examined the potential correlation between TIL density and tumour-related factors, namely, grade, molecular subtypes, and proliferation index. Using TMA sections from 49 breast tissue samples they reported that there was significant correlation of CD68, CD163, and CD25 with Ki67 but neither tumour grade nor BC subtype(226).

A metagene analysis of RNA-seq expression from 162 African American and 697 Caucasian breast cancers suggested no statistically significant racial differences in immune gene expression or TIL counts in TNBC, however, the authors emphasized that subtle differences regarding immune cell subtype distribution and activity cannot be excluded given that these may not be obvious in bulk RNA analysis(227). Here, RNA-seq expression data originated from the TCGA database, while the TIL count data was available from previously published work(227).

However, a different finding can be seen in the study by Elkhanany et al., 2020(228). Immunohistochemistry data from 182 African American (AA) and 755 Caucasian American (CA) breast cancer patients was obtained and analysed, indicating higher levels of TILs, T-Regulatory (Tregs) and T-Follicular Helper cells, as well as higher Treg: TILs ratios in African American. However, TILs served as predictor of OS advantages in the CA group only, meanwhile, Treg, Treg: TILs and Treg:CD8+ ratio were predictors of DSS and DFI disadvantages. This suggested that African American patients possessed more immune infiltrates, with immunosuppressive fractions surpassing those in Caucasian American patients, forecasting T-cell exhaustion and worse breast cancer outcomes. These findings are similar to the study by Abdou et al., 2020(229). Abdou et al., 2020(229), applied immunohistochemistry on TMAs of 688 invasive breast tumour samples (550 African Americans, 138 Caucasian) and observed that the percentage of high-grade, ER negative, and TNBC tumours, along with the density of CD8+ T cell in AA breast cancer patients

significantly outnumbered that in CA patients (median 87.6/mm² vs. 53.1/ mm²; $p < 0.001$). Although the study did not find a correlation between CD8+ T cell density and overall survival in the entire population, when focused on AA population, CD8+ T cell density demonstrated a positive correlation with overall survival(229).

In a study consisting of 462 Asian patients and 2,186 Western patients, two algorithms ESTIMATE (Estimation of Stromal and Immune Cells in Malignant Tumours Using Expression Data) and CIBERSORT (Cell Type Identification by Estimating Relative Subsets of Known RNA Transcripts) were adopted to determine and compare immune score, composition, and prognostic impact of infiltrating immune cells between Asian and Western population subgroups(230). Here, high levels of CD 4 T cells showed significant association with prolonged OS in both population groups(230). Meanwhile, in Asian patients a significant association between CD8 T cells and longer OS(230), was observed. It is also interesting to note that no significant differences in terms of compositions of infiltrating immune cells were detected between the two cohorts(230).

3.1.6.5 TIL concentration and other risk factors

The study by Loi et al(216), suggested an inversely proportional relationship between stromal TIL levels and age. A study investigating 1903 breast tumour samples (116 young and 1787 old) uncovered a significant relationship between CD8 T cells and disease-free survival (DFS), however only in patients under 40 years old(231). The authors stated that this difference may lie in other more complicated and confounding relationships than simply age alone(231)

A study on 236 lean patients and 209 overweight patients indicated a significant relationship between high stromal TIL and elevated rate of event-free (EFS) and overall survival (OS) in lean patients, however the opposite was observed in their overweight counterparts(232).

A study involving 956 breast cancer patients with smoking status identified no significant impact of tobacco consumption on TIL levels (stromal and intra tumoural) prior to and after neoadjuvant chemotherapy(233). Applying a median follow-up of 101.4 months, the study did not find any association between smoking status and relapse-free survival (RFS) and overall survival (OS) in breast cancer patients(233).

These data call for further investigation of the interaction between immune cell types within the tumour microenvironment and highlights how TIL density and localization within tumours can impact clinical outcome.

3.2 OBJECTIVES

1. Quantify tissue tumour infiltrating lymphocytes (TILs) in intratumoural stroma and the stroma at the leading edge of invasive breast cancer (IBC).
2. Describe the subsets of inflammatory cells in intratumoural stroma using a selected panel of cell specific IHC antibodies

3.3 METHODS

3.3.1 Case Selection

A total of 823 Invasive Breast Cancer cases were listed on a spreadsheet. The remaining 44 cases were removed due to significant amounts of missing data (>50%). Only those cases which had definitive surgical procedures (lumpectomies, mastectomies) were included for review (N=347). Cases with only core biopsies were excluded (due to limited material available for experiments) as were those cases with a prior history of neoadjuvant therapy. Where available, slides and corresponding Formalin Fixed Paraffin Embedded (FFPE) tissue blocks of cases were retrieved from archives. In cases where there were no slides but blocks available, new slides of tumour blocks were cut, stained with H&E, and reviewed for tumour content. In some cases, (n=85) there were no slides or blocks available; these cases were deleted from the study list. H&E-stained tumour slides were reviewed for the following:

- Adequacy of tumour tissue (invasive carcinoma covers more than 50% of total slide section)
- Preservation of tumour (more than 50% of total slide section showing well preserved tumour) 50% proportion of viable tumour was an arbitrary figure used. The study required us to have four cores on the selected tumour block, this would mean that a lot of tissue would be used, and we did not want exhaust the tumour on the block.
- Two tumour slides were selected for each case fulfilling the above criteria for a total of 262 cases.

Data for molecular subtypes of breast cancer was extracted from the previously annotated database. Molecular subtypes were defined based on the 2013 St.Gallen's definition(16); Luminal A: ER+ and PR+, HER2-, and low Ki-67 (<20%) /histologic grade (I or II); luminal B-HER2+: ER+ and/or PR+, and HER2+; luminal B-high proliferative: ER+ and/or PR+, HER2-, and high Ki-67 (>20%) histologic grade (III); HER2-enriched: ER-, PR-, and HER2+; and triple-negative (TN): ER-, PR-, HER2-. We combined the two luminal B subtypes as one group in TIL analyses because of the small sample size.

3.3.2 Visual TIL Quantification

Visual quantification of TILs within the intra-tumoural stroma and the peri-tumoural/ leading / invasive edge of the tumour were performed at AKUHN based on the international recommendations of the TIL Working group(234).

Identification and semi quantification of TILs was done for all study cases on one full thickness 4 µm H&E stained FFPE tissue section of selected study cases. The TILs were counted in the tumour stroma within the tumour borders and at the leading edge of the tumour/peritumoural area. Intratumoural stromal TILs were defined as those dispersed in the stroma between the carcinoma cells and did not directly contact carcinoma cells(234).

Leading/peritumoural/ invasive edge TILs were defined as those that were in the vicinity of the leading edge of the tumour but within the tumour boundary (**APPENDIX III**). Semi quantification of the TILs was done at a total magnification of 200x on the Olympus microscope BX43 Model U-SD03, as per current recommended international guidelines(234).

Briefly, a TIL (mononuclear cells, lymphocytes, and plasma cells) proportion score of stromal and leading edge of IBC was assessed in increments of 10 to the nearest 10%. The denominator was the total stromal area and the numerator the TIL quantity occupying the denominator area. As an example, a score result of 80% stromal TILs meant that 80% of the stromal area showed a dense mononuclear cell infiltrate. TILs were evaluated only within the borders of the invasive tumour. TILs located outside the tumour border, around DCIS and normal lobules were excluded. Additionally, TILs in tumour zones with crush artefacts, necrosis, regressive hyalinization and in previous core biopsy sites were excluded.

In tumours with heterogeneous distribution of TILs, five different fields were counted, and the scores added for an average percentage score for that case.

An intratumoural stromal or peritumoural/ leading edge of tumour TIL score of 50% and above was categorized as a lymphocyte predominant breast carcinoma (LPBC)(190). The presence of aggregates of stromal lymphocytes automatically assigned the case to a LPBC.

If tertiary lymphoid follicles were seen within the vicinity of the tumour, these were not scored as part of the stromal TILs but recorded separately(234).

For quality purposes, 10% of the scores were also rechecked randomly by a second observer (ZM).

All TILs were scored by the PI blinded to clinicopathologic characteristics. Additionally, the same slides were scanned and subjected to automated TIL estimation at the Integrative Tumour Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, USA.

3.3.3 Automated TIL quantification

3.3.3.1 *Digital pathology analysis*

Stromal cellular density (SCD) within the tumour microenvironment was assessed by applying computational pathology algorithms (Halo, Indica Labs, Albuquerque, NM), an image analysis system, to digitized H&E slides in an unbiased multi-step approach.

First, based on the Halo Tissue Classification module, a multi-data point script was trained for the segmentation of breast tumours into epithelial, stromal, and adipose tissue compartments. Next, a cell-detection script was trained to identify stromal cells based on context, detection weight, contrast threshold, roundness, and size. By combining both scripts into a universal multi-layered script and applying this to the centralized analysis of all H&E stained slide images, data on the total epithelial (tumour), stromal, and adipose tissue areas (mm²) was generated based on the first script and, by limiting the cell detection script only to the stromal compartment, data on the total number of cells within the stromal compartment was generated after the second pass(235)(**APPENDIX III**). The areas demarcated by Halo were checked by pathologists. Importantly, the process of algorithm development is iterative

with the capacity for continuous training based on pathologists' evaluation of the output. Typically, the machine's performance improves with training.

3.3.4 Subtyping of TILs

The TIL subtyping was performed using Immunohistochemistry (IHC) on Tissue Microarray (TMA) Slides

3.3.4.1 Case selection for TMA for IHC Markers (CD3, CD4, CD8, CD20, CD68 and FoxP3)

One of the two tumour slides selected previously was matched with its corresponding tissue block. The final number of cases selected for the TMA were 262.

Two approximately 6 mm circles were marked with two different coloured indelible ink pens around well-preserved invasive cancer and another two 6 mm circles of the stroma around the leading edge/peri-tumoural area.

3.3.4.2 Construction and reading of TMA

The tissue microarrays of the 262 cases were constructed at Baylor College in December 2016 with the PI in attendance. Arrays were constructed using the semi- automated Beecher TMA instrument (Estigen/Beecher, Tartu, Estonia). Tissue blocks were stored under ambient conditions, at approximately 24oC. Tissue microarray blocks were constructed by taking core samples from morphologically representative areas of paraffin-embedded tumour tissues and assembling them on a recipient paraffin block. The Beecher platform uses two separate core needles for punching the donor and recipient blocks and a micrometer-precise coordinate system for assembling tissue samples on a block. For each case, 1-mm core-diameter samples were collected from the marked 6mm circle representing the intra-tumoural stroma and the peri-tumoural stroma and each was placed on the same recipient block (**Appendix IV**). The intratumoural stromal cores were placed first on the TMA block followed by the two cores from the peritumoural/ leading edge cores. The final tissue microarray consisted of blocks containing duplicates of spots. All samples were spaced 0.5 mm apart.

A map of the proposed array was constructed in Excel to determine the final location of all cores. This sample tracking system was linked to a Microsoft Excel database containing demographic, clinicopathologic, and survival data, where available, thereby allowing linkage between histologic data and clinical features. Sections were cut from the TMA block at 5µm thickness and mounted on positively charged slides. One section was stained with H&E and reviewed to confirm the original histopathological diagnosis and to ensure presence of tumour.

The array was read according to the given tissue microarray map (**Appendix IV**). Cases in which no tumour was found, or no cores were visible were excluded from the final data analysis.

3.3.4.3 Quality Control

Each TMA included adequate internal experimental control samples, such as kidney and liver tissue, and tumour replicate punches from the same tissue block for both the intratumoural and the leading edge / peritumoural stroma. Biomarker information was considered as

uninterpretable in cases where the tissue punch was lost during sectioning or processing, or there was no tumour visible on the core Immunohistochemistry.

Biomarker information was considered as uninterpretable in cases where the tissue core was lost during sectioning or processing, or where there were no visible breast carcinoma cells. These cases (n=17) were excluded from the final data analysis. Slides were deparaffinised in Bond Dewax Solution (Leica Biosystems, Buffalo Grove, IL), and rehydrated in descending grades of 100%, 90% and 70% ethanol.

Endogenous peroxide activity was blocked by pre-treatment with 3% hydrogen peroxide for 10 minutes, and antigen recovery was achieved by 20 minutes of heat-induced epitope retrieval with the Epitope Retrieval Solution 1 from Leica Biosystems. Immunohistochemical (IHC) stains were performed on an automated tissue-staining system using the Bond Polymer Refine Detection (Leica Biosystems, Buffalo Grove, IL). Tissue sections were stained with monoclonal antibodies: CD3, clone LN10; CD4, clone 4B12; CD8, clone 4B11; CD20, clone MJ1; CD68, clone 514H12 and FoxP3 (Abcam ab20034). IHC Detection was done using polymer reagent conjugated with horseradish peroxidase followed by the 3,3'-Diaminobenzidine (DAB) Chromogen Kit (Leica Biosystems, Buffalo Grove, IL).

Appropriate positive (lymph node) controls were included. Negative controls were performed without any primary antibody.

3.4 Data Analysis

3.4.1 Interpretation of IHC stained TMA sections for inflammatory cell subtypes

The stained TMA slides were scanned on the Aperio Digital pathology slide scanner (Leica Biosystems) at the Digital and Computational Laboratory, Department of Pathology, Columbia University, NYC. Images were reviewed in the Slide Manager which delivers an effective digital pathology information management system for high resolution eSlides—that can be viewed, managed, and analysed—as an integrated element in the pathology workflow (<https://www.leicabiosystems.com/digital-pathology/scan/>).

The digitized slides were then imported into the Aperio Image Scope viewer (<https://www.leicabiosystems.com/digital-pathology/manage/aperio-imagescope/>) for scoring of the stained antibodies. Briefly, TMAs stained for the antibodies of interest were assigned a template, segmented, and annotated before analysis using the appropriate macros for final analysis (**Appendix V**). The percent positive stained cells with the total stromal TILs as a denominator were used for analysis. For this analysis total TILs were grouped into three subgroups: <10, 10-30, and >30% based on previously estimated visual TILs on whole mount sections (See 3.3.2). The bins were based on the median value for the distribution of the TILs (31.67%). Better outcomes have also been demonstrated in early TNBC when a cut off 30% has been used for high TIL (220). The distribution for each marker was presented as a mean and range.

3.4.2 Survival data

The association of tumour and leading-edge scores in respect of overall survival and recurrence-free survival time analysis was performed for all cases where this data was available.

Overall survival time (OS), in months, was calculated as the difference between cancer diagnosis date and date of death or contact date for those censored ones.

Recurrence-free survival time (RFS), in months, was calculated to be the difference between cancer diagnosis date and date of recurrence or contact date for those with no event.

At the end, the analytic set of cases with outcome data and marker information consisted of 226 women.

3.4.3 Statistical Analysis

We had four percentage values of positive cells for each tumour marker representing the four cores, for each patient. Spearman correlation was used to test for accordance between their average and maximum value. We elected to use the average of the set, after assessing a high correlation between the two values (**Appendix VI**) and based on international TIL working group recommendations(236). Total tumour infiltrating lymphocytes (TIL) were also averaged through a set of four values for each woman. Each tumour marker and total TILs were transformed either by square-root or by log transformation, to ensure the normality assumption was met when used as the outcome. Generalized linear and logistic regression models were used to assess the associations between risk factors and tumour features with tumour markers and total TILs, after adjusting for other covariates. Tumour markers were treated as the outcome and used as continuous variables after appropriate transformation, and they were categorized using their corresponding quartiles, as well (**APPENDIX VII**). Total TILs were treated as a continuous and as a categorical outcome variable. We used the established cut-off points of 10% and 30% for TILs.

The Kaplan–Meier method was used to assess recurrence-free survival and breast cancer-specific survival among patients, stratified by the different categories of tumour markers and total TILs. Multivariate Cox proportional hazards model was also used to test the differences in survival outcomes across the studied tumour markers and TILs and with the adjustment of age, body mass index, tumour subtype, tumour size and grade, TNM stage, lymph nodal involvement, and chemotherapy completion status and endocrine treatment.

All statistical tests in the present study were two-sided and performed using SAS version 9.4 (SAS Institute, Cary, NC, USA) or R version 3.6.3 (R Foundation for Statistical Computing, Vienna, Austria).

3.5 RESULTS

3.5.1 Descriptive analysis of sociodemographic and reproductive characteristics

From the original sample size of 262 cases, 226 invasive breast cancer cases were eventually included in this analysis for TILs and immune cell markers (**Figure 9**).

The mean age at diagnosis of breast cancer was 48.3 years and 62% of patients were under the age of 50 years while the mean BMI was 27 kg/m² (overweight category) with 66% of women having a BMI ≥ 25 kg/m² at diagnosis.

Our study cases showed late age at menarche (mean 15 years of age, 49% > 15 years), young age at first pregnancy among parous women (mean 23 years, 69% less than 25 years of age), high parity (mean 3 children, 69% with more than 3 children), and long cumulative breast-feeding duration (mean 67 months, 40% greater than 72 months). There was almost equal

distribution of pre- and post-menopausal women in this cohort (55% and 45% respectively). (Table 8 and 9).

Of the patients eligible for anti-HER2 therapy 83% completed this therapy while 94% of eligible patients for anti-oestrogen therapy used endocrine therapy. with 65% still to complete this treatment. In those for whom data was available, 96% and 90% of patients who commenced chemotherapy and radiotherapy completed chemotherapy and radiotherapy respectively.

There were nine recurrences and 25 breast cancer related deaths over a period of 9 years (2012-2020), with an average recurrence free survival and overall survival of 73 months (Table 8).

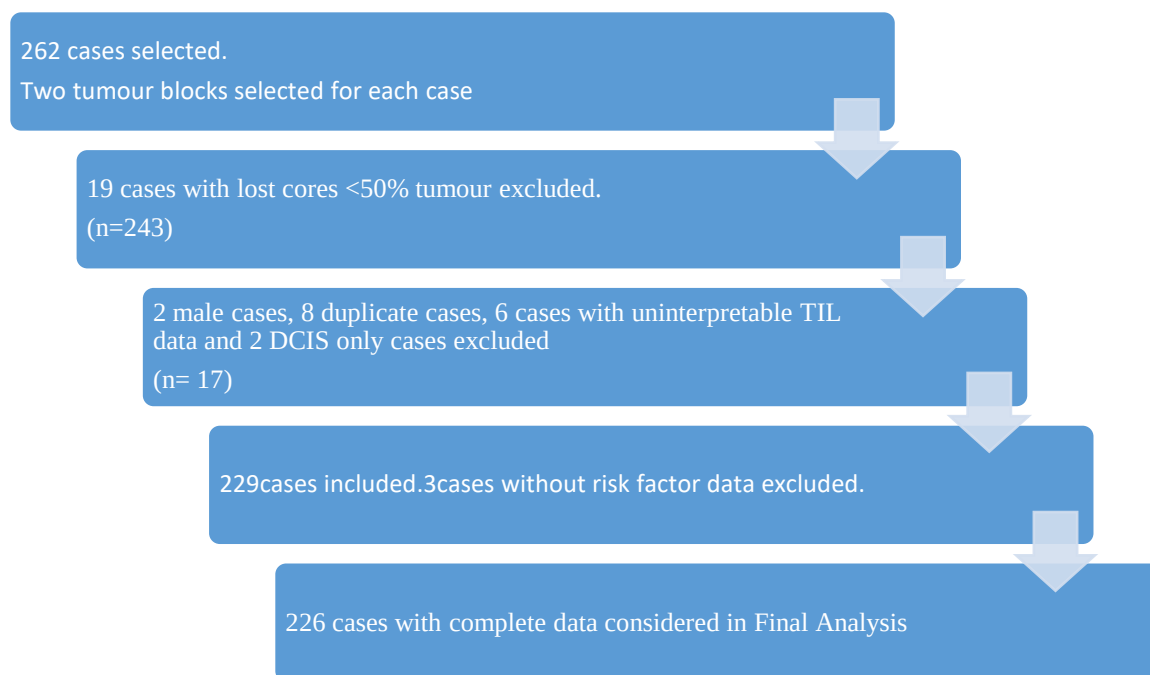
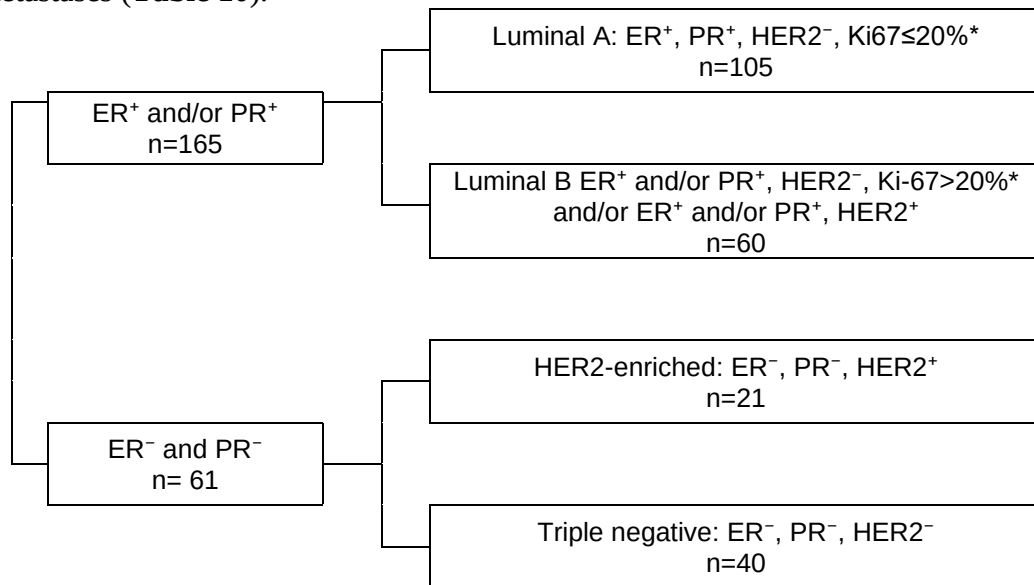


FIGURE 9: CASE SELECTION PROCESS AND FINAL INCLUSION

3.5.2 Distributions of tumour subtypes and pathologic characteristics

The distribution of the tumour subtypes defined by IHC markers is illustrated in **Figure 10** and **Table 10**.

After classifying BC into molecular subtypes, 46.5%, 26.5%, 9.3%, and 17.7 % of patients had luminal A, luminal B, HER2-enriched, and TN breast cancers, respectively (**Figure 10**). Of the 226 patients included in analysis, 49% of patients had tumours larger than 4 cm and had intermediate-to-high tumour grade (intermediate, 41%; high, 54%). Fifty-two percent of cases presented with advanced stage of disease (Stages III/IV) and 74% had lymph node metastases (**Table 10**).



*TUMOR GRADE WAS USED TO DETERMINE TUMOR SUBTYPES IN THE ABSENCE OF KI67: LOW OR INTERMEDIATE TUMOUR GRADE DEFINED AS TUMOR SUBTYPE “LUMINAL A”; HIGH TUMOUR GRADE, DEFINED AS TUMOR SUBTYPE “LUMINAL B HER2-”. ER: ESTROGEN RECEPTOR; HER2: HUMAN EPIDERMAL GROWTH FACTOR RECEPTOR-2; PR: PROGESTERONE RECEPTOR.

FIGURE 10: BREAST TUMOR SUBTYPE DEFINITION IN KENYAN BREAST CANCER PATIENTS (N=226).

3.5.3 Correlation between visual (pathologists) and automated (Halo) TILs

As shown in **Figures 11** and **12**, we observed weak but significant correlations between visual and Halo TILs (correlation coefficients = 0.18 (P = 0.005) and 0.17 (P = 0.010) for intra-tumoural stromal TILs and leading-edge TILs vs Halo, respectively).

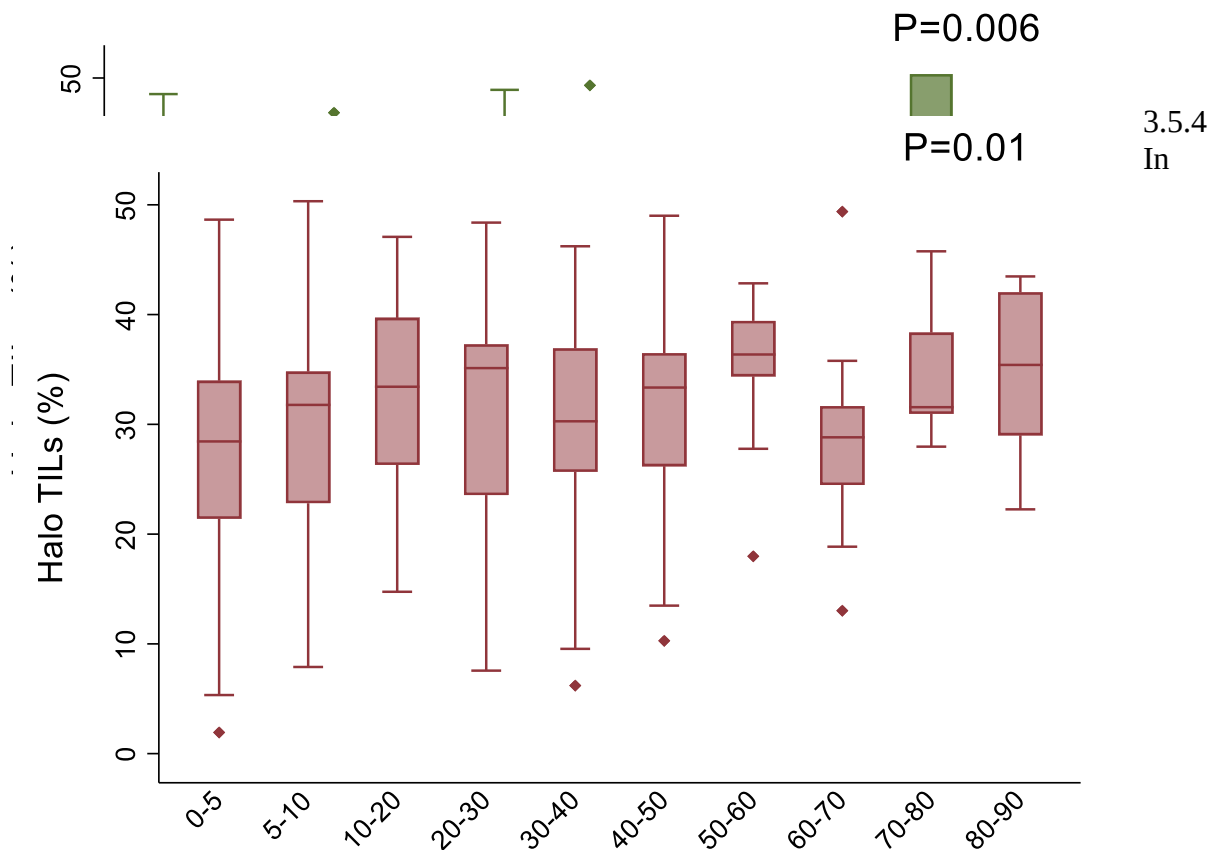


FIGURE 11: VISUAL LEADING EDGE/PERITUMOURAL TIL COUNTS COMPARED WITH HALO TILS

FIGURE 12: VISUAL LEADING INTRATUMOURAL STROMAL TIL COUNTS COMPARED WITH HALO TIL

multivariable-adjusted linear regression models (10% significance level highlighted in Table 8)

Intra-tumoural stromal TILs demonstrated associations with grade (increasing grade associated with increasing TILs), Ki-67 (increasing Ki-67 associated with increasing TILs), and BMI (overweight women had higher TILs than normal weight women).

Leading-edge TILs demonstrated associations with tumour size (increasing tumour size associated with decreasing TILs); grade (increasing grade associated with increasing TILs); and Ki-67 (increasing Ki-67 associated with increasing TILs).

Halo TILs demonstrated associations with breast cancer subtype (luminal B more likely to have higher TILs than luminal A); tumour size (increasing tumour size associated with decreasing TILs); tumour grade (increasing TILs associated with increasing grade); and suggestive associations with pregnancy age (increasing pregnancy age associated with decreasing TILs) and weight (increasing weight associated with increasing TILs).

3.5.5 In multivariable-adjusted logistic regression models (10% significance level highlighted in Table 9)

Halo was used to compare and validate with TILs on H/E since many machine learning. Algorithms are currently being validated for various biomarker analysis- the future role of machine learning has also been acknowledged in the 5th edition of the WHO classification of breast tumours (<https://www.iarc.who.int/news-events/who-classification-of-tumours-5th-edition-volume-2-breast-tumours/>). The Halo TILs detection script is a random forest algorithm that is based on supervised machine learning. It requires expert (pathologists)

labels or annotations to enable it “learn” what constitutes TILs based on parameters that the pathologist, as the expert, would be evaluating such as shape, size, nuclear-cytoplasmic ratio, compactness, density, nuclear contrast threshold, etc. Pathologist training will typically be based only on a subset of the samples e.g., 5%. An important feature of the machine learning algorithm is that it is capable of “learning” from several expert pathologists, with the capacity for continuous standardization. Another important aspect of Halo, or other automated TILs scoring methods, is that it allows high-throughput and standardized counting of TILs. An optimized algorithm can then be used to evaluate TILs on hundreds or thousands of whole slide images within a few hours. In addition to being high-throughput, Halo provides quantitative measures thereby allowing us to refine the prognostic value of TILs amongst outpatients. <https://www.sciencedirect.com/science/article/pii/S0002944020301887#fig1>; <https://ascopubs.org/doi/full/10.1200/CCI.19.00126>

Visual TILs were dichotomized at 10% while Halo TILs score was dichotomized at the median value for the distribution i.e., 31.67%.

Intra-tumoural stromal TILs demonstrated associations with grade (increasing grade associated with increasing grade), KI67 (increasing KI67 associated with increasing TILs), and BMI (overweight women had higher TILs than normal weight women). Leading-edge TILs demonstrated associations with tumour size (increasing tumour size associated with decreasing TILs); KI67 (increasing KI67 associated with increasing TILs); and suggestive associations with cumulative breastfeeding duration. Halo TILs demonstrated associations with breast cancer subtype (luminal B more likely to have higher TILs than luminal A); tumour size (increasing tumour size associated with decreasing TILs); grade (increasing grade associated with increasing TILs); and suggestive associations with pregnancy age (increasing pregnancy age associated with decreasing TILs) and weight (increasing weight associated with increasing TILs).

Note:

Missing value indicators were used to address missingness. Sample size is small, so probably not enough power to detect statistically significant associations. Observed associations will be attenuated if Bonferroni correction for multiple testing were to be applied.

TABLE 8: COMPARISONS OF VISUAL TILS AND HALO TILS BY TUMOUR CHARACTERISTICS

	Visual sTIL		Visual LE-TIL		Halo-TIL	
Characteristic	β	P value	β	P value	β	P value
Age (continuous)	0.02	0.160	0.14	0.269	-0.52	0.489
Subtypes						
Luminal A	Reference		Reference		Reference	
Luminal B	0.17	0.530	0.39	0.235	4.00	0.036
HER2-enriched	0.07	0.851	0.37	0.406	-0.12	0.962
Triple negative	0.37	0.228	-0.06	0.870	1.93	0.375
Size (continuous)/10mm	-0.02	0.861	-0.03	0.008	-0.20	0.002
Grade						
Well-differentiated	Reference		Reference		Reference	
Moderately differentiated	0.39	0.391	0.42	0.428	-1.13	0.709
Poorly differentiated	0.86	0.071	0.93	0.098	2.87	0.368
Trend	0.45	0.015	0.49	0.027	2.79	0.030
KI67 (%)						
<5	Reference		Reference		Reference	
5-10	0.25	0.421	0.37	0.314	2.41	0.256
10-15	0.64	0.036	0.18	0.619	0.27	0.897
15-20	0.33	0.321	0.46	0.248	-2.95	0.204
>20	0.98	0.002	0.74	0.052	1.73	0.439
Trend	0.22	0.003	0.18	0.047	-0.17	0.749
Ethnicity						
Bantu	Reference		Reference		Reference	
Nilote	0.29	0.377	0.17	0.666	1.07	0.642
Cushite	-0.15	0.818	-1.37	0.076	-0.13	0.978
Mixed	0.89	0.173	-0.11	0.890	-6.30	0.170
No of children						
0 or 1	Reference		Reference		Reference	
2	-0.37	0.330	0.15	0.745	-2.87	0.286
3	-0.02	0.957	-0.01	0.993	0.87	0.752
4 or more	-0.19	0.637	-0.36	0.448	-1.24	0.651
Trend	-0.01	0.973	-0.17	0.238	0.11	0.898
Pregnancy age (years)						
Q1 (<20)	Reference		Reference		Reference	
Q2 (20-22.5)	-0.35	0.245	-0.63	0.081	-5.42	0.010
Q3 (22.5-25)	-0.41	0.204	0.10	0.787	-3.44	0.125
Q4 (>25)	-0.30	0.312	-0.13	0.706	-4.44	0.034
Trend	-0.08	0.384	0.05	0.677	-0.83	0.193
Cum. BF duration (months)						
Q1 (<27)	Reference		Reference		Reference	
Q2 (27-46)	-0.23	0.496	-0.94	0.021	0.74	0.745
Q3 (46-60)	-0.41	0.204	-0.19	0.619	1.26	0.567
Q4 (>60)	-0.30	0.312	0.01	0.981	0.08	0.969
Trend	0.16	0.105	0.09	0.459	-0.03	0.963
Menarche						
≤12	Reference		Reference		Reference	
13	0.20	0.664	0.13	0.806	4.61	0.148
14	-0.23	0.606	-0.47	0.379	2.64	0.403
≥15	0.26	0.541	0.10	0.850	1.80	0.542

	Visual sTIL		Visual LE-TIL		Halo-TIL	
Characteristic	β	P value	β	P value	β	P value
Trend	0.09	0.385	0.06	0.659	-0.54	0.471
Weight (Kg)						
Q1 (<62)	Reference		Reference		Reference	
Q2 (62-70)	-0.29	0.361	-0.56	0.144	3.80	0.093
Q3 (70-80)	0.32	0.303	0.19	0.611	4.00	0.068
Q4 (>70)	0.08	0.808	0.18	0.628	4.20	0.056
Trend	0.10	0.377	0.16	0.212	0.58	0.415
Height (cm)						
Q1 (<158)	Reference		Reference		Reference	
Q2 (158-162)	-0.04	0.909	0.08	0.838	-4.69	0.051
Q3 (162-165)	0.81	0.025	0.96	0.029	0.63	0.802
Q4 (>165)	-0.30	0.931	-0.21	0.615	-0.63	0.796
Trend	0.01	0.960	-0.09	0.518	1.15	0.113
BMI						
Normal	Reference		Reference		Reference	
Overweight	0.64	0.014	0.20	0.529	1.43	0.433
Obese	0.23	0.436	0.47	0.196	0.94	0.649
Trend	0.15	0.356	0.21	0.257	0.91	0.371
Family history						
Absent	Reference		Reference		Reference	
Present	-0.30	0.446	0.10	0.832	-4.46	0.109

sTIL (stromal Tumour Infiltrating Lymphocytes); LE-TIL (Leading Edge Tumour infiltrating Lymphocytes).

TABLE 9: COMPARISON OF VISUAL TILS AND HALO TILS (N=226)

	Visual sTIL		Visual LE-TIL		Halo-TIL	
Characteristic	OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value
Age (continuous)	1.04 (0.99, 1.08)	0.059	1.01 (0.97, 1.06)	0.506	0.96 (0.93, 1.00)	0.072
Subtypes						
Luminal A	Reference		Reference		Reference	
Luminal B	1.01 (0.35, 2.90)	0.983	1.71 (0.58, 5.02)	0.329	3.29 (1.15, 9.35)	0.026
HER2-enriched	1.88 (0.50, 7.07)	0.349	2.53 (0.63, 10.11)	0.189	0.87 (0.22, 3.40)	0.839
Triple negative	1.79 (0.56, 5.66)	0.322	1.42 (0.41, 4.88)	0.578	2.22 (0.70, 7.09)	0.178
Size (continuous)/10mm	1.00 (0.97, 1.04)	0.900	0.94 (0.91, 0.98)	0.001	0.93 (0.89, 0.96)	<0.001
Grade						
Well-differentiated	Reference		Reference		Reference	
Moderately differentiated	2.02 (0.27, 15.24)	0.496	1.69 (0.36, 7.94)	0.508	1.21 (0.26, 5.67)	0.808
Poorly differentiated	4.74 (0.59, 38.34)	0.144	2.86 (0.57, 14.70)	0.208	3.36 (0.66, 17.06)	0.144
Trend	2.29 (1.10, 4.77)	0.027	1.69 (0.85, 3.37)	0.134	2.24 (1.12, 4.49)	0.023
KI67 (%)						
<5	Reference		Reference		Reference	
5-10	1.36 (0.41, 4.46)	0.615	2.73 (0.85, 8.82)	0.093	2.79 (0.88, 8.81)	0.080
10-15	3.68 (1.20, 11.34)	0.023	1.30 (0.43, 3.86)	0.642	1.76 (0.60, 5.15)	0.299
15-20	2.21 (0.67, 7.30)	0.192	3.44 (0.84, 14.07)	0.085	0.67 (0.19, 2.31)	0.523
>20	4.43 (1.25, 15.62)	0.021	3.41 (0.93, 12.57)	0.065	1.24 (0.38, 4.05)	0.717
Trend	1.43 (1.08, 1.91)	0.014	1.33 (0.99, 1.79)	0.062	0.94 (0.72, 1.24)	0.680
Ethnicity						
Bantu	Reference		Reference		Reference	
Nilote	1.28 (0.39, 4.16)	0.679	1.22 (0.32, 4.62)	0.767	0.90 (0.2, 2.95)	0.862
Cushite	1.06 (0.06, 18.28)	0.968	0.08 (0.01, 1.04)	0.053	1.17 (0.14, 9.94)	0.886
Mixed	4.38 (0.41, 46.48)	0.221	0.46 (0.03, 6.85)	0.580	0.10 (0.01, 2.21)	0.147
No of children						
0 or 1	Reference		Reference		Reference	
2	0.24 (0.05, 1.12)	0.070	0.88 (0.20, 3.91)	0.862	0.48 (0.12, 1.97)	0.307
3	0.53 (0.12, 2.34)	0.401	0.81 (0.18, 3.59)	0.781	1.62 (0.38, 6.90)	0.518
4 or more	0.35 (0.08, 1.60)	0.177	0.32 (0.07, 1.52)	0.153	1.32 (0.31, 5.64)	0.710
Trend	0.89 (0.57, 1.38)	0.600	0.67 (0.42, 1.08)	0.101	1.31 (0.79, 2.18)	0.288

Characteristic	Visual sTIL		Visual LE-TIL		Halo-TIL	
	OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value
Pregnancy age (years)						
Q1 (<20)	Reference		Reference		Reference	
Q2 (20-22.5)	0.53 (0.18, 1.62)	0.268	0.31 (0.10, 1.03)	0.057	0.47 (0.15, 1.42)	0.180
Q3 (22.5-25)	0.45 (0.14, 1.45)	0.181	2.12 (0.55, 8.23)	0.273	0.33 (0.10, 1.13)	0.077
Q4 (>25)	1.08 (0.36, 3.22)	0.893	1.13 (0.35, 3.62)	0.837	0.21 (0.07, 0.67)	0.008
Trend	1.04 (0.73, 1.47)	0.835	1.27 (0.88, 1.82)	0.190	0.62 (0.43, 0.89)	0.011
Cum. BF duration (months)						
Q1 (<27)	Reference		Reference		Reference	
Q2 (27-46)	0.60 (0.17, 2.08)	0.420	0.16 (0.04, 0.64)	0.009	1.78 (0.52, 6.07)	0.357
Q3 (46-60)	1.10 (0.34, 3.62)	0.872	0.33 (0.09, 1.20)	0.093	3.99 (1.23, 12.93)	0.021
Q4 (>60)	1.54 (0.50, 4.75)	0.449	0.83 (0.25, 2.75)	0.765	1.39 (0.46, 4.19)	0.563
Trend	1.24 (0.87, 1.75)	0.236	1.06 (0.74, 1.50)	0.751	1.21 (0.86, 1.69)	0.277
Menarche						
≤12	Reference		Reference		Reference	
13	1.52 (0.28, 8.30)	0.632	2.05 (0.36, 11.75)	0.418	1.43 (0.25, 8.09)	0.683
14	0.97 (0.18, 5.23)	0.969	0.51 (0.10, 2.67)	0.423	1.72 (0.31, 9.65)	0.539
≥15	2.04 (0.42, 9.87)	0.375	1.31 (0.27, 6.28)	0.733	0.74 (0.15, 3.66)	0.710
Trend	1.17 (0.78, 1.77)	0.450	0.98 (0.66, 1.45)	0.930	0.79 (0.52, 1.14)	0.188
Weight (Kg)						
Q1 (<62)	Reference		Reference		Reference	
Q2 (62-70)	0.78 (0.22, 2.68)	0.690	0.40 (0.12, 1.37)	0.144	1.98 (0.58, 6.67)	0.273
Q3 (70-80)	1.55 (0.46, 5.18)	0.479	1.39 (0.40, 4.88)	0.605	2.88 (0.92, 9.04)	0.069
Q4 (>70)	1.15 (0.35, 3.79)	0.815	1.46 (0.44, 4.87)	0.535	3.05 (0.97, 9.59)	0.056
Trend	1.21 (0.83, 1.78)	0.320	1.29 (0.87, 1.89)	0.205	1.20 (0.81, 1.77)	0.371
Height (cm)						
Q1 (<158)	Reference		Reference		Reference	
Q2 (158-162)	0.73 (0.20, 2.71)	0.638	1.02 (0.29, 3.59)	0.977	0.24 (0.07, 0.90)	0.034
Q3 (162-165)	3.83 (0.96, 15.25)	0.057	5.42 (1.20, 24.51)	0.028	0.51 (0.13, 1.98)	0.327
Q4 (>165)	1.37 (0.38, 4.98)	0.627	0.52 (0.14, 1.88)	0.319	0.47 (0.12, 1.80)	0.272
Trend	1.12 (0.76, 1.65)	0.982	0.82 (0.55, 1.21)	0.322	1.00 (0.67, 1.51)	0.982

Characteristic	Visual sTIL		Visual LE-TIL		Halo-TIL	
	OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value
BMI						
Normal	Reference		Reference		Reference	
Overweight	2.97 (1.09, 8.09)	0.033	1.31 (0.50, 3.47)	0.579	1.26 (0.49, 3.23)	0.628
Obese	1.70 (0.57, 5.10)	0.342	1.85 (0.62, 5.49)	0.268	1.52 (0.55, 4.22)	0.419
Trend	1.38 (0.81, 2.35)	0.235	1.34 (0.76, 2.35)	0.306	1.35 (0.79, 2.31)	0.273
Family history						
Absent	Reference		Reference		Reference	
Present	0.48 (0.09, 2.63)	0.398	1.31 (0.26, 6.60)	0.740	0.30 (0.08, 1.22)	0.093

TABLE 10: ASSOCIATIONS OF RISK FACTORS WITH PERCENT POSITIVE CELLS FOR EACH IHC MARKER IN 226 KENYAN BREAST CANCER WOMEN

Characteristic	CD3						CD4					
	Q2 OR (95% CI)	P value	Q3 OR (95% CI)	P value ^a	Q4 OR (95% CI)	P value ^a	Q2 OR (95% CI)	P value ^a	Q3 OR (95% CI)	P value ^a	Q4 OR (95% CI)	P value ^a
Age at diagnosis (year)												
<40	Reference		Reference		Reference		Reference		Reference		Reference	
40-50	0.61 (0.21-1.82)	0.3789	0.78 (0.26-2.38)	0.6655	1.10 (0.35-3.49)	0.8761	1.91 (0.64-5.72)	0.2476	1.66 (0.60-4.61)	0.3326	0.75 (0.26-2.13)	0.5832
50-60	0.46 (0.08-2.58)	0.3733	0.60 (0.12-3.07)	0.5410	0.22 (0.04-1.20)	0.0796	1.29 (0.24-7.00)	0.7693	3.44 (0.65-18.18)	0.1453	1.07 (0.20-5.84)	0.9406
60+	0.78 (0.12-5.05)	0.7937	0.32 (0.05-2.13)	0.2385	0.22 (0.03-1.45)	0.1148	1.76 (0.28-11.24)	0.5495	3.33 (0.50-22.24)	0.2153	1.63 (0.25-10.52)	0.6057
Body mass index (kg/m²)												
<25	Reference		Reference		Reference		Reference		Reference		Reference	
25+	0.87 (0.36-2.09)	0.7550	0.94 (0.37-2.37)	0.8991	0.94 (0.37-2.41)	0.9026	1.25 (0.50-3.17)	0.6320	0.83 (0.35-2.01)	0.6812	0.95 (0.38-2.38)	0.9092
Age at menarche (year)												
<15	Reference		Reference		Reference		Reference		Reference		Reference	
15+	0.39 (0.17-0.87)	0.0220	0.67 (0.30-1.49)	0.3249	2.37 (1.02-5.51)	0.0445	1.83 (0.83-4.05)	0.1350	1.37 (0.63-2.98)	0.4303	0.97 (0.44-2.14)	0.9318
Number of children												
0	Reference		Reference		Reference		Reference		Reference		Reference	
1	1.46 (0.16-13.24)	0.7349	1.66 (0.19-14.56)	0.6493	1.17 (0.12-11.76)	0.8933	-		0.71 (0.10-4.79)	0.7230	0.97 (0.11-8.61)	0.9788
2	2.66 (0.35-20.42)	0.3459	1.61 (0.22-12.10)	0.6418	1.84 (0.24-14.15)	0.5592	-		0.81 (0.14-4.54)	0.8063	1.27 (0.17-9.47)	0.8139
3+	2.36 (0.33-16.98)	0.3935	2.47 (0.36-16.82)	0.3553	2.55 (0.36-17.87)	0.3468	-		0.65 (0.13-3.33)	0.6063	1.07 (0.16-7.37)	0.9453

Characteristic	CD3						CD4					
	Q2 OR (95% CI)	P value	Q3 OR (95% CI)	P value ^a	Q4 OR (95% CI)	P value ^a	Q2 OR (95% CI)	P value ^a	Q3 OR (95% CI)	P value ^a	Q4 OR (95% CI)	P value ^a
Age at first pregnancy^b (year)												
<20	Reference		Reference		Reference		Reference		Reference		Reference	
20-25	0.98 (0.34-2.81)	0.9666	2.23 (0.72-6.91)	0.1647	3.06 (0.95-9.91)	0.0618	2.36 (0.79-7.07)	0.1264	1.54 (0.51-4.58)	0.4429	1.80 (0.61-5.30)	0.2878
25+	0.73 (0.24-2.21)	0.5787	1.14 (0.35-3.77)	0.8296	1.01 (0.28-3.66)	0.9890	0.82 (0.25-2.68)	0.7474	0.90 (0.30-2.77)	0.8584	0.66 (0.20-2.16)	0.4964
Cumulative duration of breastfeeding^c (month)												
<48	Reference		Reference		Reference		Reference		Reference		Reference	
48-72	1.23 (0.40-3.82)	0.7178	1.29 (0.40-4.20)	0.6730	1.02 (0.32-3.24)	0.9798	0.93 (0.31-2.80)	0.8935	0.72 (0.24-2.11)	0.5440	1.55 (0.50-4.75)	0.4469
72+	0.51 (0.16-1.62)	0.2512	1.02 (0.31-3.33)	0.9717	0.37 (0.11-1.20)	0.0971	0.71 (0.23-2.20)	0.5476	0.46 (0.15-1.43)	0.1794	1.06 (0.32-3.50)	0.9186
Menopausal status												
Pre	Reference		Reference		Reference		Reference		Reference		Reference	
Post	0.50 (0.12-2.05)	0.3338	0.60 (0.16-2.19)	0.4390	1.47 (0.43-4.99)	0.5403	1.31 (0.39-4.43)	0.6681	0.52 (0.14-1.96)	0.3359	1.04 (0.27-4.03)	0.9518

a Modelling each IHC marker measurement as the outcome variable and represented by its quartiles; models were adjusted for age at diagnosis, body mass index, age at menarche, number of children, menopausal status, and ER status.

b Among parous women, only. Model was adjusted for age at diagnosis, body mass index, age at menarche, number of children, cumulative duration of breastfeeding, and menopausal status.

c Among women who breastfed. Model was adjusted for age at diagnosis, body mass index, age at menarche, number of children, age at first pregnancy, and menopausal status.

Bold indicates $p \leq 0.05$

Red font in bold indicates $p \leq 0.007$ (correcting for multiple tests)

Characteristic	CD8						CD68					
	Q2 OR (95% CI)	P value ^a	Q3 OR (95% CI)	P value ^a	Q4 OR (95% CI)	P value ^a	Q2 OR (95% CI)	P value ^a	Q3 OR (95% CI)	P value ^a	Q4 OR (95% CI)	P value ^a
Age at diagnosis (year)												
<40	Reference		Reference		Reference		Reference		Reference		Reference	
40-50	1.93 (0.62-6.01)	0.2540	1.53 (0.51-4.55)	0.4456	1.29 (0.44-3.78)	0.6485	0.49 (0.16-1.52)	0.2170	1.33 (0.44-3.96)	0.6137	0.76 (0.26-2.25)	0.6168
50-60	1.04 (0.19-5.71)	0.9663	0.94 (0.17-5.17)	0.9415	1.03 (0.21-5.14)	0.9750	0.34 (0.06-2.00)	0.2345	0.27 (0.04-1.72)	0.1637	0.20 (0.03-1.23)	0.0821
60+	1.92 (0.31-12.13)	0.4866	1.90 (0.30-12.24)	0.4995	0.56 (0.09-3.62)	0.5396	0.18 (0.03-1.29)	0.0878	0.22 (0.03-1.65)	0.1392	0.07 (0.01-0.53)	0.0100
Body mass index (kg/m²)												
<25	Reference		Reference		Reference		Reference		Reference		Reference	
25+	0.91 (0.35-2.33)	0.8431	0.57 (0.23-1.43)	0.2314	0.82 (0.33-2.01)	0.6596	0.62 (0.24-1.61)	0.3290	0.84 (0.32-2.17)	0.7175	0.67 (0.26-1.73)	0.4110
Age at menarche (year)												
<15	Reference		Reference		Reference		Reference		Reference		Reference	
15+	0.86 (0.39-1.91)	0.7130	1.06 (0.48-2.34)	0.8941	1.48 (0.67-3.24)	0.3326	2.50 (1.09-5.75)	0.0304	1.69 (0.74-3.86)	0.2129	2.12 (0.93-4.83)	0.0743
Number of children												
0	Reference		Reference		Reference		Reference		Reference		Reference	
1	6.65 (0.45-99.50)	0.1697	17.30 (1.26-236.80)	0.0327	3.15 (0.39-25.57)	0.2839	-		2.15 (0.33-13.80)	0.4216	4.54 (0.35-59.27)	0.2484
2	3.82 (0.35-41.25)	0.2692	4.72 (0.44-50.99)	0.2012	1.83 (0.33-10.30)	0.4916	-		0.85 (0.17-4.29)	0.8426	2.91 (0.28-30.29)	0.3722
3+	3.82 (0.38-38.50)	0.2557	4.93 (0.49-49.29)	0.1741	1.46 (0.29-7.51)	0.6492	-		0.94 (0.20-4.49)	0.9332	5.98 (0.60-59.09)	0.1262

Characteristic	CD8						CD68					
	Q2 OR (95% CI)	P value ^a	Q3 OR (95% CI)	P value ^a	Q4 OR (95% CI)	P value ^a	Q2 OR (95% CI)	P value ^a	Q3 OR (95% CI)	P value ^a	Q4 OR (95% CI)	P value ^a
Age at first pregnancy^b (year)												
<20	Reference		Reference		Reference		Reference		Reference		Reference	
20-25	1.52 (0.53-4.37)	0.4411	2.89 (0.98-8.49)	0.0543	3.38 (1.06-10.74)	0.0394	1.27 (0.42-3.84)	0.6764	0.49 (0.16-1.50)	0.2103	1.41 (0.42-4.72)	0.5729
25+	0.63 (0.20-2.01)	0.4379	1.33 (0.42-4.21)	0.6327	1.74 (0.52-5.82)	0.3678	1.03 (0.31-3.50)	0.9575	0.57 (0.17-1.88)	0.3521	1.46 (0.40-5.25)	0.5668
Cumulative duration of breastfeeding^c (month)												
<48	Reference		Reference		Reference		Reference		Reference		Reference	
48-72	0.75 (0.23-2.49)	0.6420	1.32 (0.44-4.00)	0.6256	2.44 (0.79-7.54)	0.1204	0.55 (0.18-1.71)	0.2991	0.31 (0.10-0.96)	0.0419	0.30 (0.10-0.93)	0.0362
72+	1.58 (0.50-5.06)	0.4380	1.19 (0.38-3.70)	0.7627	1.64 (0.48-5.61)	0.4272	0.64 (0.19-2.14)	0.4647	0.43 (0.13-1.46)	0.1755	0.42 (0.12-1.44)	0.1679
Menopausal status												
Pre	Reference		Reference		Reference		Reference		Reference		Reference	
Post	0.56 (0.15-2.04)	0.3781	0.37 (0.10-1.42)	0.1472	0.81 (0.23-2.80)	0.7332	2.07 (0.44-9.66)	0.3541	4.15 (0.90-19.19)	0.0686	4.33 (0.95-19.65)	0.0579

modelling each IHC marker measurement as the outcome variable and represented by its quartiles; models were adjusted for age at diagnosis, body mass index, age at menarche, number of children, menopausal status, and ER status.

^bAmong parous women, only. Model was adjusted for age at diagnosis, body mass index, age at menarche, number of children, cumulative duration of breastfeeding, and menopausal status.

^cAmong women who breastfed. Model was adjusted for age at diagnosis, body mass index, age at menarche, number of children, age at first pregnancy, and menopausal status.

Bold indicates $p \leq 0.05$

Red font in bold indicates $p \leq 0.007$ (correcting for multiple tests)

	CD20						FOX-P3					
Characteristic	Q2 OR (95% CI)	P value ^a	Q3 OR (95% CI)	P value ^a	Q4 OR (95% CI)	P value ^a	Q2 OR (95% CI)	P value ^a	Q3 OR (95% CI)	P value ^a	Q4 OR (95% CI)	P value ^a
Age at diagnosis (year)												
<40		Reference		Reference		Reference		Reference		Reference		Reference
40-50	4.05 (1.16-14.12)	0.0281	4.39 (1.39-13.82)	0.0115	3.56 (1.01-12.56)	0.0486	1.30 (0.48-3.51)	0.6011	1.81 (0.63-5.20)	0.2694	1.88 (0.62-5.73)	0.2676
50-60	2.89 (0.46-18.22)	0.2597	4.76 (0.84-26.92)	0.0778	3.47 (0.52-23.16)	0.1989	0.99 (0.19-5.11)	0.9885	2.50 (0.48-13.06)	0.2774	5.38 (1.00-29.03)	0.0506
60+	3.94 (0.54-28.94)	0.1773	4.12 (0.61-27.93)	0.1469	7.93 (1.00-63.14)	0.0505	2.81 (0.43-18.64)	0.2836	7.80 (1.15-53.17)	0.0359	10.77 (1.49-78.01)	0.0187
Body mass index (kg/m²)												
<25		Reference		Reference		Reference		Reference		Reference		Reference
25+	1.19 (0.43-3.28)	0.7383	0.66 (0.26-1.65)	0.3689	1.44 (0.51-4.10)	0.4952	1.03 (0.42-2.55)	0.9467	1.65 (0.64-4.23)	0.2975	1.64 (0.65-4.17)	0.2992
Age at menarche (year)												
<15		Reference		Reference		Reference		Reference		Reference		Reference
15+	1.25 (0.53-2.95)	0.6062	0.73 (0.32-1.67)	0.4559	2.66 (1.09-6.51)	0.0318	1.14 (0.52-2.49)	0.7473	1.22 (0.55-2.70)	0.6243	1.18 (0.53-2.66)	0.6873
Number of children												
0		Reference		Reference		Reference		Reference		Reference		Reference
1	8.00 (0.61-105.20)	0.1137	3.81 (0.29-49.83)	0.3077	1.51 (0.06-35.90)	0.7998	0.24 (0.02-2.79)	0.2567	1.29 (0.06-27.84)	0.8719	0.24 (0.01-4.20)	0.3306
2	21.04 (1.66-267.16)	0.0188	15.16 (1.26-182.62)	0.0323	33.02 (2.54-429.82)	0.0076	0.13 (0.01-1.36)	0.0891	1.45 (0.08-27.51)	0.8042	0.66 (0.05-8.70)	0.7553
3+	2.44 (0.24-24.75)	0.4505	3.21 (0.34-30.78)	0.3121	2.99 (0.29-31.29)	0.3597	0.16 (0.02-1.52)	0.1109	0.69 (0.04-12.30)	0.7998	0.22 (0.02-2.74)	0.2384

Age at first pregnancy ^b (year)												
<20		Reference		Reference		Reference		Reference		Reference		Reference
20-25	1.28 (0.40-4.18)	0.6778	1.03 (0.34-3.10)	0.9535	1.46 (0.46-4.65)	0.5184	0.72 (0.24-2.14)	0.5579	1.01 (0.32-3.17)	0.9816	1.00 (0.34-2.92)	0.9948
25+	0.82 (0.21-3.16)	0.7691	1.57 (0.49-5.04)	0.4500	0.76 (0.20-2.93)	0.6945	0.79 (0.24-2.61)	0.6987	1.98 (0.59-6.59)	0.2660	0.43 (0.12-1.54)	0.1968
Cumulative duration of breastfeeding ^c (month)												
<48		Reference		Reference		Reference		Reference		Reference		Reference
48-72	0.78 (0.23-2.68)	0.6891	1.74 (0.56-5.44)	0.3431	1.44 (0.43-4.81)	0.5560	2.64 (0.87-7.97)	0.0860	0.86 (0.27-2.67)	0.7873	3.95 (1.17-13.32)	0.0271
72+	1.33 (0.37-4.76)	0.6575	1.65 (0.47-5.75)	0.4307	1.46 (0.41-5.28)	0.5625	0.98 (0.30-3.21)	0.9678	0.65 (0.22-1.95)	0.4465	2.36 (0.68-8.16)	0.1769
Menopausal status												
Pre		Reference		Reference		Reference		Reference		Reference		Reference
Post	1.40 (0.36-5.45)	0.6297	0.64 (0.17-2.45)	0.5155	0.71 (0.17-2.95)	0.6318	1.07 (0.31-3.69)	0.9120	1.04 (0.30-3.61)	0.9562	1.10 (0.30-4.00)	0.8821

modelling each IHC marker measurement as the outcome variable and represented by its quartiles; models were adjusted for age at diagnosis, body mass index, age at menarche, number of children, menopausal status, and ER status.

*b*Among parous women, only. Model was adjusted for age at diagnosis, body mass index, age at menarche, number of children, cumulative duration of breastfeeding, and menopausal status.

*c*Among women who breastfed. Model was adjusted for age at diagnosis, body mass index, age at menarche, number of children, age at first pregnancy, and menopausal status.

Bold indicates $p \leq 0.05$

Red font in bold indicates $p \leq 0.007$ (correcting for multiple tests)

TIL				
Characteristic	10-30 vs <10%	P value ^a	>30 vs <10%	P value ^a
Age at diagnosis (year)				
<40	Reference		Reference	
40-50	2.02 (0.86-4.74)	0.1070	1.82 (0.66-5.03)	0.2470
50-60	1.91 (0.50-7.28)	0.3443	1.75 (0.38-8.10)	0.4758
60+	1.14 (0.25-5.14)	0.8694	1.65 (0.31-8.71)	0.5566
Body mass index (kg/m²)				
<25	Reference		Reference	
25+	1.32 (0.62-2.82)	0.4787	1.96 (0.85-4.51)	0.1147
Age at menarche (year)				
<15	Reference		Reference	
15+	0.71 (0.37-1.37)	0.3014	1.58 (0.76-3.25)	0.2186
Number of children				
0	Reference		Reference	
1	0.95 (0.16-5.67)	0.9574	0.41 (0.04-4.45)	0.4656
2	1.25 (0.24-6.58)	0.7925	1.57 (0.22-11.06)	0.6510
3+	1.09 (0.22-5.37)	0.9190	1.42 (0.22-9.17)	0.7150
Age at first pregnancy^b (year)				
<20	Reference		Reference	
20-25	0.83 (0.32-2.12)	0.6906	1.08 (0.41-2.82)	0.8751
25+	1.28 (0.47-3.49)	0.6288	1.12 (0.37-3.33)	0.8434
Cumulative duration of breastfeeding^c (month)				

TIL				
Characteristic	10-30 vs <10%	P value ^a	>30 vs <10%	P value ^a
<48	Reference		Reference	
48-72	1.35 (0.54-3.41)	0.5238	1.24 (0.47-3.31)	0.6617
72+	2.37 (0.87-6.42)	0.0906	1.45 (0.52-4.05)	0.4780
Menopausal status				
Pre	Reference		Reference	
Post	0.78 (0.27-2.22)	0.6384	1.17 (0.37-3.71)	0.7924

modelling each IHC marker measurement as the outcome variable and represented by its quartiles; models were adjusted for age at diagnosis, body mass index, age at menarche, number of children, menopausal status, and ER status.

*b*Among parous women, only. Model was adjusted for age at diagnosis, body mass index, age at menarche, number of children, cumulative duration of breastfeeding, and menopausal status.

*c*Among women who breastfed. Model was adjusted for age at diagnosis, body mass index, age at menarche, number of children, age at first pregnancy, and menopausal status.

Bold indicates $p \leq 0.05$

Red font in bold indicates $p \leq 0.007$ (correcting for multiple tests)

TABLE 11: ASSOCIATIONS OF TUMOR CHARACTERISTICS WITH PERCENT POSITIVE NUCLEI FOR EACH IHC MARKER IN 226 KENYAN BREAST CANCER WOMEN

Characteristic	CD3						CD4					
	Q2 OR (95% CI)	P value ^a	Q3 OR (95% CI)	P value ^a	Q4 OR (95% CI)	P value ^a	Q2 OR (95% CI)	P value ^a	Q3 OR (95% CI)	P value ^a	Q4 OR (95% CI)	P value ^a
Oestrogen receptor status												
Negative	Reference		Reference		Reference		Reference		Reference		Reference	
Positive	0.89 (0.22-3.50)	0.8613	0.98 (0.25-3.89)	0.9724	0.87 (0.21-3.67)	0.8519	1.45 (0.36-5.79)	0.5973	1.04 (0.24-4.49)	0.9572	0.18 (0.03-1.17)	0.0719
Progesterone receptor status^b												
Negative	Reference		Reference		Reference		Reference		Reference		Reference	
Positive	2.34 (0.65-8.39)	0.1932	1.23 (0.34-4.45)	0.7518	2.08 (0.55-7.87)	0.2791	0.34 (0.10-1.18)	0.0881	1.04 (0.29-3.77)	0.9568	2.71 (0.45-16.28)	0.2761
Human epidermal growth factor receptor 2 status^b												
Negative	Reference		Reference		Reference		Reference		Reference		Reference	
Positive	1.67 (0.67-4.16)	0.2698	0.67 (0.26-1.74)	0.4074	0.86 (0.33-2.25)	0.7590	0.98 (0.36-2.64)	0.9647	1.90 (0.73-4.94)	0.1885	1.92 (0.71-5.23)	0.2005
KI67^c (%)												
<20	Reference		Reference		Reference		Reference		Reference		Reference	
20+	0.64 (0.25-1.60)	0.3361	0.79 (0.32-1.95)	0.6073	1.15 (0.47-2.80)	0.7549	2.57 (1.03-6.40)	0.0424	1.18 (0.46-3.06)	0.7277	1.52 (0.57-3.99)	0.4014
Tumour subtype^c												
Luminal A	Reference		Reference		Reference		Reference		Reference		Reference	
Luminal B	1.19 (0.43-3.30)	0.7398	0.32 (0.10-1.07)	0.0648	0.46 (0.15-1.44)	0.1824	1.29 (0.41-4.07)	0.6606	2.05 (0.70-5.97)	0.1888	2.17 (0.67-6.97)	0.1944

CD3							CD4					
	Q2 OR (95% CI)	P value ^a	Q3 OR (95% CI)	P value ^a	Q4 OR (95% CI)	P value ^a	Q2 OR (95% CI)	P value ^a	Q3 OR (95% CI)	P value ^a	Q4 OR (95% CI)	P value ^a
HER2- enriched	1.09 (0.25- 4.79)	0.9075	1.09 (0.26- 4.56)	0.9103	0.87 (0.20- 3.81)	0.8522	2.29 (0.50- 10.56)	0.2882	1.74 (0.35- 8.67)	0.4966	3.75 (0.80- 17.45)	0.0925
Triple negative	0.57 (0.19- 1.71)	0.3181	0.50 (0.17- 1.49)	0.2148	0.36 (0.11- 1.12)	0.0781	2.14 (0.73- 6.31)	0.1685	0.99 (0.28- 3.43)	0.9833	2.74 (0.87- 8.69)	0.0864
Tumour size ^c (mm)												
<2	Reference		Reference		Reference		Reference		Reference		Reference	
2+	0.75 (0.33- 1.71)	0.5001	0.71 (0.30- 1.67)	0.4364	0.78 (0.34- 1.80)	0.5583	0.64 (0.28- 1.50)	0.3070	1.14 (0.49- 2.66)	0.7548	1.03 (0.43- 2.47)	0.9390
Tumour grade ^c												
1/2	Reference		Reference		Reference		Reference		Reference		Reference	
3	1.25 (0.53- 2.95)	0.6058	1.25 (0.52- 3.00)	0.6266	1.24 (0.52- 2.97)	0.6233	0.98 (0.42- 2.30)	0.9710	0.74 (0.31- 1.73)	0.4811	0.47 (0.19- 1.17)	0.1040
TNM stage ^c												
I/II	Reference		Reference		Reference		Reference		Reference		Reference	
III/IV	0.81 (0.31- 2.12)	0.6610	3.29 (1.07- 10.07)	0.0373	0.94 (0.35- 2.53)	0.8973	1.27 (0.48- 3.35)	0.6328	1.05 (0.40- 2.78)	0.9253	1.22 (0.43- 3.49)	0.7088
Lymph nodal involvement ^c												
No	Reference		Reference		Reference		Reference		Reference		Reference	
Yes	1.12 (0.37- 3.41)	0.8362	0.33 (0.10- 1.10)	0.0711	1.06 (0.34- 3.32)	0.9249	0.49 (0.15- 1.62)	0.2410	0.62 (0.18- 2.19)	0.4608	0.18 (0.06- 0.62)	0.0062

modelling each IHC marker measurement as the outcome variable and represented by its quartiles.

*b*Model includes ER, PR, and HER2 status, KI67, tumour size, grade, TNM stage, lymph nodal involvement, and age at diagnosis.

*c*Model includes tumour subtype, tumour size, grade, TNM stage, lymph nodal involvement, and age at diagnosis.

Characteristic	CD8						CD68					
	Q2 OR (95% CI)	P value ^a	Q3 OR (95% CI)	P value ^a	Q4 OR (95% CI)	P value ^a	Q2 OR (95% CI)	P value ^a	Q3 OR (95% CI)	P value ^a	Q4 OR (95% CI)	P value ^a
Oestrogen receptor status^b												
Negative	Reference		Reference		Reference		Reference		Reference		Reference	
Positive	0.94 (0.19-4.57)	0.9339	1.88 (0.42-8.45)	0.4113	1.62 (0.34-7.67)	0.5467	0.46 (0.11-2.01)	0.3034	0.36 (0.08-1.65)	0.1880	0.71 (0.16-3.10)	0.6433
Progesterone receptor status^b												
Negative	Reference		Reference		Reference		Reference		Reference		Reference	
Positive	0.22 (0.05-0.97)	0.0460	0.33 (0.08-1.32)	0.1176	0.49 (0.12-2.04)	0.3245	1.10 (0.30-4.07)	0.8835	1.41 (0.35-5.59)	0.6291	1.03 (0.28-3.71)	0.9693
Human epidermal growth factor receptor 2 status^b												
Negative	Reference		Reference		Reference		Reference		Reference		Reference	
Positive	0.30 (0.11-0.85)	0.0233	0.77 (0.31-1.92)	0.5798	0.62 (0.25-1.55)	0.3071	0.88 (0.34-2.27)	0.7922	1.14 (0.44-2.97)	0.7847	0.44 (0.16-1.25)	0.1225
KI67^c (%)												
<20	Reference		Reference		Reference		Reference		Reference		Reference	
20+	0.64 (0.23-1.79)	0.3946	1.74 (0.71-4.27)	0.2250	1.20 (0.49-2.95)	0.6965	2.67 (0.92-7.75)	0.0708	4.01 (1.41-11.42)	0.0092	2.65 (0.91-7.68)	0.0737
Tumour subtype^c												
Luminal A	Reference		Reference		Reference		Reference		Reference		Reference	
Luminal B	0.40 (0.11-1.49)	0.1714	1.02 (0.38-2.78)	0.9662	0.50 (0.17-1.44)	0.1979	1.04 (0.35-3.07)	0.9465	1.47 (0.51-4.27)	0.4750	0.29 (0.08-1.02)	0.0545
HER2-enriched	1.37 (0.34-5.57)	0.6609	1.09 (0.25-4.82)	0.9141	1.01 (0.26-3.98)	0.9907	1.78 (0.42-7.53)	0.4365	2.13 (0.50-9.15)	0.3103	0.92 (0.21-4.07)	0.9090
Triple negative	2.94 (0.93-9.32)	0.0667	1.48 (0.46-4.81)	0.5105	0.63 (0.17-2.24)	0.4712	2.91 (0.83-10.28)	0.0967	3.53 (0.97-12.87)	0.0565	1.07 (0.29-3.97)	0.9146

Characteristic	CD8						CD68					
	Q2 OR (95% CI)	P value ^a	Q3 OR (95% CI)	P value ^a	Q4 OR (95% CI)	P value ^a	Q2 OR (95% CI)	P value ^a	Q3 OR (95% CI)	P value ^a	Q4 OR (95% CI)	P value ^a
Tumour size^c (mm)												
<2	Reference		Reference		Reference		Reference		Reference		Reference	
2+	0.55 (0.23- 1.30)	0.1754	1.01 (0.44- 2.33)	0.9776	0.83 (0.36- 1.89)	0.6523	0.84 (0.37- 1.94)	0.6855	0.81 (0.35- 1.89)	0.6274	1.02 (0.43- 2.42)	0.9563
Tumour grade^c												
1/2	Reference		Reference		Reference		Reference		Reference		Reference	
3	0.56 (0.22- 1.40)	0.2153	0.98 (0.42- 2.30)	0.9598	1.50 (0.64- 3.52)	0.3549	1.65 (0.68- 4.01)	0.2672	1.45 (0.59- 3.56)	0.4188	3.91 (1.57- 9.73)	0.0034
TNM stage^c												
I/II	Reference		Reference		Reference		Reference		Reference		Reference	
III/IV	1.75 (0.58- 5.25)	0.3178	0.59 (0.22- 1.55)	0.2815	1.01 (0.37- 2.78)	0.9868	0.74 (0.26- 2.09)	0.5650	0.75 (0.26- 2.16)	0.5964	0.57 (0.20- 1.61)	0.2848
Lymph nodal involvement^c												
No	Reference		Reference		Reference		Reference		Reference		Reference	
Yes	0.34 (0.10- 1.16)	0.0855	1.81 (0.56- 5.91)	0.3239	0.64 (0.20- 2.07)	0.4592	1.42 (0.46- 4.42)	0.5432	1.16 (0.37- 3.66)	0.7948	4.28 (1.17- 15.68)	0.0280

modelling each IHC marker measurement as the outcome variable and represented by its quartiles.

*b*Model includes ER, PR, and HER2 status, KI67, tumour size, grade, TNM stage, lymph nodal involvement, and age at diagnosis.

*c*Model includes tumour subtype, tumour size, grade, TNM stage, lymph nodal involvement, and age at diagnosis.

Characteristic	CD20						FOX-P3					
	Q2 OR (95% CI)	P value ^a	Q3 OR (95% CI)	P value ^a	Q4 OR (95% CI)	P value ^a	Q2 OR (95% CI)	P value ^a	Q3 OR (95% CI)	P value ^a	Q4 OR (95% CI)	P value ^a
Oestrogen receptor status^b												
Negative	Reference		Reference		Reference		Reference		Reference		Reference	
Positive	0.73 (0.15-3.61)	0.7020	0.91 (0.21-3.97)	0.9001	0.56 (0.12-2.66)	0.4631	0.51 (0.12-2.10)	0.3474	0.84 (0.20-3.50)	0.8135	0.44 (0.10-1.90)	0.2691
Progesterone receptor status^b												
Negative	Reference		Reference		Reference		Reference		Reference		Reference	
Positive	1.24 (0.29-5.32)	0.7732	0.69 (0.18-2.60)	0.5782	0.94 (0.22-4.05)	0.9356	1.89 (0.52-6.87)	0.3352	1.07 (0.30-3.83)	0.9127	1.41 (0.36-5.50)	0.6256
Human epidermal growth factor receptor 2 status^b												
Negative	Reference		Reference		Reference		Reference		Reference		Reference	
Positive	1.11 (0.39-3.18)	0.8437	1.80 (0.66-4.93)	0.2536	2.43 (0.88-6.71)	0.0868	0.98 (0.37-2.62)	0.9716	1.62 (0.63-4.18)	0.3144	1.20 (0.45-3.15)	0.7161
KI67^c (%)												
<20	Reference		Reference		Reference		Reference		Reference		Reference	
20+	0.95 (0.37-2.42)	0.9154	0.66 (0.25-1.74)	0.3980	1.58 (0.63-4.01)	0.3326	0.58 (0.23-1.48)	0.2553	1.16 (0.47-2.87)	0.7535	1.35 (0.55-3.34)	0.5133
Tumour subtype^c												
Luminal A	Reference		Reference		Reference		Reference		Reference		Reference	
Luminal B	1.22 (0.34-4.42)	0.7620	3.81 (1.17-12.41)	0.0267	3.99 (1.17-13.61)	0.0271	1.05 (0.34-3.18)	0.9376	1.36 (0.45-4.06)	0.5842	1.22 (0.38-3.92)	0.7338
HER2-enriched	1.01 (0.22-4.56)	0.9886	0.98 (0.16-6.09)	0.9801	2.76 (0.66-11.56)	0.1657	0.68 (0.13-3.50)	0.6407	1.76 (0.42-7.38)	0.4421	1.73 (0.41-7.33)	0.4574
Triple negative	1.26 (0.37-4.32)	0.7153	2.73 (0.82-9.02)	0.1006	2.89 (0.89-9.39)	0.0775	1.19 (0.39-3.66)	0.7604	0.74 (0.21-2.65)	0.6468	1.85 (0.60-5.72)	0.2862

Tumour size ^c (mm)												
<2	Reference		Reference		Reference		Reference		Reference		Reference	
2+	0.51 (0.21-1.22)	0.1288	0.55 (0.23-1.31)	0.1764	0.42 (0.17-1.02)	0.0554	0.82 (0.36-1.90)	0.6473	0.44 (0.18-1.06)	0.0680	0.49 (0.20-1.16)	0.1038
Tumour grade ^c												
1/2	Reference		Reference		Reference		Reference		Reference		Reference	
3	0.99 (0.39-2.47)	0.9771	0.55 (0.22-1.38)	0.2032	0.98 (0.38-2.55)	0.9681	1.61 (0.68-3.83)	0.2778	1.42 (0.59-3.42)	0.4374	2.10 (0.85-5.15)	0.1063
TNM stage ^c												
I/II	Reference		Reference		Reference		Reference		Reference		Reference	
III/IV	0.81 (0.28-2.38)	0.7049	0.42 (0.15-1.21)	0.1074	0.73 (0.24-2.27)	0.5886	0.79 (0.30-2.14)	0.6466	1.20 (0.42-3.39)	0.7352	1.10 (0.39-3.11)	0.8603
Lymph nodal involvement ^c												
No	Reference		Reference		Reference		Reference		Reference		Reference	
Yes	1.53 (0.40-5.82)	0.5325	1.42 (0.43-4.67)	0.5664	0.62 (0.18-2.15)	0.4530	1.19 (0.36-3.98)	0.7743	0.56 (0.17-1.84)	0.3400	0.66 (0.20-2.22)	0.5063

modelling each IHC marker measurement as the outcome variable and represented by its quartiles.

bModel includes ER, PR, and HER2 status, KI67, tumour size, grade, TNM stage, lymph nodal involvement, and age at diagnosis.

cModel includes tumour subtype, tumour size, grade, TNM stage, lymph nodal involvement, and age at diagnosis.

Characteristic	TIL			
	10-30 vs <10%	P value ^a	>30 vs <10%	P value ^a
Oestrogen receptor status^b				
Negative	Reference		Reference	
Positive	0.94 (0.28-3.22)	0.9251	0.48 (0.13-1.81)	0.2746
Progesterone receptor status^b				
Negative	Reference		Reference	
Positive	0.89 (0.30-2.67)	0.8346	0.96 (0.28-3.38)	0.9532
Human epidermal growth factor receptor 2 status^b				
Negative	Reference		Reference	
Positive	1.22 (0.54-2.75)	0.6270	1.56 (0.66-3.67)	0.3131
KI67^c (%)				
<20	Reference		Reference	
20+	1.49 (0.66-3.35)	0.3337	2.97 (1.30-6.82)	0.0101
Tumour subtype^c				
Luminal A	Reference		Reference	
Luminal B	1.15 (0.45-2.93)	0.7743	2.14 (0.79-5.79)	0.1345
HER2-enriched	1.58 (0.45-5.52)	0.4748	2.71 (0.73-10.13)	0.1381
Triple negative	1.05 (0.38-2.93)	0.9264	2.97 (1.10-8.01)	0.0315
Tumour size^c (mm)				
<2	Reference		Reference	
2+	0.29 (0.14-0.61)	0.0011	0.61 (0.28-1.33)	0.2136
Tumour grade^c				
1/2	Reference		Reference	
3	1.70 (0.81-3.55)	0.1579	2.46 (1.10-5.52)	0.0288
TNM stage^c				
I/II	Reference		Reference	
III/IV	1.14 (0.47-2.72)	0.7756	0.48 (0.19-1.20)	0.1161
Lymph nodal involvement^c				
No	Reference		Reference	
Yes	0.75 (0.27-2.09)	0.5828	0.69 (0.25-1.91)	0.4761

modelling each IHC marker measurement as the outcome variable and represented by its quartiles.

bModel includes ER, PR, and HER2 status, KI67, tumour size, grade, TNM stage, lymph nodal involvement, and age at diagnosis.

cModel includes tumour subtype, tumour size, grade, TNM stage, lymph nodal involvement, and age at diagnosis.

TABLE 12: HAZARD RATIO (HR) AND 95% CONFIDENCE INTERVAL (CI) FOR THE ASSOCIATIONS BETWEEN IHC MARKERS AND TIL AND OVERALL SURVIVAL (OS) AND RECURRENCE-FREE SURVIVAL (RFS)

		OS (n= 225, number of events= 25)		RFS (n= 225, number of events= 9)	
Characteristic		HR (95% CI)	P value ^a	HR (95% CI)	P value ^a
TIL	Continuous TIL (%)	1.00 (0.98-1.03)	0.9773	1.00 (0.95-1.04)	0.7888
	Categories				
	<10	1.00 (reference)		1.00 (reference)	
	10-30	0.82 (0.29-2.31)	0.7125	0.51 (0.07-3.52)	0.4967
	30+	1.22 (0.44-3.40)	0.6988	1.50 (0.29-7.75)	0.6295
	Trend	1.10 (0.65-1.86)	0.7306	1.20 (0.49-2.96)	0.6893
CD3	Continuous CD3 (%)	1.03 (1.01-1.05)	0.0112	1.02 (0.97-1.06)	0.4900
	Quartiles				
	Q1	1.00 (reference)		1.00 (reference)	
	Q2	0.14 (0.02-1.12)	0.0638	0.34 (0.03-3.91)	0.3876
	Q3	0.63 (0.19-2.09)	0.4540	1.09 (0.12-9.62)	0.9414
	Q4	1.88 (0.71-5.03)	0.2068	1.53 (0.23-10.10)	0.6600
CD4	Trend	1.36 (0.93-1.98)	0.1105	1.28 (0.65-2.53)	0.4762
	Continuous CD4 (%)	1.02 (0.97-1.07)	0.3977	1.00 (0.89-1.11)	0.9299
	Quartiles				
	Q1	1.00 (reference)		1.00 (reference)	
	Q2	1.12 (0.31-4.07)	0.8684	15.38 (0.94-250.00)	0.0551
	Q3	2.83 (0.88-9.09)	0.0813	3.26 (0.18-58.82)	0.4248
CD8	Q4	1.84 (0.52-6.54)	0.3469	5.10 (0.28-90.91)	0.2719
	Trend	1.31 (0.90-1.91)	0.1617	1.19 (0.61-2.31)	0.6155
	Continuous CD8 (%)	1.01 (0.99-1.04)	0.3789	1.02 (0.97-1.06)	0.4895
	Quartiles				
	Q1	1.00 (reference)		1.00 (reference)	
	Q2	0.91 (0.25-3.38)	0.8892	-	
CD20	Q3	0.64 (0.19-2.16)	0.4741	2.28 (0.35-14.93)	0.3866
	Q4	1.94 (0.66-5.68)	0.2276	1.78 (0.24-13.33)	0.5759
	Trend	1.19 (0.82-1.73)	0.3573	1.36 (0.69-2.69)	0.3686
	Continuous CD20 (%)	0.98 (0.91-1.06)	0.6419	1.09 (0.97-1.23)	0.1371
	Quartiles				
	Q1	1.00 (reference)		1.00 (reference)	
CD68	Q2	1.18 (0.38-3.72)	0.7757	0.34 (0.01-15.63)	0.5813
	Q3	0.76 (0.22-2.62)	0.6695	6.02 (0.61-58.82)	0.1249
	Q4	1.14 (0.34-3.82)	0.8290	3.05 (0.23-40.00)	0.3994
	Trend	1.00 (0.67-1.47)	0.9824	1.82 (0.86-3.84)	0.1156
	Continuous CD68 (%)	1.00 (0.97-1.02)	0.7355	0.99 (0.94-1.03)	0.5733

		OS (n= 225, number of events= 25)		RFS (n= 225, number of events= 9)	
Characteristic		HR (95% CI)	P value ^a	HR (95% CI)	P value ^a
Quartiles					
Q1		1.00 (reference)		1.00 (reference)	
Q2		1.07 (0.32-3.58)	0.9116	2.77 (0.40-19.23)	0.3019
Q3		1.56 (0.48-5.05)	0.4573	0.36 (0.02-6.02)	0.4806
Q4		0.67 (0.16-2.79)	0.5790	1.23 (0.13-11.90)	0.8583
<i>Trend</i>		0.95 (0.64-1.42)	0.8066	0.88 (0.44-1.76)	0.7100
Continuous FOX-P3 (%)		0.80 (0.53-1.21)	0.2828	0.96 (0.61-1.51)	0.8518
FOX-P3	Quartiles				
	Q1		1.00 (reference)	1.00 (reference)	
	Q2		6.58 (1.36-31.25)	0.81 (0.05-12.50)	0.8826
	Q3		7.52 (1.48-38.46)	11.36 (1.05-125.00)	0.0451
	Q4		2.13 (0.30-15.15)	2.04 (0.11-37.04)	0.6295
	<i>Trend</i>		1.21 (0.80-1.84)	1.61 (0.74-3.53)	0.2315

3.5.6 Associations between breast cancer risk factors, TILs, and immune cell subtypes

The distributions of TILs and immune cell subtypes by risk factors are presented in **Table 10** and **Table 11**. When adjusted for age at diagnosis and ER status (ER negative used as a reference), there was no significant relationship between mean total TILs and BC risk factors (BMI, age at menarche, menopausal status, parity, age at first pregnancy, cumulative duration of breast feeding and family history of BC). However, when considered individually, patients with onset of menarche at more than 15 years of age and age at first pregnancy between 20-25 years (reference <20 years of age) were more likely to have higher mean CD3 numbers (18.7 and 19.1 compared with 13.3 and 13.2 respectively, $p < 0.05$), while mean CD8 numbers were significantly higher in women with age at first pregnancy greater than 20 years (18.7 versus 12.5, $p < 0.05$) compared with women whose age at first pregnancy was < 20 years. Compared to women with onset of menarche 15 years or younger, those with age at menarche 15 years older were 2.37-fold more likely to belong to the group with the highest CD3. However, an opposite association was observed with age at menarche when comparing the two lowest CD3 categories ORQ2 vs. Q1 = 0.39; 95% CI = 0.17-0.87. However, when adjusting for multiple testing, these results did not attain statistical significance (**Table 10**).

Adjusted results for CD8 on the other hand showed that compared with nulliparous women there was higher odds for women with one child to have CD8 counts distributed in the third quartile (CD8 counts 19.6-31.5, ORQ3 vs Q1 17.30; CI-1.26-236.80). This association did not maintain significance when corrected for multiple tests.

Compared with women who breastfed cumulatively for < 48 months, women who had a cumulative breastfeeding duration of more than 48 months were less likely to have higher mean numbers of CD68 positive stromal cells (37.6 versus 46.1 respectively, $p = 0.01$). The association was only significant for cumulative breast-feeding duration of > 48 months and only for the two highest quartiles Q3 vs Q1 & Q4 vs Q1. This association did not maintain significance when corrected for multiple tests.

The mean levels of CD20 expressing immune cells were more likely to be higher in women between the ages of 40-50 years at diagnosis (CD 20 counts 39.5-55.5, OR Q2 vs Q1 4.05, CI- 1.16-14.12, OR Q3 vs Q1 4.39; CI-1.39-13.82, Q4 vs Q1 3.56; CI- 1.01-12.56) when compared with women less than 40 years of age and those women with 2 children (reference nulliparous women), (CD 20 counts 39.5-55.5, OR Q2 vs Q1 21.04; CI-1.66-267.16, OR Q3 vs Q1 15.16; CI-1.26-182.62, Q4 vs Q1 33.02; CI-2.54-429.82).

Regarding FoxP3, compared with women less than 40 years at diagnosis, women who were diagnosed at more than 40 years of age were progressively more likely to have higher mean number of Fox-P3 cells (mean values of 0.45 at 40-50 years and 0.72 at > 60 years, $p = 0.039$ and 0.0017 respectively). This significance remained after correction for multiple tests, but when adjusted for age at diagnosis, BMI, parity, age at menarche, menopausal status and ER expression, the association was only significant for age at diagnosis of more than 60 years and only for the two highest quartiles Q3 vs Q1 & Q4 vs Q1 (**Table 11**).

3.5.7 Associations between tumour characteristics, molecular subtypes and TILs and immune cell subtypes

Results of adjusted associations between tumour characteristics, total TILs and immune cells are shown in **Tables 10** and **11**.

The distribution of TILs and immune cells are presented in **Figure 10**

Compared to ER positive tumours and luminal A breast cancer, a significantly higher mean total TILs were observed with ER and PR negative tumours, TNBC, Grade 3 tumours, high Ki67 index and in the absence of lymph node involvement (**Table 10**). Also, the observation of a high mean TIL count in TNBC (22.7 compared to 13.4 in ER positive tumours, $p=0.003$) and tumours with a high Ki67 ($>20\%$ Ki67 index) were still observed when adjusted for multiple variables. When adjusted for covariates and compared with luminal A Tumours, there was a likely progressive increase in total TILs from luminal B to HER2 enriched to TNBC, however only the TNBC subtype retained significance (OR 2.97; CI-1.10-8.01, $p=0.03$).

When compared to the luminal A subtype, the odds of a high total TIL (>30 vs $<10\%$) was more likely with a high Ki67 index (OR 2.97; CI- 1.30-6.82, $p=0.01$) **Table 11**.

Compared to tumours less than 2 cm in size, tumours greater than 2 cm were more likely to have TIL counts $< 10\%$ rather than TIL counts of 10-30% (OR 0.29; CI- 0.14-0.61, $p=0.001$).

With regards to individual immune cell types, compared to cases without lymph node metastases, the CD4 mean numbers were lower (3.35 versus 7.06) in the presence of lymph node involvement ($p=0.0017$) even after correction for multiple tests. However, the association of CD4 being less likely to have lymph node involvement was only significant in Quartile 4 (CD 4 counts $>11\%$, OR Q4 vs Q1 0.18; CI- 0.06-0.62) when adjusted for covariates.

Lower counts in the presence of lymph node metastases were also observed with CD20 (2.49 versus 4.59, $p=0.028$) but this association was not observed when adjusted for covariates (**Table 11**). When compared to luminal A, and after adjusting for covariates, luminal B tumours were more likely to have a higher CD20 count but this significance was only observed for CD20 counts in Q3 and Q4 (CD20 counts 2.5-7.6, ORQ3 vs Q1 3.81: CI-1.17-12.41, ORQ4 vs Q1 3.99:CI-1.17-13.61).

When adjusted for covariates, CD8 counts were less likely to be associated with PR and HER2 positive tumours, but this observation was limited to Q2 only compared with Q1 (CD 8 counts 8.0 - 19.6, OR Q2 vs Q1 0.22; CI - 0.05-0.97 and OR Q2 vs Q1 0.30; CI- 0.11-0.85).

The mean scores for CD68 were significantly higher in Grade 3 tumours compared with Grade 1 tumours, higher in Ki67 scores $>20\%$ and in Stage II tumours compared with Stage I tumours (**Table 10**). However, after adjusting for covariates, the only association that remained significant was that with Ki67 index and this was only observed in Q3 of the CD68 count (CD68 count 39.47 - 55.5, OR 4.01; CI-1.41-11.42)

The mean score for FoxP3 was seen to increase with increasing tumour grade, however this association was not observed when adjusted for covariates.

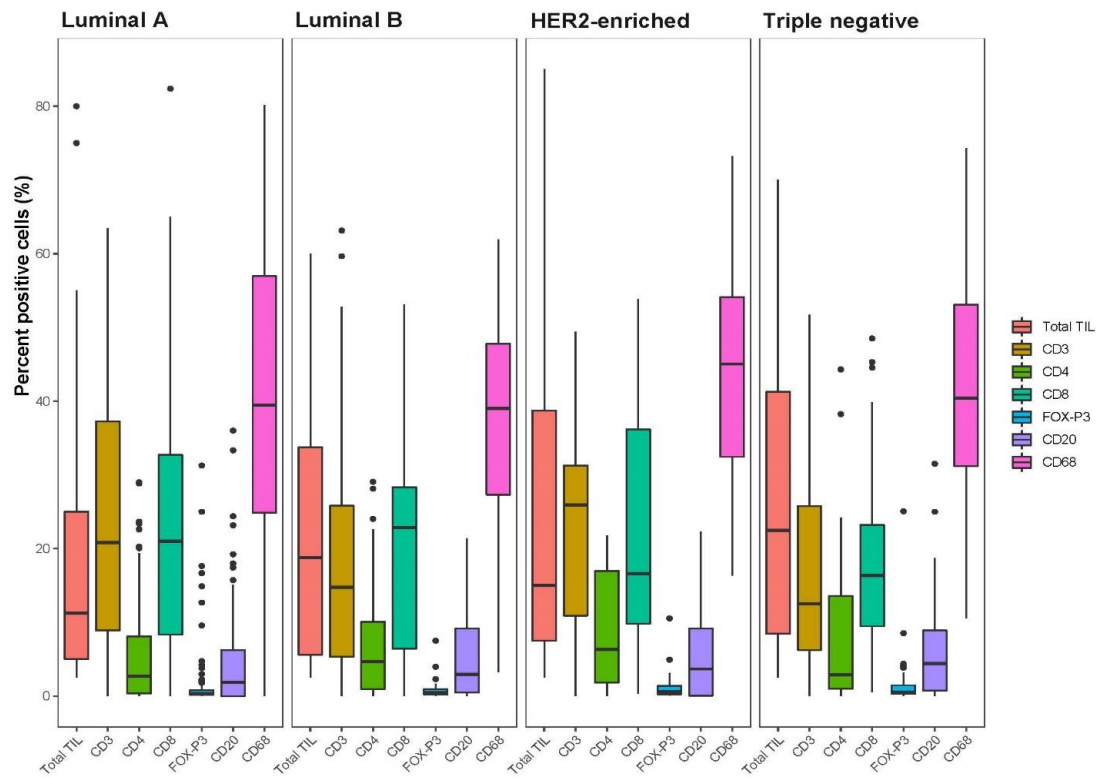


FIGURE 13: DISTRIBUTION OF TILs AND TIL SUBTYPES BY BREAST CANCER MOLECULAR SUBTYPES

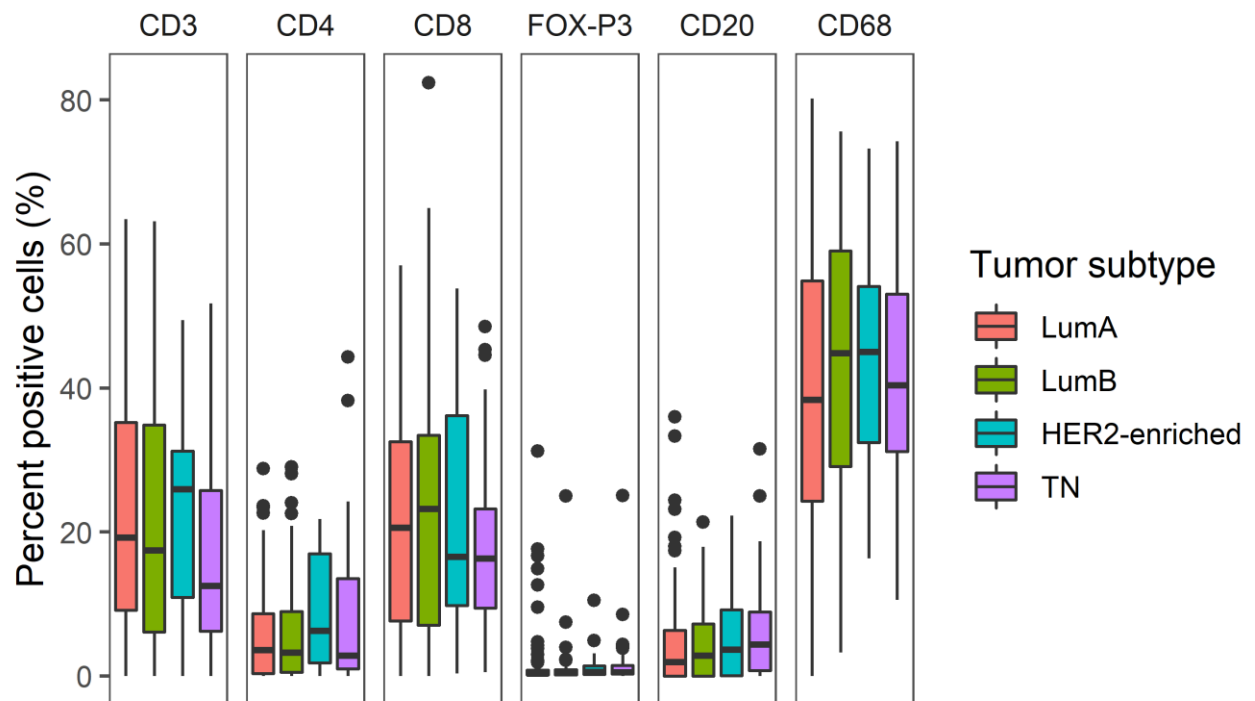


FIGURE 14: IMMUNE CELL SUBTYPES BY BREAST CANCER MOLECULAR SUBTYPES

3.5.8 Associations between molecular subtypes and TILs and outcome data

The Overall Survival (OS) and Recurrence Free Survival (RFS) data are presented in **Table 10** and illustrated with the Kaplan Meier curves (**Figures 14 to 25**).

There were 25 deaths and 9 recurrences out of a total of 225 patients for whom this data was available.

There was no significant association noted between total TIL, overall survival, and recurrence free survival.

Taken individually, there was no association noted between OS, RFS and C4, CD8, CD20 and CD68 as continuous variables or quartiles.

When analysed as a continuous variable CD3 showed a slight but significantly increased hazard ratio with OS (HR1.03; CI- 1.01-1.05, P= 0.01), but this association was not observed when CD3 was analysed by quartiles (**Table 11**).

The hazard ratio for OS was increased for Fox P3 counts in both Quartiles 2 and 3 when compared with Quartile 1 (FoxP3 count 0.169-0.985, HR Q2 vs Q1 6.58; CI-1.36-31.25, HR Q3 vs Q1 7.52; CI-1.48-38.46).

For RFS, the hazard ratio was increased for Fox P3 counts in Quartile 3 when compared with Quartile 1 (FoxP3 count 0.460.985, HR Q3 vs Q1 11.36; CI- 1.05-125.00)

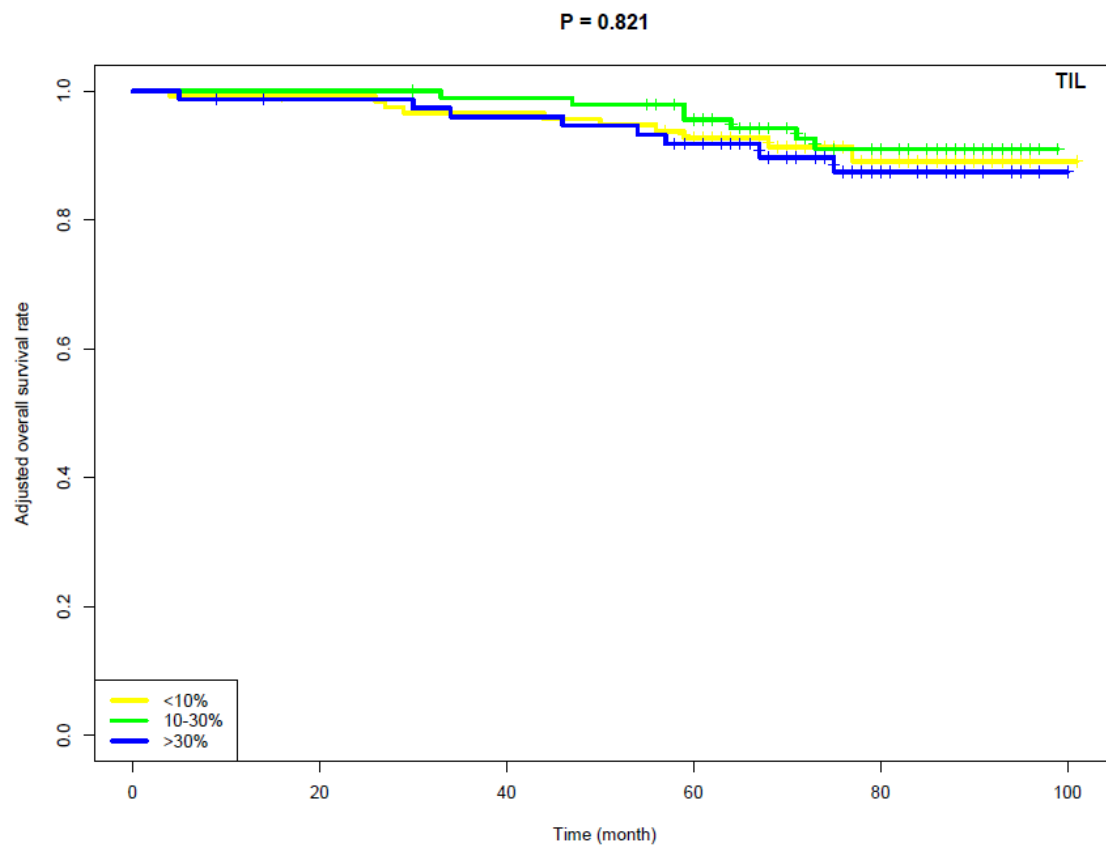


FIGURE 15: OVERALL SURVIVAL AND TOTAL TILS

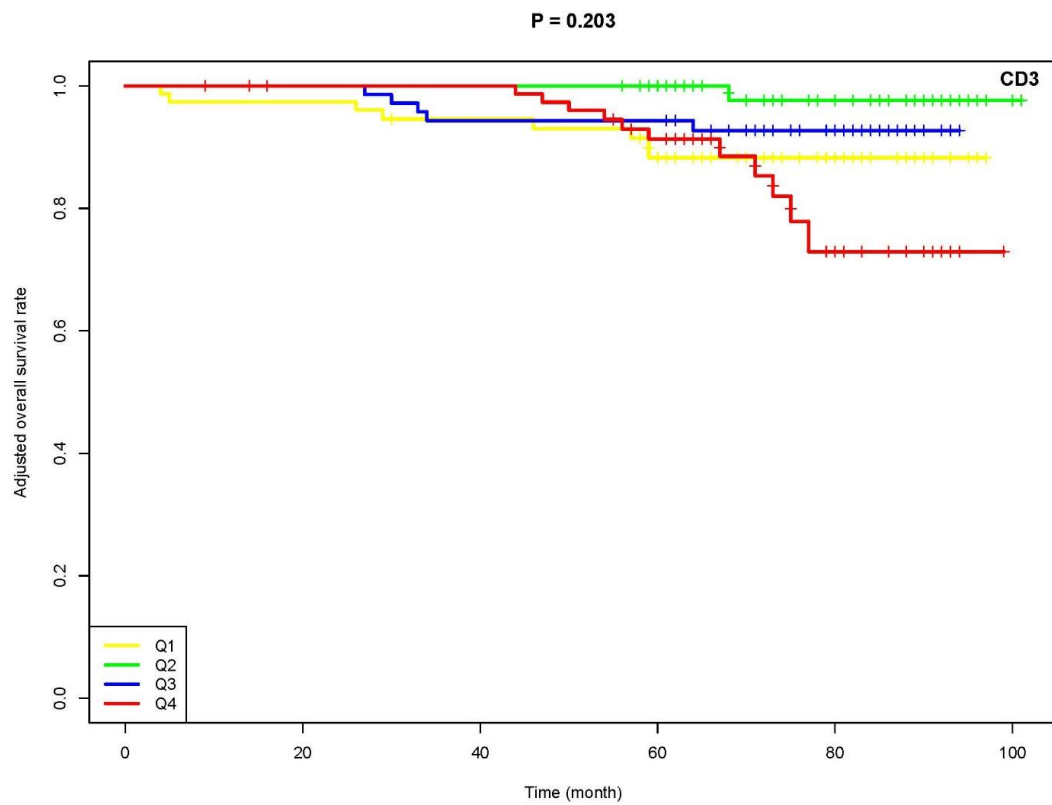


FIGURE 16: OVERALL SURVIVAL AND CD3

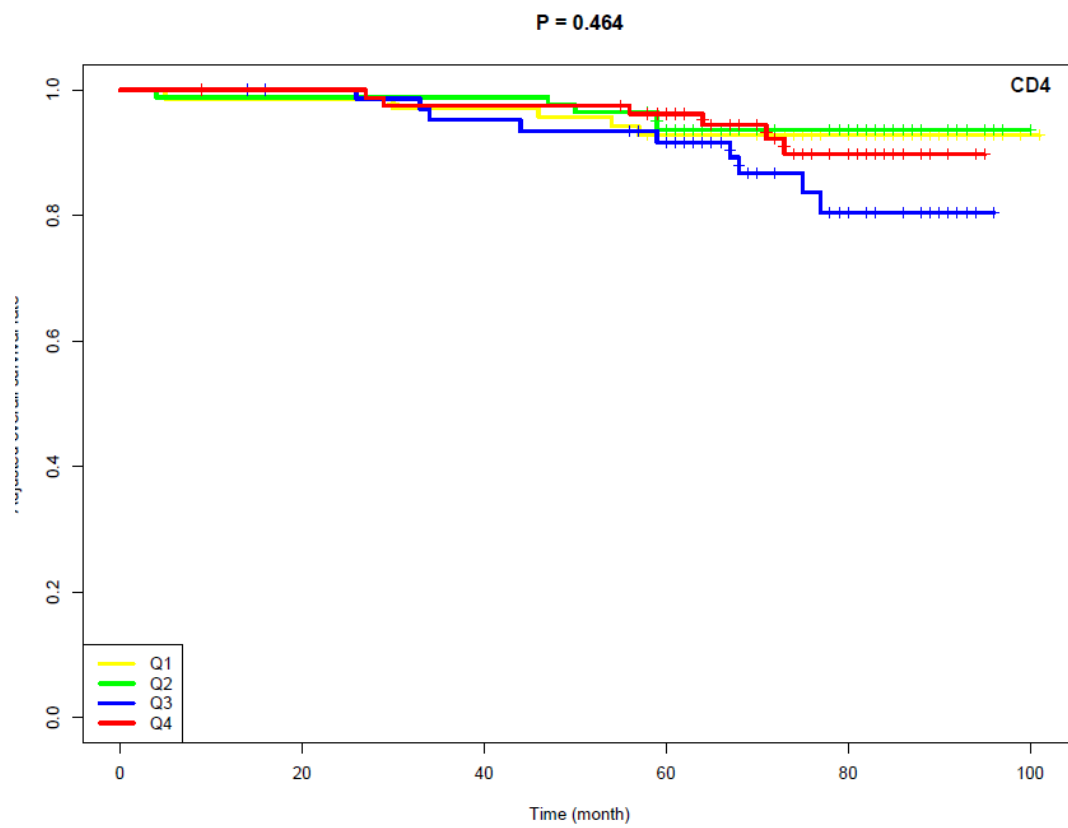


FIGURE 17: OVERALL SURVIVAL AND CD4

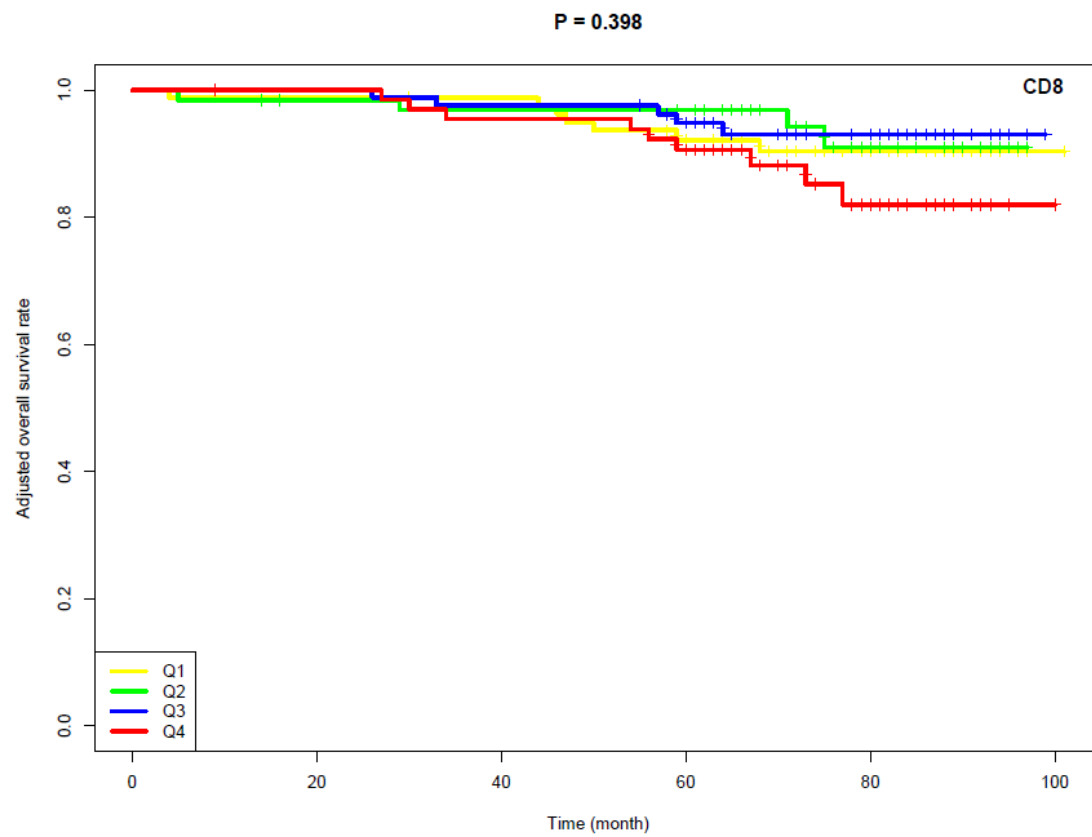


FIGURE 18: OVERALL SURVIVAL AND CD8

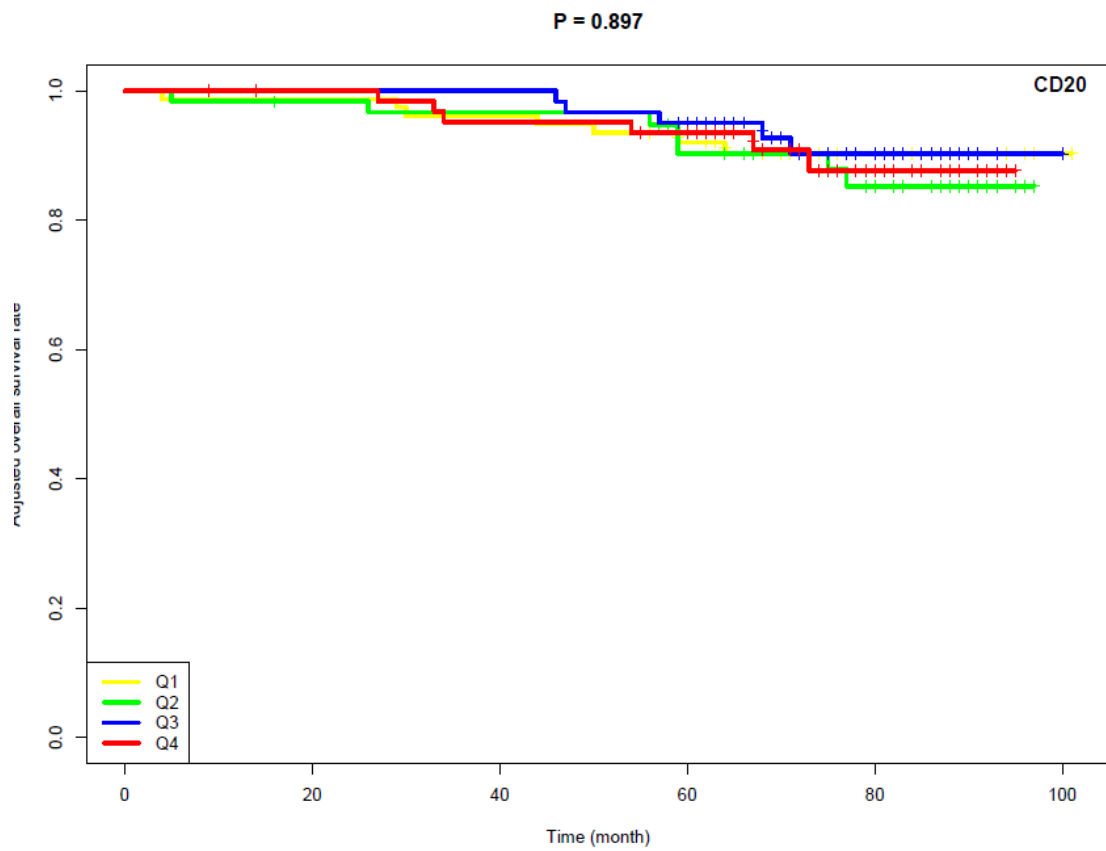


FIGURE 19: OVERALL SURVIVAL AND CD20

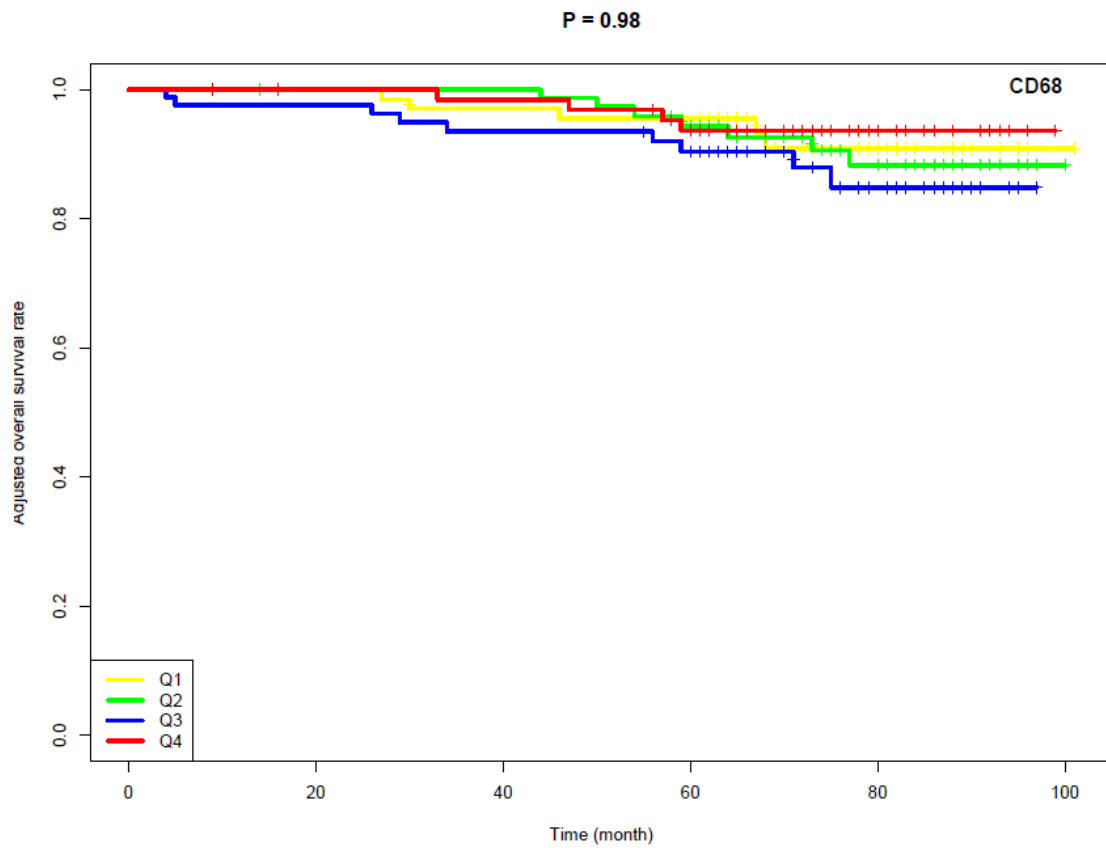


FIGURE 20: OVERALL SURVIVAL AND CD68

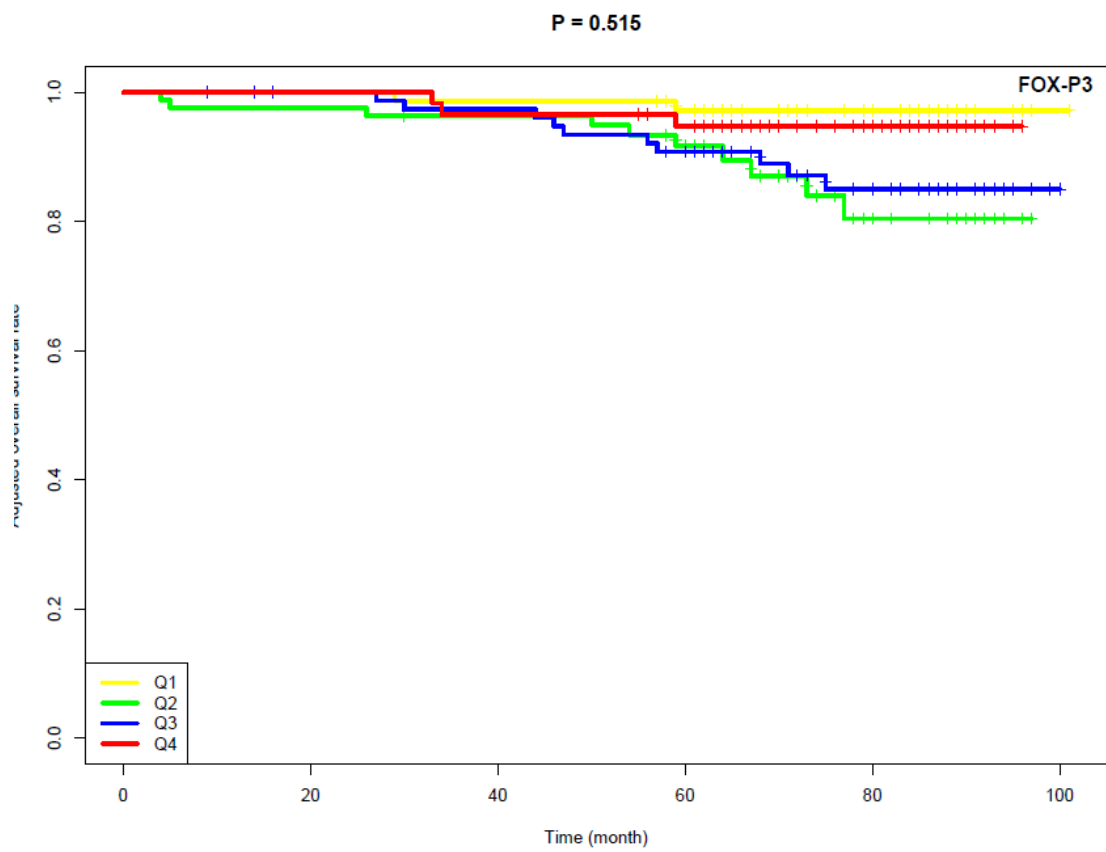


FIGURE 21: OVERALL SURVIVAL AND FOX P3

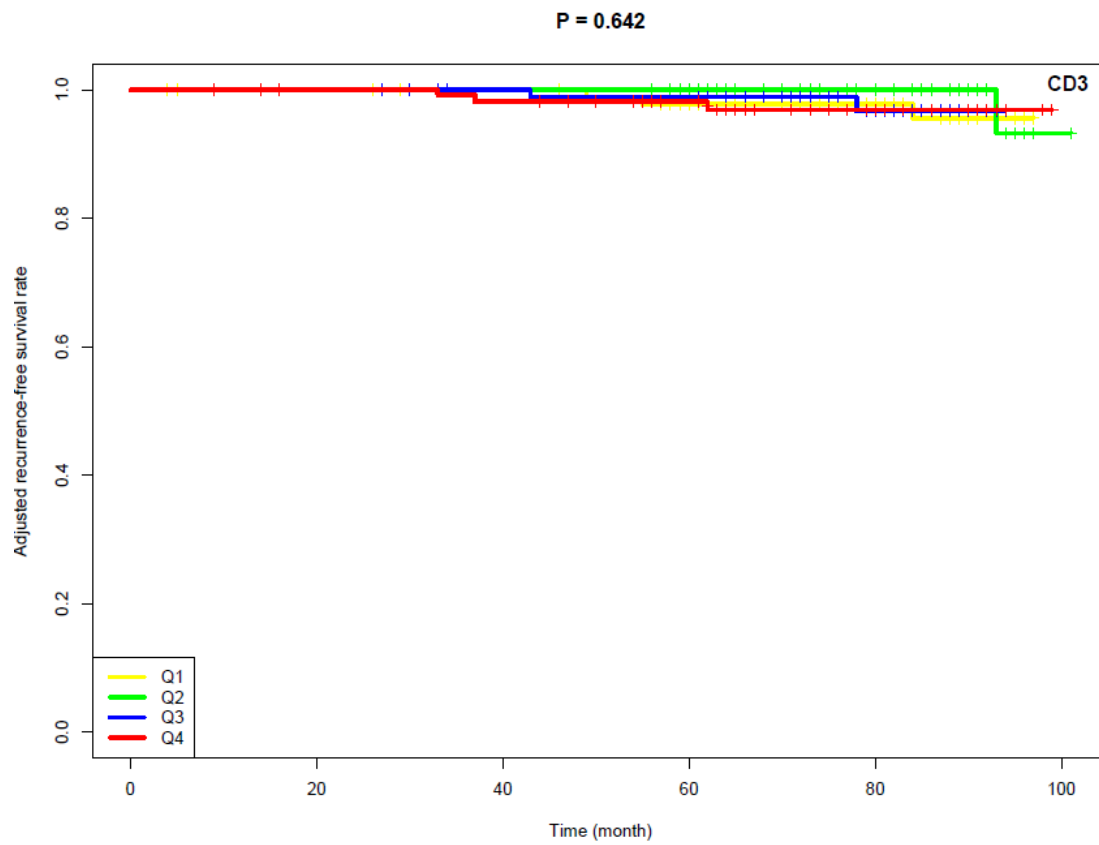


FIGURE 22 : RECURRENCE FREE-SURVIVAL RATE AND CD3

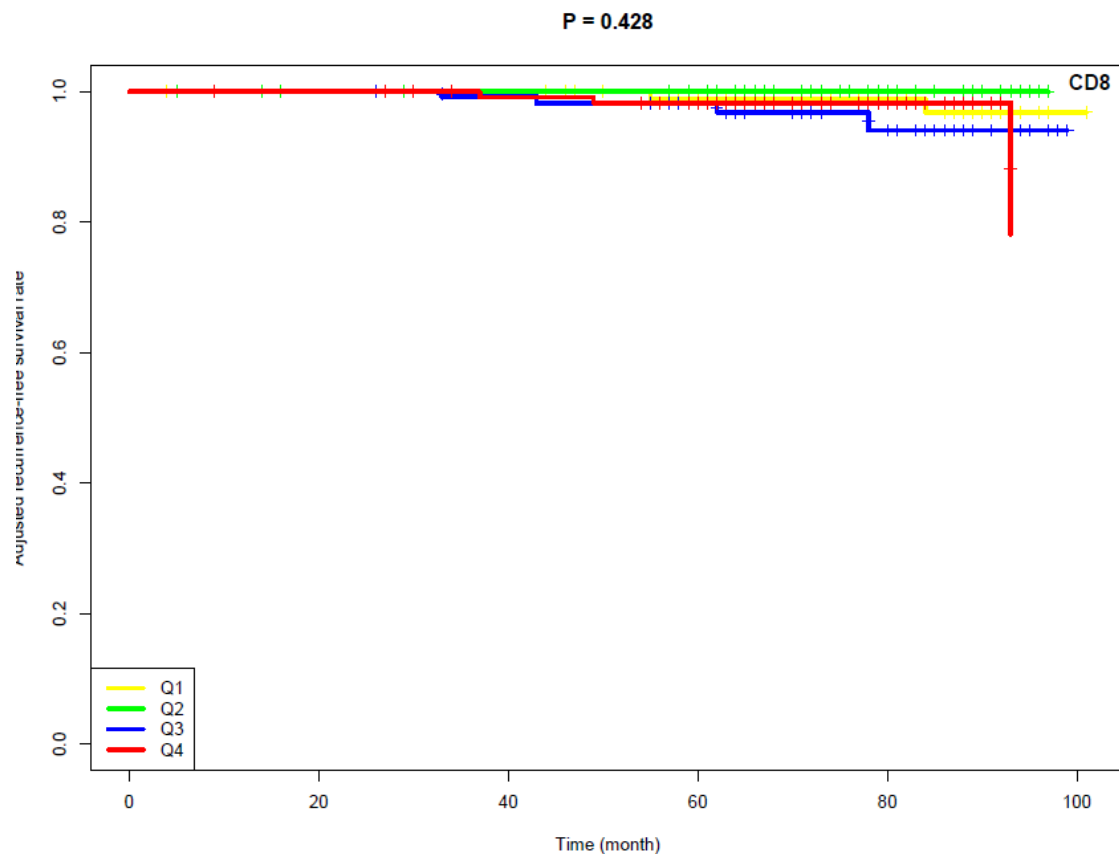


FIGURE 23: RECURRENCE FREE-SURVIVAL RATE AND CD4

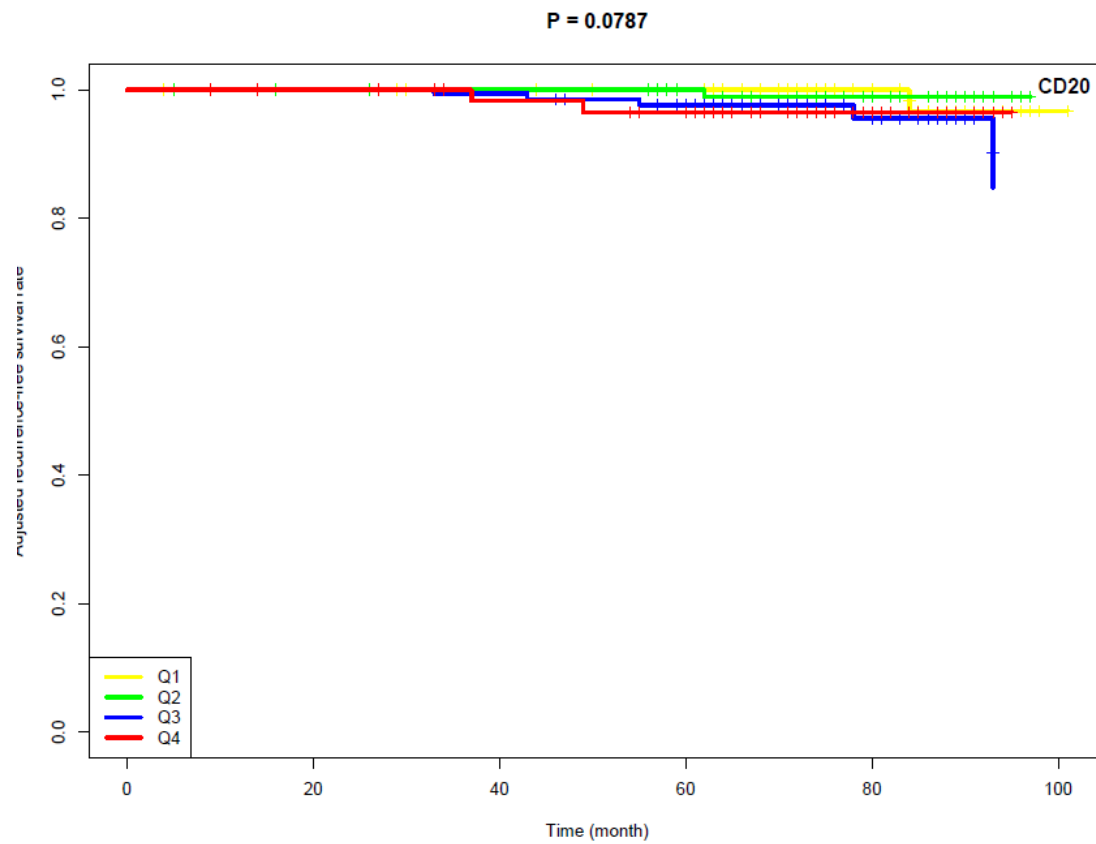


FIGURE 24: RECURRENCE FREE-SURVIVAL RATE AND CD8

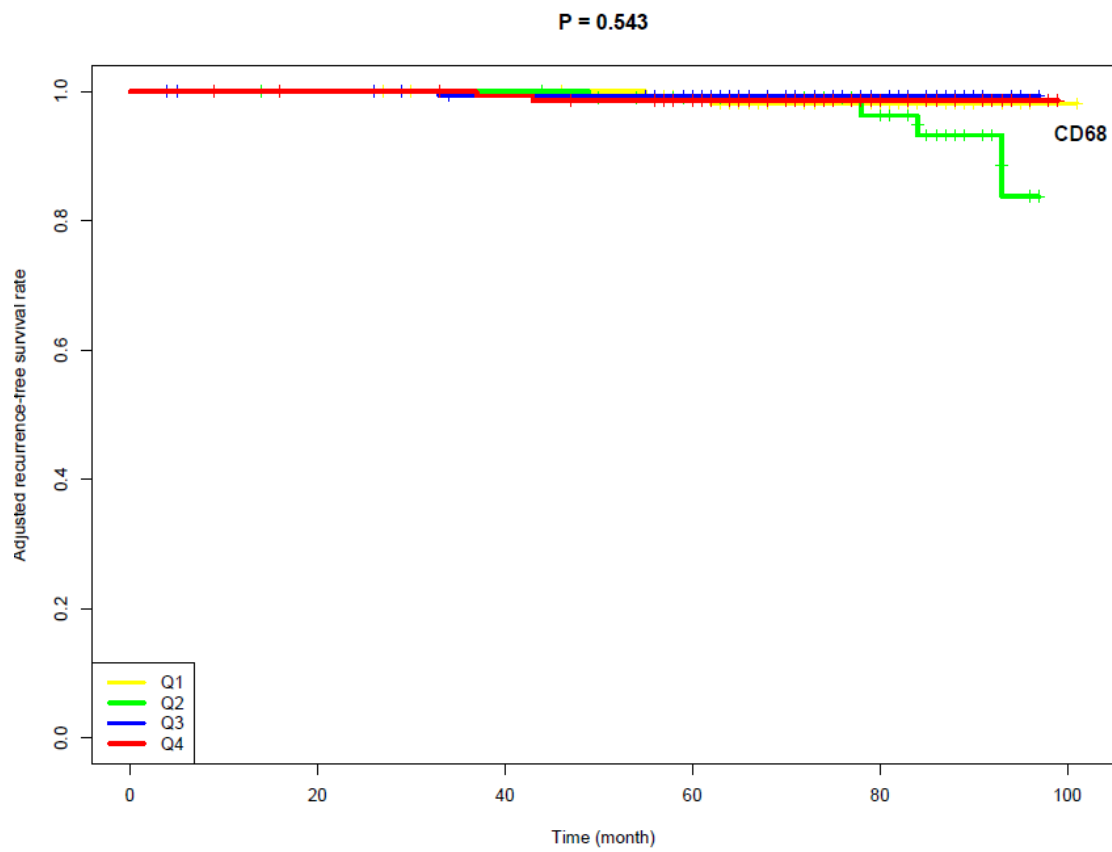


FIGURE 25: RECURRENCE FREE-SURVIVAL RATE AND CD68

3.6 DISCUSSION

There are limited data from Sub Saharan Africa on the association of the immune response in breast cancer subtypes with BC outcomes. Pathways that are enriched for signals of adaptation tend to be involved in immune response and wound healing(168), suggesting that African breast cancer may have distinct immune profiles. Due to the consistent evidence of the prognostic value of TILs in TNBC and HER2 positive breast cancers(218, 237-241), the 17th St Gallen International Breast Cancer consensus guidelines recommend routine pathologic evaluation of stromal TILs in early TNBC(242).

In this case series of Kenyan patients with breast cancer, we show that although some associations with stromal TILs were observed, the most consistent were in relation to tumour grade and size. We found that TIL counts by both the visual and automated count increased with BMI, tumour grade and Ki67 index while there was an inverse relationship with tumour size. We observed a significantly higher mean TIL count in ER and PR negative tumours, and TNBCs. Our study findings are similar to those of Mremi et al(122), which examined specific prognostic and predictive features of BC that included TILs in a population of 90 Tanzanian, 120 African American and 120 White American women, and observed that in the former group the TIL counts were typically higher in TNBC which conferred the tumours in these women a better prognosis. These findings similar to ours suggests that other yet unexplored exposures could explain the heterogeneity in anti-tumour immune response. Using similar methodology to our study, previous studies from Kenya have reported increased CD4 lymphocytes and tumour associated macrophages in 68 breast cancer cases when compared with benign breast tissue samples (225) and a subsequent study of 49 breast cancer tissues (226) showed correlation of TIL with Ki67 but not tumour grade nor subtype. This is likely due to the small sample size.

We did not observe any association of TILs with OS or RFS in BC subtypes which may reflect the limited events reported (225 cases with 34 events). A recent meta-analysis involving 15,676 patients in whom outcome data was available for 6697 patients demonstrated that each 10% increment of TILs in breast tumours improved overall survival in HER2 enriched and TNBC subtypes but not in luminal BC subtypes(243). Our study did not find any significant associations between immune cell subtypes with OS and RFS, except for Fox P3 and CD3.

When we examined TILs using the visual method and compared that with the Halo platform, there was a weak but significant correlation between visual and Halo TILs. We realised that the two methods cannot directly be compared because machine learning algorithms generate stromal cellular density (SCD) which is the percentage of tumour-stroma that is occupied by nucleated cells and includes fibroblasts, inflammatory, endothelial, and other cells in the tumour(235). The clinical significance of SCD in the tumour microenvironment will require further studies.

We used immunohistochemistry to characterise the TIL subtypes and utilized automated image analysis for enumerating the TILs in the breast cancer microenvironment of this well characterised cohort. We analysed the association of these TILs with clinical and pathologic factors, overall survival (OS) and recurrence free survival (RFS).

Significantly, we observed that some demographic and reproductive factors specifically older ages at diagnosis were associated with increased mean FoxP3 counts while older ages at menarche and first pregnancy were likely to be associated with higher mean scores for CD3

while CD68 counts were reduced with prolonged cumulative breast-feeding duration. These findings remained significant after adjusting for covariates but lost significance after multiple testing probably a result of the small sample size. FoxP3 is a marker of T- regulatory cells (Tregs), which are known to increase with age (244) and cause T cell dysregulation leading to increased susceptibility to disease including cancer (244). Our findings of increasing FoxP3 with age older than 60 years were maintained after adjusting for covariates signifying this may not be an age-related effect but rather a reflection of the adaptation of the immune system to unique infectious exposures through natural selection. We observed FoxP3 levels to be significantly associated with worse overall and recurrence free survival which is consistent with other studies showing FoxP3 to be an independent prognostic factor of overall survival in breast cancer(245).

Anti-tumour immunity is mediated through a complex interaction of immune cells within the tumour microenvironment. Cytotoxic CD8 positive T cells play a major role in effecting anti-tumour cell mediated immune response and have been widely studied in breast cancer as a biomarker of anti-tumour activity (246-248) .

We found that the average CD8 count was not correlated with tumour subtype however when divided into quartiles, CD8 was inversely correlated with PR and HER2 expression but only for average CD8 counts of 8-19.6 compared to those less than 8. We did not observe any correlation between CD8 counts and overall or recurrence free survival. Similarly, Mahmoud et al(173), who used similar methodology to ours (stained TMA sections), in study involving 1334 breast cancer patients showed that the total number of CD8 lymphocytes correlated with negative ER and PR expression ($p < 0.001$), but not with HER2 status. Interestingly, the study found a survival advantage with higher CD8 in triple negative and HER2 positive BC. A study using stained TMA sections, and enriched for African Americans (550 African Americans, 138 Caucasian), observed racial differences in CD8 density with black patients having a higher median CD8 density, across all subtypes compared to white women (87.6/mm² and 53.1/mm² respectively, $p < 0.001$), however this did not translate into a survival advantage. Similar to our results, Meng et al (249) reported that the percentage of CD8+T cells was significantly reduced in HER2+ and TNBC patients compared with patients with the luminal subtype ($P = 0.019$ and $P = 0.003$, respectively). Some of the differences we observed could be explained by the interaction and influence of other immune cells on the functions of CD8 which would need to be further studied.

Higher CD3 score when evaluated on a continuous form was correlated to worse overall survival in our study. These findings are contrary to those of 1165 invasive BCs from Nottingham City Hospital conducted over a 10-year period, which demonstrated that CD3+ cells constituted the main subtype of all TILs and that CD3+ and CD8 + cells were associated with a longer breast cancer specific survival in TNBC(185). The inconsistent findings of the association of CD3 with breast cancer survival among different populations highlight the complexity of immune cell interactions with the tumour microenvironment and the importance of conducting large immunological studies in diverse African populations.

We also observed that CD68+ tumour associated macrophages (TAMs) were more likely to be associated with higher tumour grade and a high proliferation index but not to overall survival or recurrence free survival. Hypoxic conditions within tumour microenvironment also promotes M2 differentiation, the main subtype of macrophages within the tumour microenvironment which leads to an immunosuppressive environment through production of IL10, Arginase 1 and Vascular Endothelial Growth Factor (VEGF), all of which suppress Th1 T cell function(196). A recent meta-analysis of TAMs and primary breast cancer over a

six-year period that included 1576 patients showed that both CD68 and CD163 were significantly associated with lymph node metastasis (CD68, $P = 0.003$; CD163, $P < 0.001$); high Ki67 (CD68, $P = 0.026$; CD163, $P < 0.001$), poor histological grade (CD68, $P < 0.001$; CD163, $P < 0.001$) and hormonal receptor negativity (CD68, $P < 0.001$; CD163, $P < 0.001$), however only CD163 positive TAMs showed significantly worse survival outcome(250). Larger studies to further confirm the role of TAMs in breast cancer subtypes and outcomes is recommended.

The strength of our study is the well annotated study sample with availability of outcome data. We also used digital scanned TMA slides to annotate the TIL subtypes within the tumoural stroma manually with subsequent automated analysis thus reducing the chance of inter-observer variations.

One major limitation of our study is that our sample size is small, so it is hard to make inferences on the associations that were observed or those that were not. The other limitation was the use of TMAs instead of whole section slides thus precluding assessment of TIL subtype heterogeneity across the cores, however the 1 mm size of each core for a total of 4 cores per case is more than adequate for accurate assessment of TILs and of TIL subtypes(185).

In our study, the TILs were visually assessed on whole H/E sections as per the International TIL Working group guidelines. TMAs were only used for assessment of the distribution of the various immune cells. For the latter, duplicate punches each of 1 mm each were used for the leading edge and intratumoural stroma.

A 1 mm size of each core for a total of 4 cores per case is representative of the whole tissue for accurate assessment of TILs and of TIL subtypes(185)

Furthermore, there have been several studies that have analyzed the prognostic impact of immune-variables in large cohorts and on TMA's(251).

In conclusion, the immune landscape of breast cancer in patients from Sub Saharan Africa is understudied and findings are mainly extrapolated from studies in Caucasians and African Americans. We have previously documented risk factor heterogeneity in Kenyan women with breast cancer that is different from that in African Americans (unpublished data). Although our findings with regards to enrichment of TILs in ER negative and TN breast cancers are similar to what has been previously published, the unexpected association of CD3 with worse survival outcomes points to yet unexplained immune cell modifications that require further study.

Our findings add to the body of knowledge regarding TILs and TIL subtypes in Kenyan women with breast cancer and underscores the need to conduct studies specific to African populations comparing various methods and measures of TILs and TIL subsets to define the contribution of the immune system to tumour progression in this population and ultimately offer personalised treatment, monitor treatment responses and outcomes in breast cancer.

Chapter 4

4.0 CYTOKINES

4.1 INTRODUCTION

Cytokines are a large group of proteins that are secreted by specific cells of the immune system (particularly macrophages and activated T lymphocytes). Usually peptides or glycoproteins of low molecular weight they act on a variety of target cells in an additive, synergistic, or antagonistic manner (252). Cytokines are divided into several categories: Interleukins, Chemokines and Lymphokines.

Cytokines are signalling molecules produced by cells for specific biological functions. Interleukin (IL) is a type of cytokine produced by white cells responsible for proliferation, maturation, and migration of immune cells. They are produced as “chemo-attractant molecules” i.e., specific cytokines that function by attracting immune cells to sites of infection/inflammation. Chemokines are polypeptides that form the largest family of cytokines with approximately 41 human members forming a complex network that is chemotactic for all leucocytes with individual functions depending on type and location. Based on their name, lymphokines originate from lymphocytes, and are further subdivided into interferons, Tumour Necrosis Factor(TNF), and Tumour Growth Factor (TGF) (253). Cytokines can be categorized as those that effect the cells of the adaptive immune response, or as those that have pro- or anti-inflammatory functions. They can be further classified based on the receptors used for signalling (254). In neoplastic processes, the role of chemokines became apparent in experimental animals’ null for natural killer (NK) cell function, which when challenged with a tumour, showed a typical inflammatory infiltrate.; These experimental models suggest that neoplastic cells either produce chemotactic factors or induce their expression in nearby ‘host’ cells. Moreover, chemokines induce direct effects on stromal and neoplastic cells in addition to their roles in regulating leukocyte recruitment (12)

The main communication between host cells and cancer cells is a key function of cytokines and these are associated with shaping the tumour microenvironment and metastasis (255).Cytokines are highly active proteins, and hence very tightly regulated. The concentration in blood in healthy individuals is usually undetectable. A higher concentration usually indicates inflammation or is associated with disease progression making cytokines good biomarkers to monitor disease and treatment progression (256).Plasma cytokine levels have been associated with relapse and metastases of breast cancer(257) . Cytokines can regulate tumour progression, and also alter the tumour micro- environment inducing tumour promotion and affecting the sensitivity to chemotherapeutic agents (258).Some of the functions of cytokines include immune mediation and modulation, development of hematopoietic stem cells and communication between cells. They also affect how the host responds to different stimuli and various infectious pathogens(256).

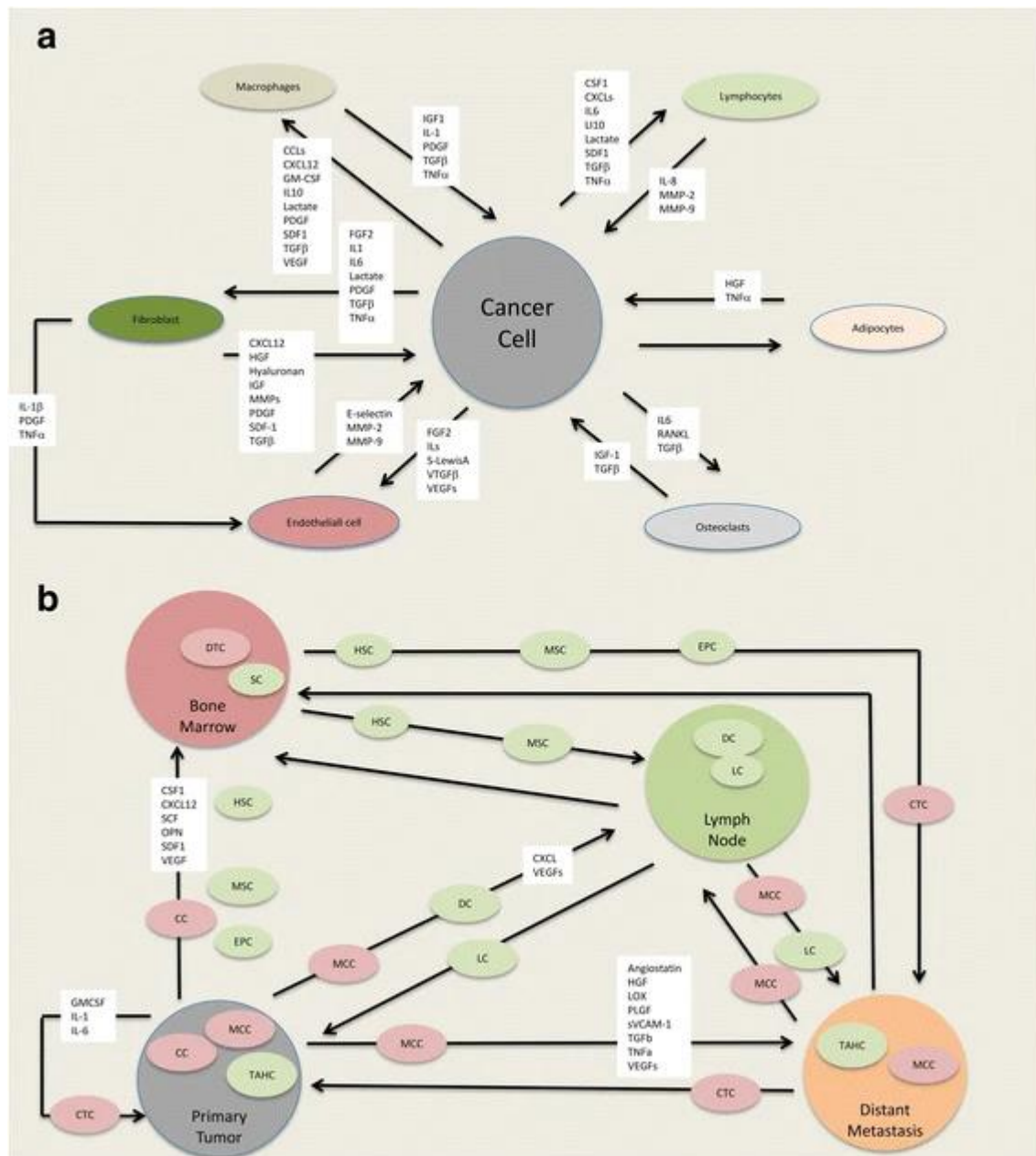


FIGURE 26: SCHEMATICS OF MUTUAL COMMUNICATION BETWEEN CANCER CELLS, TUMOR ASSOCIATED HOST CELLS AND ECOSYSTEMS(259).

©by, Nora Sundahl et al, Molecular Medicine.2018

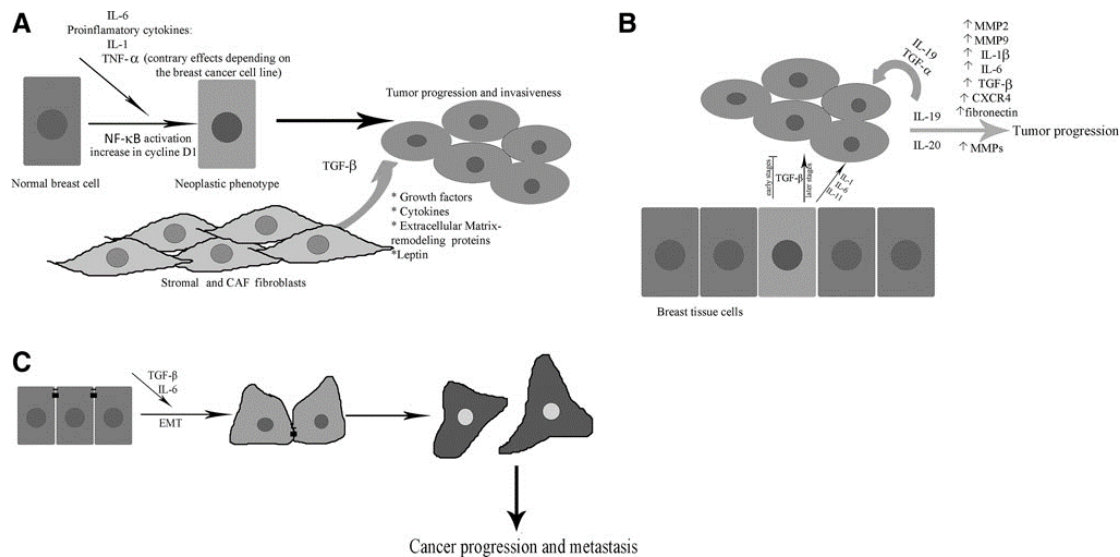


FIGURE 27: ROLE OF CYTOKINES IN THE DIFFERENT STAGES OF BREAST CANCER (260)

© *Journal of Interferon & Cytokine Research*. 2015

Studies have shown that other than the immune cells in the tumour micro-environment producing cytokines, tumour cells also produce cytokines. The most commonly studied cytokines that play a part in tumour progression are IL-1, IL-4, IL-6 while IL-2 has been known to boost the host's immunity against the tumour(261). Poor prognosis of most breast cancers is thought to be associated with over expression of cytokines. While IL-1, IL-6, IL-11, and TGF-beta drive breast cancer growth and/or invasion, IL-12, IL-18, and IFNs suppress these processes. Serum levels of some cytokines appear to be favorable prognostic indicators (soluble IL-2R), whereas others (IL-1beta, IL-6, IL-8, IL-10, IL-18, gp130) appear to be unfavorable(252). Nuclear Factor Kappa Beta (NF- κ B), an inducible signalling transcription factor, contributes in cytokine receptor activation and intracellular signalling leading to acceleration of tumour progression(262).

TABLE 13: CYTOKINE EFFECTS ON THE BREAST TUMOR MICROENVIRONMENT

CYTOKINE	SOURCE	CYTOKINE EFFECT ON TUMOR	CLINICAL AGENTS IN DEVELOPMENT FOR USE IN BREAST CANCER
IL-1	B-cells Mature DCs NK cells Macrophages	Pro-Tumor • Promotes tumor invasion • Promotes inflammation	IL-1 Receptor Antagonist • Anakinra (NCT01802970) • IL-1 Neutralizing mAB • Canakinumab (NCT03742349)
IL-2	Activated T-cells NK cells DCs Macrophages	Antitumor • Activates NK cells • Activates CD8+ T-cells	IL-2 Analogues • Aldesleukin (NCT00006228) • Bempegaldesleukin (NCT02983045 ; NCT03435640)
IL-6	B-cells TAMs Breast tumor	Pro-Tumor • Treatment resistance • Increased tumor plasticity • Promotes metastasis	IL-6 Antagonist • Tocilizumab (NCT03135171)
IL-8	TAMs Breast tumor	Pro-Tumor • Increased tumor plasticity • Increased MDSCs	IL-8 Neutralizing mAB • Humax IL-8 (NCT02536469) • IL-8 Receptor Antagonist • SX-682 (protocol in development) Reparaixin (NCT02370238)
IL-10	TAMs B-cells	Antitumor (pleiotropic) • Pro-Tumor Promote Tregs • Promote M2 phenotype	IL-10 Antagonist • Pegilodecakin (NCT02009449)
IL-21	CD4 ⁺ T-cells Breast tumor	Antitumor • Stimulates NK cells • Stimulates CTLs	IL-21 Analogues • Recombinant IL-21 (no further clinical development at this time)
IFN	DCs NK cells Macrophages B-cells T-cells	Antitumor • Enhanced antigen presentation • Increased CTL-killing	IFN Analogues • Human-leukocyte IFN (NCT03328026) IFN Stimulating Agents • MIW815 (NCT02675439)

	Breast tumor		
TGF-β	TAM	Antitumor (Early-Stage Disease;	TGF- β Receptor 1 Inhibitor
	Breast tumor	pleiotropic)	• Galunisertib
	Macrophages	Pro-Tumor (Late-Stage Disease)	(NCT02538417; NCT02672475)
		• Promotes M2 • Therapeutic resistance • Immune suppression	TGF- β Receptor 2 Inhibitor • M7824 (NCT03579472 ; NCT-3524170; NCT03620201)

Analogues increase the activity (or concentration) of a specific cytokine within the TME
 Antagonists and neutralizing antibodies decrease the activity (or concentration) of active cytokine within the TME

CTL cytotoxic T lymphocyte, DC dendritic cell, IFN interferon, IL interleukin, mAB monoclonal antibody, MDSC myeloid derived suppressor cells, NK natural killer, TAM tumor associated macrophages, TGF- β transforming growth factor beta

© *npj Breast Cancer* 5, 37 (2019)

4.2 TYPES OF CYTOKINES

TABLE 14: CYTOKINES' UTILITY FOR BREAST CANCER PROGNOSIS AND SURVIVAL

<i>Cytokine</i>	<i>Location</i>	<i>Levels</i>	<i>Prognosis</i>	<i>Impact on survival</i>	<i>References</i>
TGF- β	Tumor	↑	Is predictive of local and distant metastases. Loss of Wnt5a, regulated by TGF-b, is associated with early relapse and metastases	Worse survival	Grau and others (2008); Ivanovic' and others (2009); Yu and others (2010); Zu and others (2012)
TGF- β 2	Serum, tumor	↑	Favorable: In circulation, higher levels correlate with later relapse. TGF-b2 mRNA expression correlates with early relapse in advanced stage than in early stage of breast cancer Unfavorable: Higher transcript levels in tissue is a marker of unfavorable prognosis	Longer relapse-free survival mRNA levels correlate with increased survival in early breast cancer than in advanced stage	Dave and others (2011)
TGF- β 3	Tumor	↑↓	Reduced expression correlates with poor prognosis and higher tumor grade, while increased expression is indicative of good prognosis	Not determined	Desmedt and others, (2007); Ivshina and others (2006); Miller and others (2005); Sotiriou and others (2006); van de Vijver and others (2002)
TGF- β 3	Serum	↑	Higher TGF-b3- TGF-b3 receptor complexes are linked to lymph node metastases: poor prognosis Nonetheless, upregulation of this cytokine is involved in the parity-induced protection against breast cancer in women who gave birth before the age of 24 years	Not determined	D'Cruz and others (2002); Laverty and others (2009)
IL-1 and IL-1B	Serum, tumor	↑	High in tumor and circulation: poor prognosis, more aggressive and invasive tumor. High IL-1ra and low IL-1 at tumor site: good prognosis	Not determined	Chavey and others (2007); Goldberg and Schwertfeger (2010); Jin and others (1997); Rafiq and others (2007)
IL-6	Serum	↑	Poor prognosis: diminished response to endocrine therapy and to chemotherapy Lower levels correlate with better response to therapy: good prognosis	Worse survival	Baumgarten and Frasar (2012); Culig, (2011); Knu" pfer and Preiss (2007); Liu and others (2012); Zhang and Adachi (1999)
IL-19	Tumor	↑	Associated with higher mitotic rate, advanced tumor stage, and metastasis: poor prognosis	Worse metastasis-free survival	Hsing and others (2012)
IL-8	Tumor	↓	Low levels are associated with lymph node-positive status, lack of hormone receptors, accelerated clinical progression, and metastasis: poor prognosis.	Increased risk of death	Culig (2011); Zuccari and others (2012)
TNF and IL-13	Serum	↑	Higher levels correlate with higher tumor stage and lymph node metastasis	Not determined	Sheen-Chen and others (1997); Srabovic and others (2011)

CYTOKINES IN BREAST CANCER

IL-10	Serum	↑	May inhibit tumor growth by suppressing IL-6. However, higher levels of IL-10 correlate with metastatic disease: poor prognosis	Not determined	Kozlowski and others (2003); Merendino and others (1996)
IL-20	Tumor	↑	Associated with advanced tumor stage, metastasis, and poor clinical outcome: poor prognosis	Worse survival	Hsu and others (2012)
CSF-1	Tumor	↑	Expression in tumor correlates with ipsilateral recurrence of breast cancer: poor prognosis	Not determined	Maher and others (1998); Scholl and others (1994); Kluger and others (2004)
soluble IL-2R	Serum	↑	Good prognosis	Not determined	Gangemi and others (2012); Nicolini and others (2006)
IL-1B	Tumor	↑	Correlates with invasiveness, aggressiveness, and higher tumor grade: poor prognosis	Not determined	Chavey and others (2007); Jin and others (1997)

Different cytokines are listed in this table along with their prognostic values in breast cancer. The location in which each cytokine was evaluated is also indicated as well as their impact on survival (260).

© *Journal of Interferon & Cytokine Research*. 2015

4.2.1 Interleukin 1 (IL-1) is a cytokine known to have multiple effects in the inflammatory process. Human IL-1Ra is composed of 153 amino acid residues with a molecular weight of 17.3 kDa in the absence of glycosylation(263). Various studies have demonstrated that IL-1 is also involved in the activation of IL-8 - a pro-tumourigenic cytokine (264).

4.2.2 Interleukin 2 (IL-2) is from the gamma cytokine family which is an important activator of the immune system in response to cancer. It is composed of 153 amino acids (265)..According to some studies IL-2 targeted immunotherapy can destroy tumour cells(266). Under specific circumstances IL-2 is known as an immunoactivator since it activates T cell immune response causing differentiation of naive CD8⁺ cells into effector and memory effector cells. It also has an effect on NK cells in terms of its activation, cytolytic activity, homeostasis, and survival. The activated cells (T cells and NK cells) lead to cytokine production and increasing cellular response to perforins and granzyme B. On the other hand, clinical trials showed that IL-2 targeted immunotherapies can lead to devastating side effects, as very high doses of IL-2 are required to achieve a beneficial effect. Some of the side effects include hypotension, pulmonary edema, and heart toxicities. Various other strategies are being adopted to reduce these side effects by targeting the tumour microenvironment directly. IL-2 plays a part in developing autoimmunity; T regulatory (T- reg) cells are the only T cells that have a high affinity for IL-2(261).

4.2.3 Interleukin 4 (IL-4), is an immunosuppressive cytokine produced by the Th2 CD4 + T cell phenotype as well as mast cells and eosinophils(261). IL-4 has 129-amino-acids(267). It has dual functions acting both as a tumour suppressor and a tumour promotor(268). Its main role involves the regulation of lymphocyte proliferation and survival. However, it is also involved in cancer cell proliferation and is over expressed in various neoplasms e.g. gliomas, ovarian, breast, pancreatic, colon and bladder cancers(268). In a study by Venmar et al (269),it was demonstrated that higher IL-4 levels can increase chances of breast cancer tumour metastasis to the lung and liver, mediated through the overexpression of IL-4 Receptor alpha (IL4R α)(269).

4.2.4 Interleukin 6 (IL-6) is considered a small cytokine, consisting of 185 amino acids. Compared to other cytokines, IL-6 has the highest production frequency. It is known to cause breast cancer proliferation through formation of new blood vessels, resistance to therapy and cachexia (270).

It is also associated with promotion of multi drug resistance (MDR) to various cancer therapies(258, 271). Usually, IL-6 is secreted by different types of cells e.g., T and B cells, monocytes, fibroblasts, endothelial cells, keratinocytes, adipocytes, and mesangial cells. However, in several carcinomas including breast carcinoma, it is secreted by the tumour cells(271).

IL-6 is known to have two types of receptors onto which it exerts its action; transmembrane IL-6 receptor (mIL-6R) which is expressed on the cell membrane surface and soluble IL-6 receptor (sIL-6R) which is present in circulation(271). In patients with breast carcinomas, data shows increased levels of IL-6 in the circulation, as high as a tenfold increase compared to healthy controls (17). An increased level of IL-6 is also associated with recurrent disease, metastasis, lymph node infiltration and tumour stage. On the other hand, at the tissue level, IL-6 inhibits the growth of estrogen receptor positive breast cancer cells(272). Poor prognosis was observed in those patients who had raised IL-6 with a high tumour stage, thus reflecting the

role of IL-6 as a biomarker of tumour burden (272). Suppression/blocking of IL-6 or the IL-6/JAK/STAT3 pathway is potentially a good strategy to inhibit growth of breast carcinoma and for cancer therapy(271).

4.2.5 Interleukin 8 (IL-8), from the chemokine family of cytokines acts in various ways to influence tumour progression and metastases(273). Comprises of 72 amino acids (274). It is predominantly secreted by cancer cells(273) and involved in the formation of the tumour blood vessel supply by promoting Matrix Metalloproteinase (MMP) production, endothelial cell proliferation and survival (257). IL-8 has a multifaceted role in the progression on breast cancer. It has an autocrine/paracrine role, it is elevated significantly in higher stage tumours and higher levels in serum is related to lymph node status(275). Higher levels of IL-8 were found in patients with stage II and III carcinomas, and lymph node metastasis(275). In another study, it was shown that patients with breast carcinoma had a lower survival rate and shorter disease free survival when treated with chemotherapeutic drugs if they expressed higher levels of IL-8 or its receptor, CXCR1/CXCR2(258).

4.2.6 Interleukin 10 (IL-10) gene is found on chromosome 1. Human IL-10 has been shown to exist in solution predominantly as a homodimer, composed of two polypeptide chains of 160 amino acids each with a molecular weight of 18kDa(276). It is an anti-inflammatory cytokine that is involved in inhibition of T-cell and macrophage cytokine synthesis specifically inhibiting IL-2 and IFN- γ by Th1 and IL-4 and IL-5 by Th2 T cells(276). However, in cancer cells, it is known to have both pro- and anti-tumour effects in tumour proliferation(277). IL-10 is a cytokine that is also associated with tumour progression in breast cancer. Tumour Associated Macrophages (TAMs) are the main source of IL-10. TAMs that are in the tumour microenvironment secrete higher levels of IL-10 compared to TAMs from non-breast cancer cells(278).

4.2.7 Tumour Necrosis Factor α (TNF- α) is a multi-functional cytokine that is involved in tumour growth and tumour migration. Higher levels of TNF- α are also related with higher tumour stages and lymph node metastasis (279). It is produced by mononuclear macrophages and endothelial cells. It takes part in phagocytosis by activating and inducing T, B and NK cells through its receptors, TNFR1 and TNFR2. Elevated TNF- α levels are associated with higher stages of breast cancer, while stage I and healthy controls had normal TNF- α levels. It is also involved in tumour invasion and metastasis by increasing the expression of MMPs in breast tumours(257).

4.2.8 Transforming Growth Factor- β (TGF- β) is regarded as vital in cellular communication and controls a wide range of cellular activities e.g., cell proliferation, differentiation, and apoptosis. It has a dual role as a tumour suppressor and a tumour promoter (280). In the early stages it contributes to preventing the spread of cancer due to its anti-proliferative effects and causes cell cycle arrest (particularly in epithelial, endothelial, and haematopoietic cells). In the later stages it has a counteractive effect as tumour cells evade it and instead can proliferate further in its presence(260). This was demonstrated using various mouse models(260). In bone metastases, which is a regular site of metastases for breast carcinomas, TGF- β stimulates the neoplastic cells to produce osteolytic cytokines e.g. Parathyroid Hormone related Protein (PTHrP) leading to detrimental effects(280).

4.3 CYTOKINES IN OESTROGEN RECEPTOR POSITIVE AND TRIPLE NEGATIVE BREAST CANCER.

The hormone status of breast carcinomas is vitally important in the expression of various cytokines. Oestrogen receptor (ER) positive tumours have been shown to have a better prognosis compared to ER negative tumours due to associated hormone therapies that are able to target these tumours(125). With regards to IL-6, hormone sensitive cells (ER and PR positive tumours) show a much higher response to IL-6 than hormone insensitive cells. The sIL-6R is mainly expressed by ER positive tumours while ER negative tumours secrete mIL-6R(271).

IL-6 also suppresses ER negative cells in an autocrine fashion. oestrogen levels in the tumour microenvironment are also influenced by IL-6 levels. It activates enzymes e.g., oestrone sulphatase, aromatase and 17 β -hydroxysteroid dehydrogenase which are oestrogen generating enzymes.

Oestrogen sulphate is regarded as an oestrogen reservoir as it can survive in circulation for a longer time compared to oestrogen(271). In terms of quantity, Oestrogen sulfates are the most significant type of circulating estrogens both during and after menopause. Breast tissues from mammary cancer patients contain significant amounts of estrone sulfate and estradiol sulfate. Different hormone-dependent mammary cancer cell lines (MCF-7, R-27, and T-47D) convert estrone sulfate to estradiol (E2) at a very high proportion, but no or very little conversion has been observed in the hormone-independent mammary cancer cell lines (MDA-MB-231, MDA-MB-436). The control of sulfatase and 17 beta-hydroxysteroid dehydrogenase activities, which are crucial steps in the formation of estradiol in the breast, may open up new possibilities in the treatment of hormone-dependent mammary cancer (281). Higher oestradiol levels in ER positive cells is related to IL-6 as it converts oestrone to oestradiol in these types of malignant cells(272). The mechanistic pathway by which IL-6 effects its action on ER-positive breast cancer is through the phosphorylation of tyrosine 705 of the Signal transducer and activator of transcription (STAT)3. The ER-positive breast cancer cells are dependent on paracrine IL-6 from the tumor microenvironment(282). A study of ninety-six metastatic breast cancer patients observed that patients with multiple metastatic sites had higher serum IL-6 values compared to those with only 1 metastatic site (8.15 +/- 1.7 pg/ml and 3.06 +/- 6.6 pg/ml; $p < 0.001$) however, serum IL-6 levels were not associated with age, menopausal status, nodal involvement, body-mass index, tumour grade, histological and receptor status in breast cancers. Survival was three times better in those patients with lower serum IL-6 at the end of the follow up period(283). A meta-analysis involving 10,137 cases and 15,566 controls did not find any correlation between breast cancer risk and the commonly accepted functional Single Nucleotide Polymorphism (SNP) -174G within the IL-6 gene promoter region. Furthermore, there were no ethnic differences found for this SNP as a breast cancer risk factor(284). However, in another study by Ma et al,(285) it was shown that serum levels of IL-6 were significantly higher in ER positive breast carcinomas compared to ER negative breast cancers. It was speculated that the IL-6 produced by the breast carcinoma did not contribute as much to the total serum IL-6, even though it was in large amounts.

Triple Negative Breast Cancer (TNBC) is notably the most aggressive type of breast carcinoma; it usually presents at a younger age, higher cancer stage at presentation, higher risk of metastases, and a high density of lymphocytic infiltrate on histological analysis. Because of its lack of therapeutic targets and poor clinical outcome it is considered to be a major clinical challenge(286).

In TNBC, IL-6 and IL-8 are the major cytokines involved in tumour cell proliferation and poor prognosis (287) since these two cytokines are important for anchorage-independent colony formation as well as resistance to apoptosis (288). IL-8 serum levels in ER negative tumours were significantly higher as compared to ER positive tumours (275). Serum levels of IL-6 and IL-8 were found to be critical for growth of TNBCs but not ER positive BC cells; both IL-6 and IL-8 are critical for anchorage-independent colony formation as well as resistance to apoptosis and the combined expression of IL-6 and IL-8 is critical for tumor growth and correlates with poor prognoses in BC patients (287). IL-8 is highly expressed in ER negative and/or HER2 + breast cancers and increases the invasiveness and metastatic potential of both ER positive and ER negative breast cancer cells (288). IL-8 promotes Cancer Stem cell (CSC) self-renewal and invasion by binding to its cognate receptor CXCR1 on CSCs (289). Decrease in IL-6 levels leads to increased CD4⁺ and CD8⁺ T cells response which enhances the anti-tumour response (286).

Levels of IL-10 are significantly higher in patients with ER negative tumours as compared to ER positive tumours due to over expression of this cytokine in ER negative tumours (277). PR positive tumours express much lower levels of IL-10 compared to other breast cancer subtypes (20). However, in one study it was shown there was no association between serum IL-6, IL-8 levels, and PR status (275).

Other than cytokines and chemokines, stromal/tumour cells also produce profuse amounts of immunosuppressive TGF- β , IL-10, and PGE2 all of which combined promote a state of chronic inflammation within the tumour microenvironment. The inflammatory milieu thus contributes to tumour proliferation and metastases through an enabling immunosuppressive environment by suppression of effector immune cells and the recruitment of regulatory T cells and myeloid-derived suppressor cells (290).

4.4 MEASUREMENT OF CYTOKINES

Cytokines are regarded as important biomarkers in monitoring disease progression, inflammation and the effect of drugs affecting the immune system (256).

Various in-vitro assays for cytokines are available that measure the analyte at both the mRNA and protein level. These assays can be performed on body fluids, cells, or biopsy material. Cytokine protein assays tend to be more instructive due to the unstable nature of mRNAs which many preclude detection by currently available methods (291). Due to the limitation of lack of blood samples, we utilized FFPE tissue material for mRNA of cytokines in this study.

4.4.1 mRNA methods:

Both in research labs and in regular clinical testing, formalin-fixed paraffin-embedded (FFPE) tumor material is an invaluable resource for the investigation of RNA-based biomarkers; although RNA in FFPE samples may degrade with time into smaller fragments, as shown by the much lower RNA quality indicators than in fresh/frozen tissues, it can still be used to successfully analyze gene expression, according to Ribeiro-Silva et al. (292) since only small fragments of mRNAs are required to reconstruct an expression profile of a given gene (293). On the other hand, during the formalin fixation process, RNA is vulnerable to cross-linking to proteins. To address the issues with quantification, amplification, and isolation Specht et al. (294) developed a technique that combines laser-capture microdissection (LCM) with quantitative reverse transcription polymerase chain reaction (QRT-PCR), using Proteinase-K

for RNA digestion and extraction. This has made it possible to analyze quantitative gene expression from tissue banks that have been archived. In breast cancer as an example, routine clinical testing using QRT-PCR assays from FFPE based gene expression analysis panels such as Oncotype DX test (<http://www.genomichealth.com>), MammaPrint test (<http://www.agendia.com>) and Prosigna Breast Cancer Prognostic Gene Signature Assay (<https://www.nanostring.com>) are now readily available(293).

4.4.2 Flow cytometry: Using this method, cytokines can be studied at single cell level. It is a combination of intra cellular staining methods and polychromatic flow cytometry. The main advantage of using polychromatic flow cytometry is that it allows for the detection of multiple cytokines and its various characteristics within a single cell, hence complex cytokine phenotype detection. Limitations for this method include; not all cells produce a given cytokine at the same time, some cytokines are undetectable in cells that are not stimulated by in vitro stimulation and if a cytokine population is present within the cells it does not necessarily mean that it will be released and exhibit its effects in vivo(295).

4.4.2 Enzyme linked immunosorbent assay (ELISA): These assays are the most common and easily available ways for analyzing and measuring cytokines. Approximately 100µL of sample is required. ELISA-based cytokine assays can be performed using direct or indirect methods. The indirect or “sandwich” ELISA, in which a targeted cytokine is sandwiched between a capture antibody bound to the plate and an enzyme- labelled detection antibody, is the most frequently utilized method today(291).

Cytokine Immunoassays may give false positive or negative results depending upon the quality of the capture and detection antibodies. Since these assays are performed on complex biologic fluids, background assay noise with other proteins in these samples may interfere with immunoassay interpretation (291).

4.4.3 Bead based Immunoassays: The solid support for the immunoassay is the microsphere beads that have been embedded with fluorophores. Low amounts of cytokines can be measured with this method. It is also highly sensitive and specific. Flow cytometry is then used to detect the fluorophore intensity. This method is highly sensitive and reproducible due to the high affinity antibodies and fast reaction kinetics. Alternatively, bioluminescent proteins can be used as an alternative for fluorophores - as the light is made from biochemical reaction hence eliminates the background light from auto fluorescence(296).

4.4.4 Chip/Slide based protein arrays: This method is a variation of the bead-based immunoassay and can measure minute amounts of cytokines in a sample. Using this method, cytokines can be detected from cell culture, wound fluids, pancreatic fluid and more. The downside of these methods (bead based and chip based) includes low precision and lack of reference methods (297).

4.4.5 Rolling circle amplification: This method is used to detect cytokines proteins on a microarray-based chip device. The specific antibody to the cytokine is attached to biotin. A secondary antibody that binds to the biotin-cytokine-antibody complex is then added. Circular DNA, nucleotides and DNA polymerase is then added. This leads to the production of a long DNA concatamer, a process known as rolling circle amplification. This method is regarded as equally or higher in sensitivity as compared to a standard ELISA protocol. Other methods

include the Biobar code and Proximity Ligation. Limitations of immunoassays include long preparation times, potential cross-reactivity of antibodies and the fact that some immunoassays can only analyse one analyte at a time(298).

4.4.6 Antibody Based Flow Assays

4.4.6.1 *Immunochemistry based liquid chromatographic separation*

In this method liquid chromatography is used and several different columns with different properties are set up to separate cytokines. Once the cytokines have been separated, they pass through a reactor containing fluorescently labelled antibodies, and the cytokine-antibody complex is measured(299).

4.4.6.2 *Capillary Electrophoresis/Immunohistochemistry based separation*

In this method, cytokines that are labelled with a fluorescent dye are matched with antibodies that are bound to a silica capillary or chip device.

The cytokines are then detected using LIF (Laser Induced Fluorescence). This method has very low detection limits and high productivity as the beads can be used about 200 times(299).

4.4.7 Aptamer Assays: Aptamers are described as oligonucleotide strands of either DNA or RNA that are specific to the analyte in question (in this case particular cytokines). The advantages of using aptamers are that they are very stable, can be synthetically prepared hence eliminating the need to use animals to generate them and can be detected by single step assays. The major disadvantage is the number of aptamers available for the different types of cytokines is limited.

4.4.8 Mass spectrometry: Separates particles based on their mass and charge; however, for cytokines this technique can pose significant challenges. One challenge is that there are many matrix interferences and suppression of particles. Mass cytometry is a process whereby selective mass tags can be used to allow for mass spectrometric detection of different analyses, hence more specific for cytokine detection.

4.4.9 Electrochemical Methods: These methods include Amperometric based protein measurements and electrochemiluminescence. These methods involve electrochemical immunoassays for measurements of cytokines. Advantages of this method include high sensitivity and that the equipment is inexpensive(256).

4.5 OBJECTIVES

The objectives of this study were to:

- a) Describe the pro- and anti-inflammatory cytokine profiles in the tumour microenvironment of ER positive and Triple Negative breast carcinomas of Kenyan women using microarray gene expression
- b) Validate the results using an Enzyme Linked Immunosorbent Assay (ELISA) platform.

4.5.1 Strategy of Approach

Quantitative RT-PCR (qRT-PCR) offers a precise relative quantification of the mRNA levels of biomarkers in formalin-fixed paraffin-embedded (FFPE) tumor tissue samples and enables the selective determination of mRNA expression levels in cancer cells (300).

First, we conducted gene expression analysis of mRNA extracted from FFPE breast cancer tissues using the expression of 84 key genes involved in mediating communication between tumour cells and the cellular mediators of inflammation and immunity (Human Cancer Inflammation & Immunity Crosstalk PCR Array (<https://www.qiagen.com/~media/genetable/pa/hs/18/pahs-181za>))

We then validated selected unregulated and down regulated cytokine genes at the protein level by using protein extracts from FFPE tissues on the Enzyme Linked Immunosorbent Assay platform. Using a sheep model, Addis et al (301) have previously demonstrated that protein extracts from FFPE tissue provide outcomes that are comparable to those of fresh tissue extracts. Enzyme Linked Immunosorbent assays (ELISA) is a common immunoassay method that offers several significant benefits, including amenability to automated ion, diffusion of compatible instrumentation and assay platforms, and availability of antibodies optimised for this methodology. Addis et al(301) demonstrated that all antibodies tested in their experiment yielded encouraging results and noted that ELISA application may be an effective method for testing protein extracts form FFPE tissues.

The limitation of this approach was that plasma or serum samples were unavailable to conduct these experiments. We thus ensured reliability of our results by using an assay that had a high level of sensitivity and linearity with screened and verified capture and detection antibodies as well as ensuring that the control Optical Density (D) readings were within the pre-determined cut-off values.

4.6 METHODS

4.6.1 Objective 1: To describe the pro- and anti-inflammatory cytokine profiles in the tumour microenvironment of ER positive and Triple Negative breast cancers. Triple negative breast cancer is the most aggressive breast cancer subtype- it is also the subtype that is considered to afflict women of African ancestry in whom it occurs at a much younger age(25, 41, 302). Currently, there are no effective therapeutic options for this subtype hence identifying unique biomarkers that may have a potential therapeutic benefit is important. On the other hand ER positive breast cancers are the most prevalent subtype. HER2 positive tumours although have multiple therapeutic targets. Thus, comparison between Triple negative and ER positive tumours is of clinical and research interest.

4.6.2 Selection of Cases

Tissues blocks of tumour tissue from ten cases each of TNBC and ER positive breast cancer were randomly selected from the list of 232 pathologically confirmed breast cancers. (https://drive.google.com/file/d/1VFGgTP1_PsQeyHSbofzW8jIcc-yxP4_-/-/view?usp=sharing).

4.6.3 Preparation of Cases

The excess paraffin wax from each of the selected tumour blocks was trimmed off with a scalpel. For each of the ten cases, 3 sections of 10 µm thickness were cut on a rotary microtome.

The sections were floated in a water bath at 56°C and picked onto glass slides. The slides were then placed on a hot plate at 60°C for approximately 10 minutes. The sections were brought down through three xylene baths for dewaxing, and three absolute ethanol baths to facilitate clearing and rehydration of tissues. Finally, tissue sections were washed well in running tap water for 5 minutes. Once dried, the tumour area on the previously stained Haematoxylin and Eosin (H&E) slide was marked out with indelible ink (**Figure 28**). The unstained slide sections were superimposed on the marked H/E slide to match tumour area (**Figure 29**) which was scrapped off with a sterile scalpel into a clean Eppendorf tube.



FIGURE 28: TUMOUR AREA
MARKED WITH INDELIBLE INK

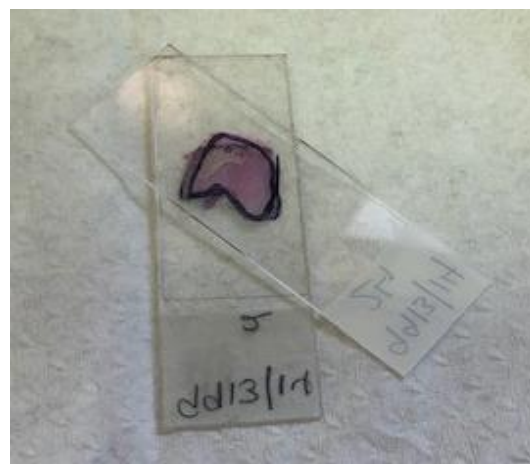


FIGURE 29: TUMOUR AREA SUPERIMPOSED ON
UNSTAINED SLIDE

4.6.4 Extraction of RNA

This was performed according to the manufacturer's instructions using the Qiagen RNEasy FFPE kit (**APPENDIX VIII**). RNAse free environment was maintained by ensuring the bench spaces were cleaned with detergent, using RNAse free water, and regularly changing gloves, and using disposable sterile RNAse free plastic ware.

4.6.5 RNA Integrity testing

RNA yield and purity were determined by spectrophotometry (A260/A280) using the Nano Drop 1000 instrument (Thermo Scientific, Wilmington DE) (**APPENDIX IX**). The RNA integrity (RIN) analysis was performed at the Skin and Hair Research laboratory (University of Cape Town) on the Bio analyser Agilent instrument using the RNA 6000 Pico kit (Agilent Technologies, Santa Clara, CA) according to the manufacturer's protocol.

4.6.5 Final selection of cases for downstream conversion of total RNA to cDNA: There were a total of 12 cases out of the original 20 that had more than 40ng/μL of RNA extracted for downstream experiments as recommended by the manufacturer instructions (41) (**APPENDIX X**).

Four cases each of TNBC and ER/PR positive breast cancers were eventually selected as follows based on the amount, and purity of RNA (**Tables 15 and 16**). The cases were selected based on the concentration and purity of the RNA and available budget. We started with ten randomly selected cases each of ER and TNBC of which 12 cases each of TNBC and ER had more than 40ng/μL of RNA extracted for downstream experiments as recommended by the manufacturer. Only four cases each passed the criteria for selection based on the purity and highest quantity of RNA.

TABLE 15: SELECTED LIST OF ER POSITIVE CASES

Amount of RNA (ng/μL)	ng (ng/μL)	A260/A280	A260/A230
269.6		1.96	2.13
152.9		1.95	2.07
161.6		1.94	1.94
201.5		1.96	2.16

TABLE 16: SELECTED LIST OF TRIPLE NEGATIVE CASES

Amount of RNA (ng/μL)	ng (ng/μL)	A260/A280	A260/A230
82.1		1.94	2.08
64.5		1.91	2.01

91.4	1.96	2.18
57	1.91	2.05

4.7 FIRST STRAND cDNA SYNTHESIS FROM TOTAL RNA

4.7.1 Preparation of Genomic DNA elimination mix.

For each case 250 ng of RNA was placed in a sterile PCR tube as per manufacturer's recommendations and mixed with buffer GE (2ul) and made up to a total volume of 10ul with water. The contents of the tube were mixed gently with a pipette and followed by brief centrifugation. The DNA elimination mix was then incubated at 42°C for 5 min and placed immediately on ice for 1 min on the Gene Amp PCR system 9700 from Applied Bio systems.

4.7.2 Preparation of Reverse Transcription (RT) mix for cDNA Synthesis

The RT mix (**Table 17**) was added to PCR tubes containing 10uL of the genomic DNA elimination mix to make a total of 20uL as starting material. This was gently mixed followed by brief centrifugation. The mixture was incubated at 42°C for exactly 30 min followed immediately by incubation at 95°C for 5 min. The reactions were placed on ice and stored overnight at -20°C.

TABLE 17: REVERSE TRANSCRIPTION (RT) MIX COMPONENTS

Component	Volume per reaction (μL)
5XBufferBC3	4
Control P2*	1
cDNA synthesis Enzyme mix	1
RNase Inhibitor	1
RNase free water	3
Total volume	10

**The inbuilt external RNA Control P2, allows detection of inhibitors of reverse transcriptase reaction that could compromise the efficiency of cDNA synthesis.*

4.7.3 Pre-amplification of cDNA Target templates

Five microliters of the cDNA synthesis reaction from above was mixed with 20uL of the pre-amp mix (**Table 18**). The contents of the tube were mixed gently by pipetting up and down followed by a brief spin. This step was repeated for all the 8 cDNA samples. The PCR tubes were then placed in the thermocycler which was programmed for the PCR conditions (**Table 19**).

TABLE 18: PRE-AMPLIFICATION MIX COMPONENTS

Component	Amount per sample (μL)
RT2 PreAMP PCR master mix	12.5
RT2 Pathway Primer mix	7.5
Total volume	20

TABLE 19: PCR CONDITIONS

Cycles	Duration	Temperature	Comments
1	10min 15 secs	95° C	Hot start Taq Polymerase is activated by this step.
8	2min	40° C	Hold

The tubes were removed from the cycler and placed on ice. 2μL of side reduction reducer (to eliminate residual primers) was added to each pre-amplified reaction, mixed gently by pipetting up and down and spun briefly. Tubes were incubated at 37°C for 15 minutes followed by heat inactivation at 95°C for 5 min. Immediately 84μL of nuclease free water was added to the tubes giving a total preamplifier cDNA volume of 109 μL; this was mixed well. Reactions were stored overnight at -20°C

4.7.4 Application of Pre-Amplified cDNA to RT-PCR Profiler microarray plates

The RT2 SYBR Green Mastermix was centrifuged briefly to mix the contents; this was performed at room temperature as the master mix contains Hot start DNA Taq Polymerase which is active only after heat activation. The reaction mix below was prepared in a loading reservoir (**Table 20**).

TABLE 20: PCR COMPONENT MIX

Component	Amount per sample (μL)
2XRT2 SYBR Green Mastermix	1350
Pre-amplification cDNA reaction mixture from step above 102 per sample	
RNase free water	1248
Total volume	2700

25 μL of the reaction mixture was dispensed in each well of the RT2 profiler PCR array using an 8 tip multi-channel pipette. The RT2 PCR profiler array was sealed carefully with an optical adhesive film.

The array plate was centrifuged at 1000g for 1 minute at RT to remove any air bubbles. The PCR array was then placed on ice while the cycling program on the Roche Light cycler 480 was set up as below (**Table 21**). The array was then placed in the real time cycler and the run started.

TABLE 21: LIGHT CYCLER 480 CONDITIONS

Cycles	Duration	Temperature	Comments
1	10min	95° C	Hot start Taq Polymerase is activated by this step.
45	15 seconds	45° C	Fluorescence data acquired
	1 min	60°C	

Components of the RT profiler are illustrated in APPENDIX IV.

4.7.5 Data quality Control (QC)

This method automatically selects an optimal set of internal control / housekeeping / normalization genes for the analysis from the full plate on the PCR Array. The software measures and identifies the genes with the most stable expression via a non-normalized calculation. The Cycle Threshold (CT) values for these genes are then geometrically averaged and used for the delta – delta CT (ddCT) calculations, (**APPENDIX X1**).

4.8 DATA MANAGEMENT AND ANALYSIS

The data analysis related to the RT profiler PCR arrays was done using the free web-based RT profiler PCR data analysis software (<https://geneglobe.qiagen.com/ke/my-geneglobe/analysis>). The software using the delta Ct ($\Delta\Delta Ct$) method to normalize the expression of the genes of interest with an optimal set of internal control / housekeeping / normalization genes from the full plate on the PCR Array (**APPENDIX XII**). The software generated scatter plots and cluster–grams.

Correlation of the molecular data with tumour subtypes was performed using Stata version 13.1 (Stata Corp LP, 4905 Lakeway Drive, College Station, TX 77845, USA).

4.9 OBJECTIVE 2

To validate the microarray gene expression results using an Enzyme Linked Immunosorbent Assay (ELISA) platform.

This part of the experiment involved validation of the gene expression results obtained from the RT profiler experiment. Validation of 12 selected upregulated and downregulated genes based on the microarray gene expression experiment results and the published literature was performed using a custom designed Multi-Analyte ELISA array Kit (<https://www.qiagen.com/dk/products/discovery-and-translational-research/functional-and-cell-analysis/elisa-assays/multi-analyte-elisarray-kits/>).

A panel of 12 cytokines (**Table.22**) was analysed on the custom designed Multi-Analyte ELISA array plates.

The smallest sample size was selected for each group to demonstrate a fourfold change in cytokine levels. A sample size of 33 cases per subgroup (ER positive and TNBC) were required for demonstration of a fourfold change (**APPENDIX XIII**).

TABLE 22: MULTI-ANALYTE CYTOKINE ELISA ARRAY PLATES

1	2	3	4	5	6	7	8	9	10	11	12
IL6	IL8	IL10	IL12	IL17A	IFN γ	GM-CSF	TGF β 1	RANTES	MIG	MDC	TNF α

4.10 MATERIALS AND METHODS

4.10.1 Protein extraction from FFPE tissues:

This was performed with the Qiagen Qproteome® FFPE Tissue kit as per the manufacturer's instructions. Briefly 33 cases each of TNBC and ER positive BC tissue were selected based on tumour preservation and > 80% tumour on the slide. Three 10 μ m scrolls were cut from paraffin wax embedded tissue blocks and placed in Eppendorf tubes. The protocol below was followed for protein extraction.

The protein yield and purity were determined by spectrophotometry (A260/A280) using the Nanodrop 1000 instrument (Thermo Scientific, Wilmington DE). (**Table 23**).

TABLE 23: CONCENTRATION OF PROTEINS IN THE PROTEIN EXTRACTS

Study No.	MG/ML	A280	A260/A280	Study No.	MG/ML	A280	A260/A280
4000	12.101	12.11	1.7	4043	50.628	50.63	2.22
4001	5.912	5.91	3.99	4044	29.725	29.73	2.69
4002	18.691	18.69	2.99	4045	15.35	15.35	4.07
4003	46.126	46.13	2.21	4046	30.04	30.09	2.19
4004	8.58	8.58	2.88	4047	19.008	19.01	2.5
4005	15.3	15.30	2.59	4048	8.008	8.01	3.13
4006	18.999	18.99	3.49	4049	25.559	25.55	2.7
4007	18.081	18.08	3.17	4050	18.198	18.20	3.78
4008	5.081	5.08	1.59	4051	12.908	12.91	3.14
4009	22.278	22.28	2.68	4052	6.675	6.67	2.81
4010	19.602	19.60	3.57	239	17.178	17.18	2.49
4011	9.502	9.50	2.61	1	18.245	18.24	2.79
4012	8.64	8.64	2.61	11	18.617	18.62	2.65
4013	7.681	7.68	3.19	12	36.611	36.61	2.99
4015	4.188	4.19	3.24	21	9.473	9.97	

Study No.	MG/ML	A280	A260/A280	Study No.	MG/ML	A280	A260/A280
4016	16.379	16.38	3.5	47	8.699	8.69	2.94
4017	49.93	49.93	2.24	56	11.917	11.92	2.62
4018	5.517	5.52	0.32	72	17.727	17.73	2.43
4019	10.02	10.02	2.91	90	10.373	10.37	2.91
4020	12.547	12.55	3.09	125	14.299	14.30	2.79
4021	22.626	22.63	2.89	142	29.204	24.20	2.29
4022	9.068	9.07	2.78	147	14.89	14.89	2.49
4023	19.344	19.34	2.76	60	14.611	14.61	3.13
4024	17.639	17.64	4	136	11.011	11.01	2.26
4025	12.376	12.38	3.85	225	19.108	19.11	2.76
4026	28.715	28.72	2.29	98	20.792	20.74	2.62
4027	11.508	11.51	2.51	173	65.299	65.30	2.99
4028	14.191	14.19	3.24	292	24.646	24.65	2.82
4029	19.775	19.77	2.94	169	13.715	13.72	2.83
4030	17.811	17.81	3.57	192	26.423	26.42	2.99
4031	12.538	12.54	3.2	260	6.358	6.36	3.53
4032	13.383	13.38	2.88	276	13.021	13.02	2.95
4033	6.945	6.95	3.4	261	11.992	11.99	2.97
4034	19.229	19.23	2.7	170	12.947	12.95	2.82
4035	46.879	46.87	2.37	251	16.024	16.02	3.19
4036	37.652	37.65	2.4	270	8.87	8.87	2.76
4037	15.679	15.68	2.17	177	20.889	20.89	2.72
4038	20.243	20.49	3.15	293	36.857	36.86	4.1
4039	19.927	19.73	3.05	183	17.658	17.66	2.79
4040	7.691	7.69	2.86	273	8.862	8.86	2.6
4041	26.913	26.91	3	262	9.969	9.97	2.44
4042	20.629	20.62	2.36	279	4.582	4.55	2.66
317	13.526	13.53	2.71	319	16.551	16.55	2.3
190	2.09	2.09	4.24	4053	11.697	22.70	2.45
316	13.309	13.31	2.85				

Further confirmation of the protein content was performed on three cases each of ER positive and triple negative breast cancer protein extracts in FFPE extraction buffer at the Proteomics lab, Stellenbosch University as per below:

4.10.3 Samples

Six (6) randomly chosen protein extract samples, labelled 53, 69, 75, 138, 203 and 226 were submitted for analyses.

4.10.4 Analysis Details

An aliquot of the protein solutions was precipitated using methanol/chloroform and re-suspended in 50mM Tris-HCl pH7.8 containing 4M urea and 2% SDS. This solution was

diluted ten times for protein concentration determination using the DirectDetect® system (Merck). The concentrations were: 53 = 7.5 mg/mL, 69 = 5.26 mg/mL, 75 = 7.94 mg/mL, 138 = 4.31 mg/mL, 203 = 3.25 mg/mL, and 226 = 5.45 mg/mL. The samples were standardized to sample the lowest concentration (sample 203) and 20 µg removed for on bead digest. The system suitability sample (a four-protein mixture) showed the peptides were both retained and eluted as expected judging by the total ion chromatogram (TIC) and the blank injection before the analysis showed no major contaminants (**Figure 30**). The samples were digested with trypsin and the total ion chromatograms (TICs) indicated that all the samples digested successfully (**Figure 31**).

RT: 10.00 - 120.00

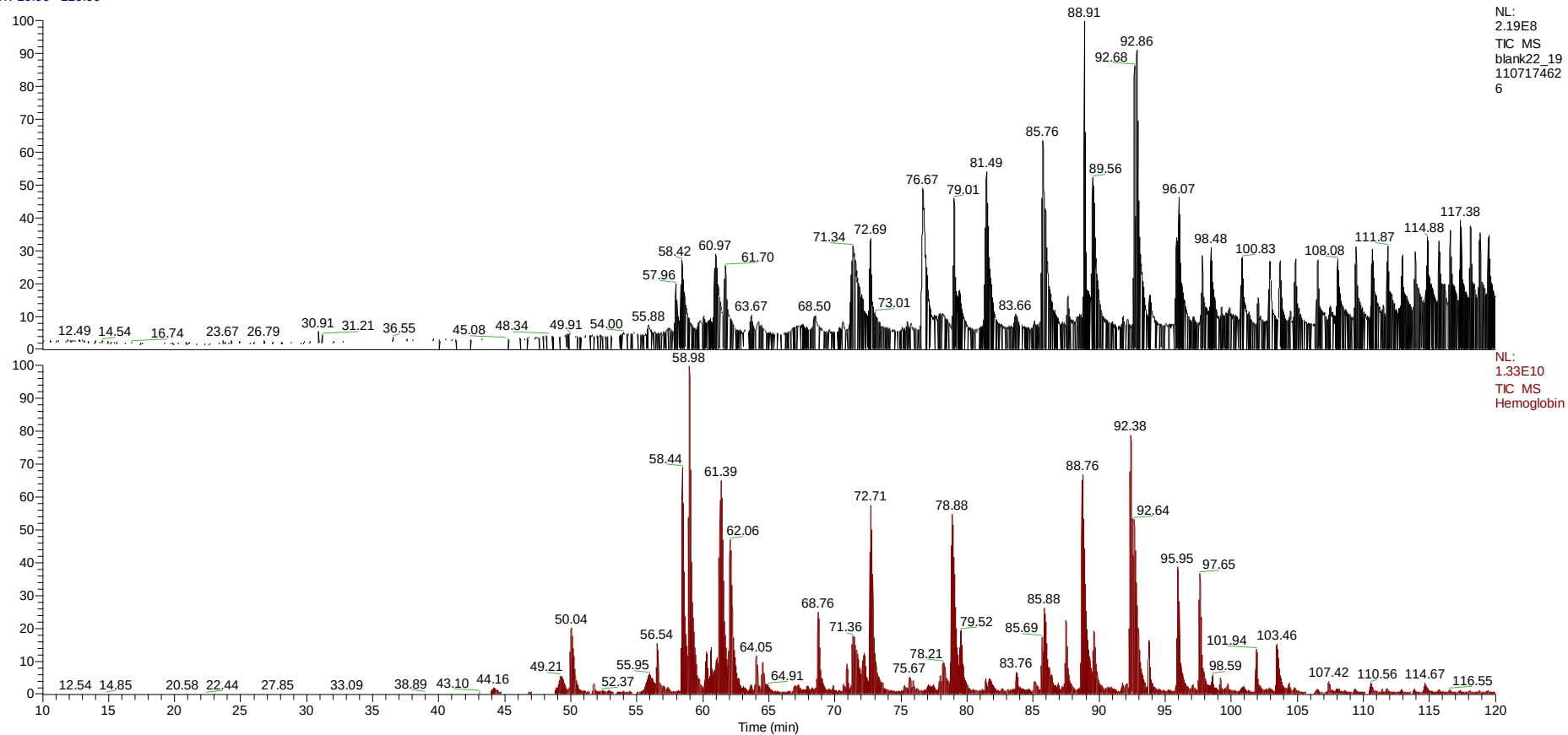


FIGURE 30: TICs OF THE SYSTEM BLANK AND SUITABILITY MIXTURE INJECTION.

RT: 30.00 - 110.00

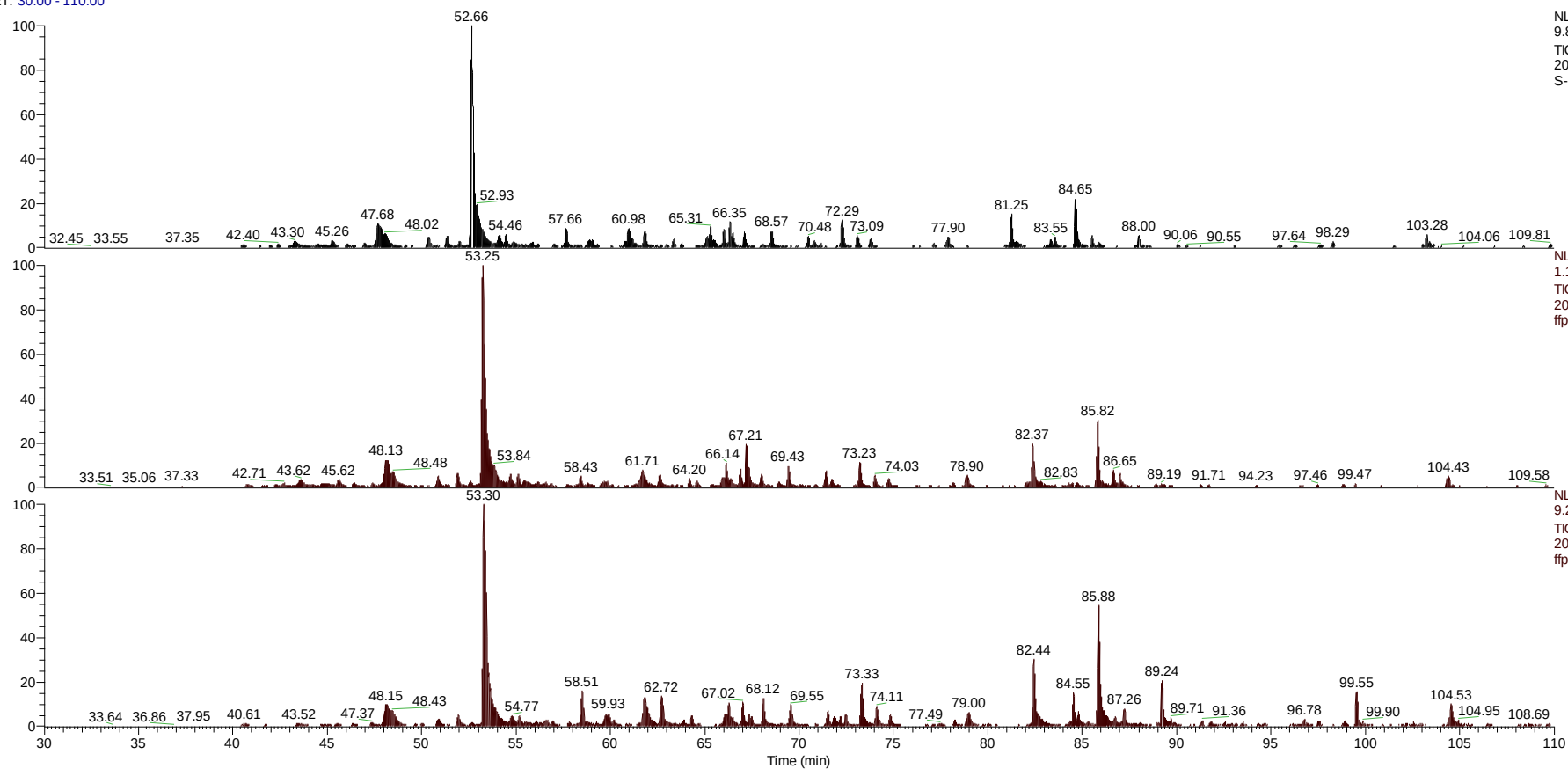
NL:
9.86E9
TIC MS
201912_S
S-FFPE-53

FIGURE 31: TICs OF THE SAMPLES IN ASCENDING NUMERICAL ORDER WITH SAMPLE 53 AT THE TOP.

CYTOKINES IN BREAST CANCER

RT: 30.00 - 110.00

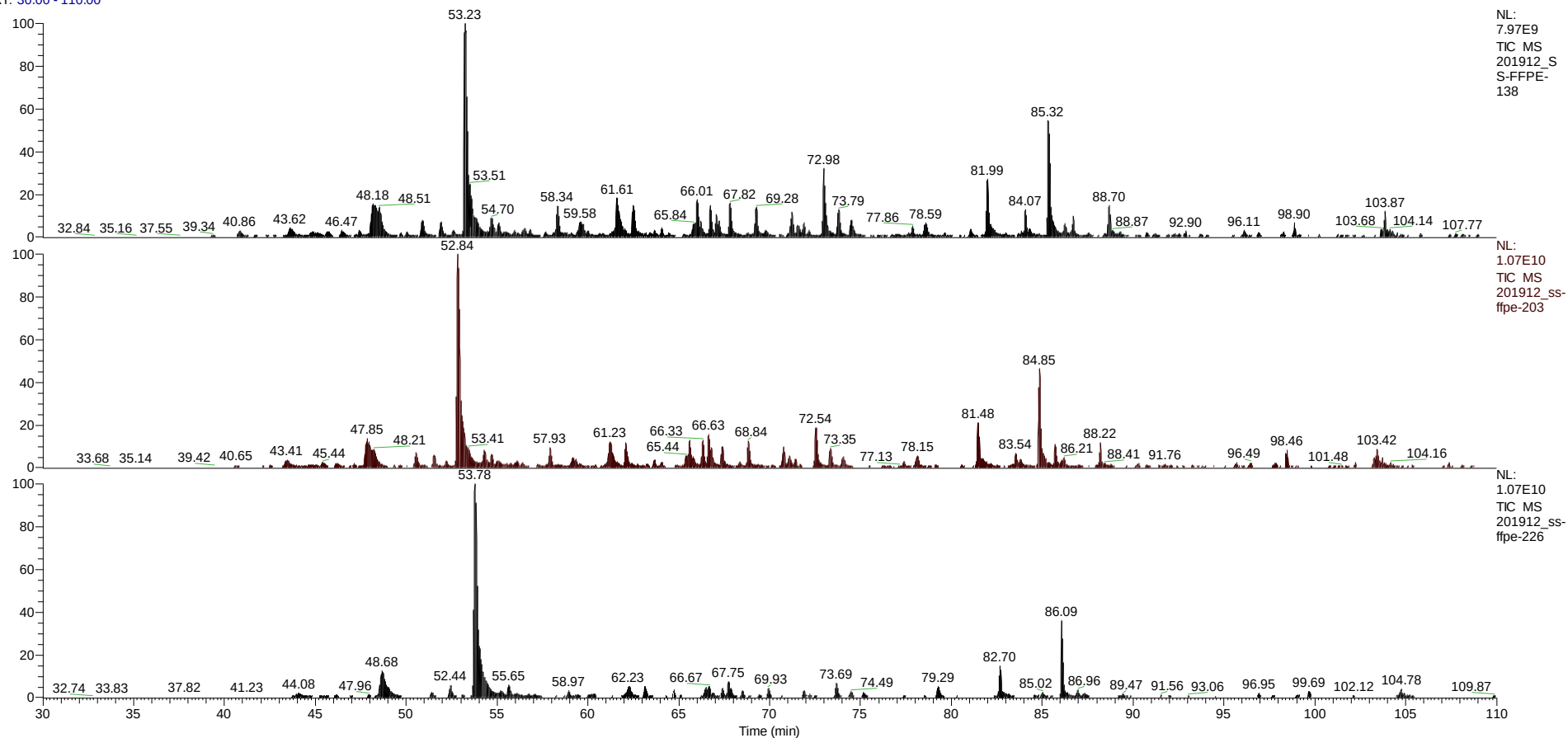


FIGURE 32: TICs OF THE SAMPLES IN ASCENDING NUMERICAL ORDER WITH SAMPLE 53 AT THE TOP.

4.10.5 Cytokine Validation Experiment using the Multi-Analyte ELISA array

The panel of 12 cytokines was analysed on the custom designed Multi-Analyte ELISA array plates as per manufacturer's instructions (**APPENDIX XIV**). Briefly, 50 µL assay buffer was added into each well of ELISArray plate. 50 µL of samples and control samples were transferred into the appropriate wells of the ELISArray plate and incubated for two hours at room temperature in lab bench drawer protected from light. The wells were washed three times using wash buffer diluted 1:10. This was followed by 100 µL of Detection Antibody Solution. and Incubation of 2 hours at room temperature in lab bench drawer protected from light after which the plate was washed three times using wash buffer diluted 1:10. 100 µL Avidin – HRP was added, and plate incubated for 30 minutes followed by four washes using wash buffer diluted 1:10. 100 µL Development Solution was then added and plate incubated for 15 minutes in the dark followed by 100 µL Stop Solution. The plate was read at an OD of 450 nm within 30 minutes.

4.10.6 Data Analysis

This was performed as per manufacturer's instructions provided in the handbook (www.sabiosciences.com/ELISA.php). We started with the first set of data in each plate. To get the second set of data (Background OD removal), figures in the second row were subtracted from the figures in the first row of each case. The third set of data (Negative control removal) was arrived at by subtracting the figures in Neg-A row from the figures in all other rows. Raw data are presented as Tables (**APPENDIX XV**).

4.11 RESULTS

4.11.1 Results Objective 1: To describe the pro- and anti-inflammatory cytokine profiles in the tumour microenvironment of ER positive and Triple Negative breast cancers using a microarray gene expression method

The ER positive group was defined as the control group and the TNBC group the test group. All four samples each of the TNBC and ER positive cases passed PCR Array reproducibility, Real Time (RT) efficiency and DNA Genomic Contamination analysis (**APPENDIX XVI**). The STAT1, CCL18, GZMB, CCL21 and TP53 genes were automatically selected as internal control/Housekeeping/normalization genes.

The expression of the genes in the test group, TNBC, were normalized per the expression of the control group, ER positive (**Table 24**) IL10 and NOS10 were the most upregulated genes in TNBC with expression levels more than 60x those of ER positive group. This was followed by IL12A, IDO1, HGDC, CXCL5, CCR10 with more than 10x the expression levels found in the ER positive group (**Tables 24 and 25**). VEGFA, CCL5, RPLP0 and CCR9 were the most downregulated genes with more than 10x less expression levels in TNBC (**Tables 24 and 26**)

TABLE 24: FOLD REGULATION COMPARISON BETWEEN GROUP 1 (TNBC) AND CONTROL GROUP (ER POSITIVE) WITH CORRESPONDING P-VALUE

Control Group	Test Group	Fold Regulation cut off	p-Value cut off
Control Group	Group 1	2	0.05

Position	Gene Symbol	Fold Regulation	p-Value	Comments
A01	ACKR3	-2.68	0.000000	
A02	AICDA	8.91	0.000000	
A03	BCL2	-3.77	0.000000	
A04	BCL2L1	-5.33	0.000000	
A06	CCL2	-2.47	0.000000	
A10	CCL28	-8.48	0.000000	
A11	CCL4	4.61	0.000000	
A12	CCL5	-12.86	0.000000	
B02	CCR10	11.68	0.000000	
B04	CCR4	-2.26	0.000000	
B05	CCR7	-2.18	0.000000	
B06	CCR9	-10.44	0.000000	
B09	CSF2	7.65	0.000000	
C03	CXCL12	-5.15	0.000000	
C04	CXCL2	2.65	0.000000	
C05	CXCL5	13.23	0.000000	
C06	CXCL9	-2.14	0.000000	
D01	EGFR	-2.01	0.000000	
D07	HIF1A	6.71	0.000000	
D08	HLA-A	-2.45	0.000000	
D11	IDO1	22.09	0.000000	
D12	IFNG	6.71	0.000000	
E02	IL10	66.97	0.000000	
E03	IL12A	31.24	0.000000	
E04	IL12B	-4.33	0.000000	
E05	IL13	5.22	0.000000	
E06	IL15	4.21	0.000000	
E07	IL17A	10.38	0.000000	
E09	IL1B	5.72	0.000000	
E11	IL23A	10.82	0.000000	
E12	IL4	5.33	0.000000	
F03	IRF1	-2.56	0.000000	
F07	MIF	-6.84	0.000000	
F08	MYC	-2.09	0.000000	
F09	MYD88	-6.94	0.000000	

Position	Gene Symbol	Fold Regulation	p-Value	Comments
F11	NOS2	69.33	0.000000	
F12	PDCD1	2.08	0.000000	
G04	STAT3	-5.48	0.000000	
G05	TGF B1	-5.19	0.000000	
G07	TLR3	-5.41	0.000000	
G12	VEGFA	-15.08	0.000000	
H02	B2M	-3.96	0.000000	
H04	HPRT1	2.39	0.000000	
H05	RPLP0	-11.51	0.000000	

The results are also illustrated in the scatter plot. (*Figure 30*)

Test Group	Control Group	Fold Regulation Threshold
Group 1	Control Group	2

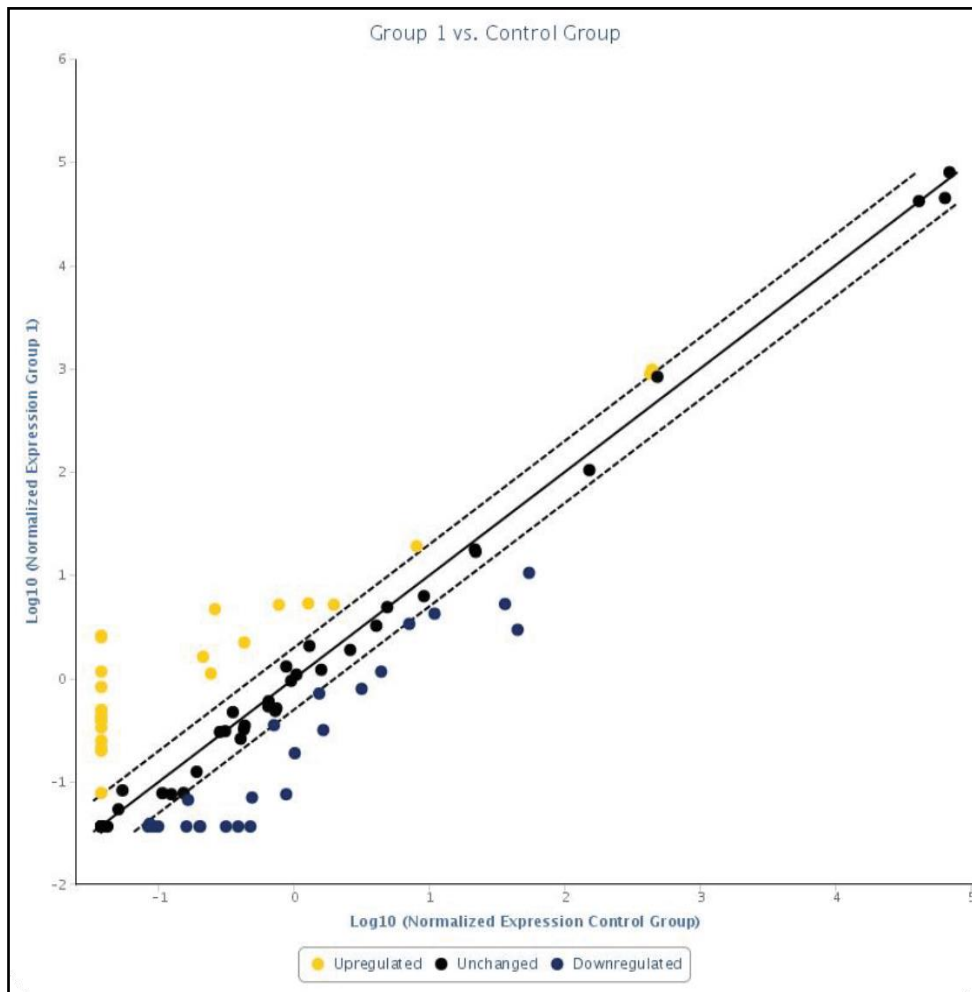


FIGURE 33: SCATTER PLOT NORMALIZED TO DEMONSTRATE FOLD DIFFERENCE BETWEEN GROUP 1 (TNBC) AGAINST CONTROL (ER POSITIVE)

TABLE 25: GENES OVEREXPRESSED IN GROUP 1 (TNBC) VS. CONTROL GROUP (ER POSITIVE)

Position	Gene Symbol	Fold Regulation
F11	NOS2	69.33
E02	IL10	66.97
E03	IL12A	31.24
D11	IDO1	22.09
H06	HGDC	18.07
C05	CXCL5	13.23
B02	CCR10	11.68
E11	IL23A	10.82
E07	IL17A	10.38
A02	AICDA	8.91
B09	CSF2	7.65
D07	HIF1A	6.71
D12	IFNG	6.71
E09	IL1B	5.72
E12	IL4	5.33
E05	IL13	5.22
A11	CCL4	4.61
E06	IL15	4.21
C04	CXCL2	2.65
H04	HPRT1	2.39
H08	RTC	2.23
H09	RTC	2.11
F12	PDCD1	2.08

TABLE 26: GENES UNDER EXPRESSED IN GROUP 1(TNBC) VS. CONTROL GROUP (ER POSITIVE)

Position	Gene Symbol	Fold Regulation
G12	VEGFA	-15.08
A12	CCL5	-12.86
H05	RPLP0	-11.51
B06	CCR9	-10.44
A10	CCL28	-8.48
F09	MYD88	-6.94
F07	MIF	-6.84
G04	STAT3	-5.48
G07	TLR3	-5.41
A04	BCL2L1	-5.33
G05	TGF B1	-5.19
C03	CXCL12	-5.15
E04	IL12B	-4.33
H02	B2M	-3.96
A03	BCL2	-3.77
A01	ACKR3	-2.68
F03	IRF1	-2.56
A06	CCL2	-2.47
D08	HLA-A	-2.45
B04	CCR4	-2.26
B05	CCR7	-2.18
C06	CXCL9	-2.14
F08	MYC	-2.09
D01	EGFR	-2.01

Volcano Plot

The volcano plot (**Figure 34**) displays both large and small gene expression changes that are statistically significant.

Test Group	Control Group	Fold Regulation Threshold	p-Value Threshold
Group 1	Control Group	2	0.05

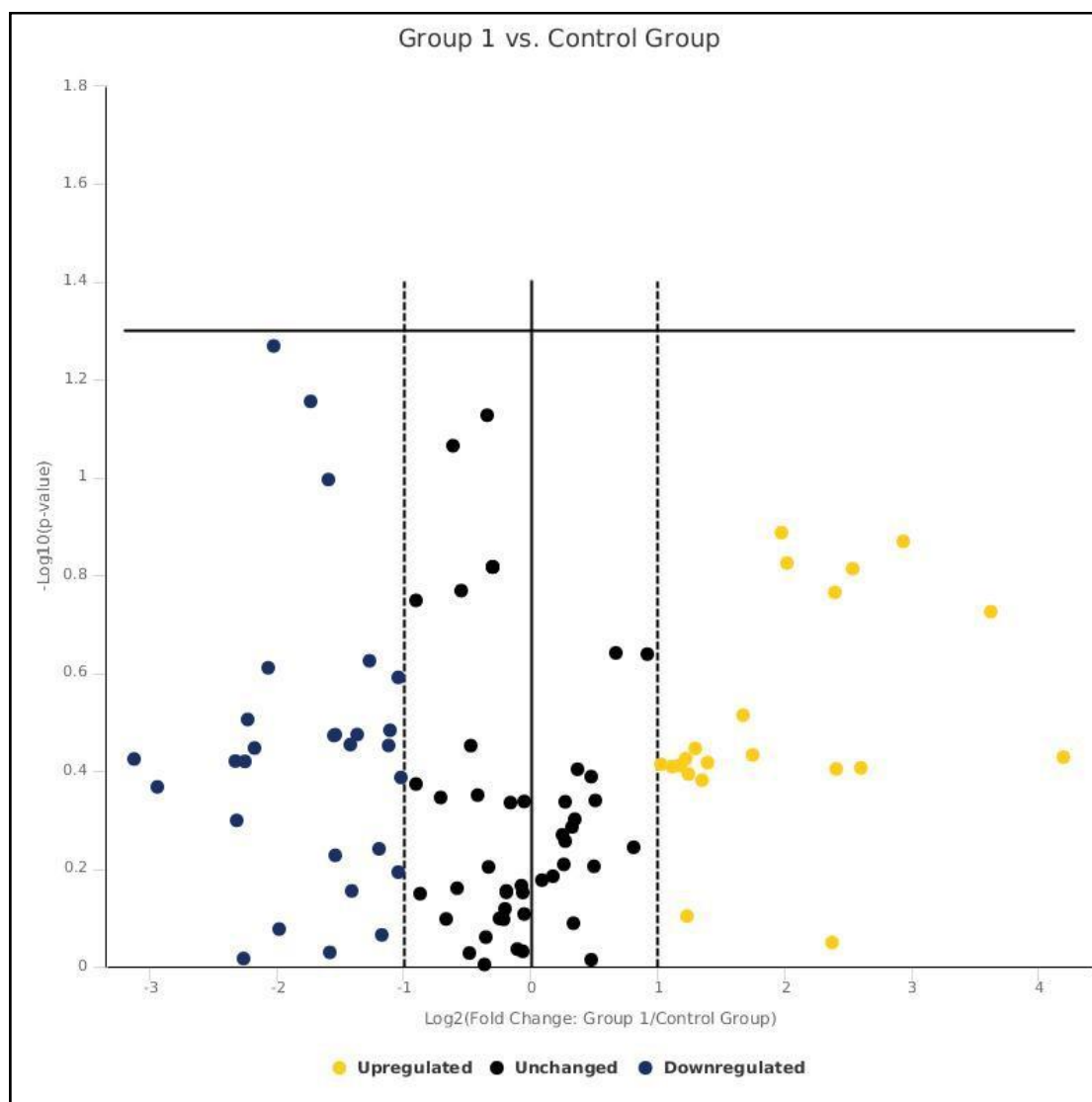


FIGURE 34: VOLCANO PLOT

The Heat Map (**Figure 35**) provides a visualization of the fold changes in expression between the selected groups for every gene in the array in the context of the array layout.

Test Group	Control Group
Group 1	Control Group

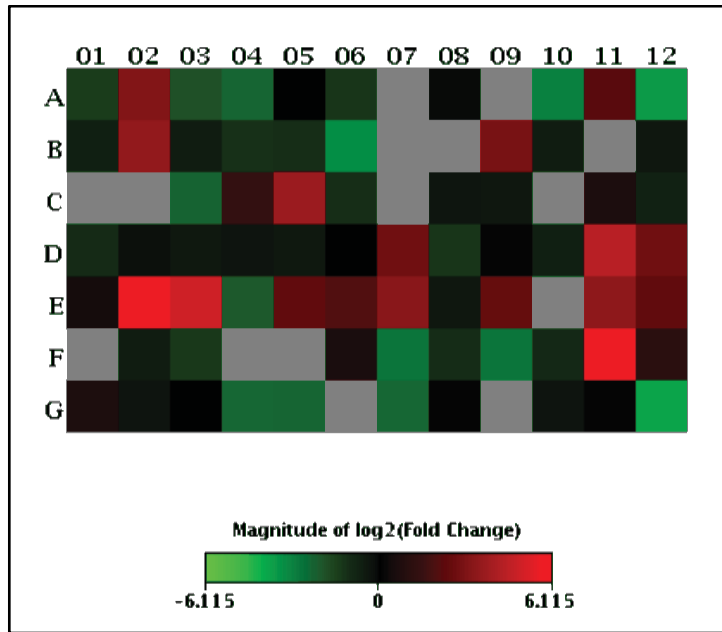


FIGURE 35: HEAT MAP

4.11.2 Results Objective 2: To compare microarray cytokine gene expression results with cytokine expression on the Enzyme Linked Immunosorbent Assay (ELISA) platform.

We found differential expression of cytokines between ER positive BC and TNBC (**Figure 36**). Prominently there was a greater upregulation of IL10 in ER positive cancer when compared with TNBC. Both MDC and IL17a were upregulated in ER positive breast cancers but downregulated in TNBC. On the other hand, MIG expression was upregulated in TNBC but downregulated in ER positive BC.

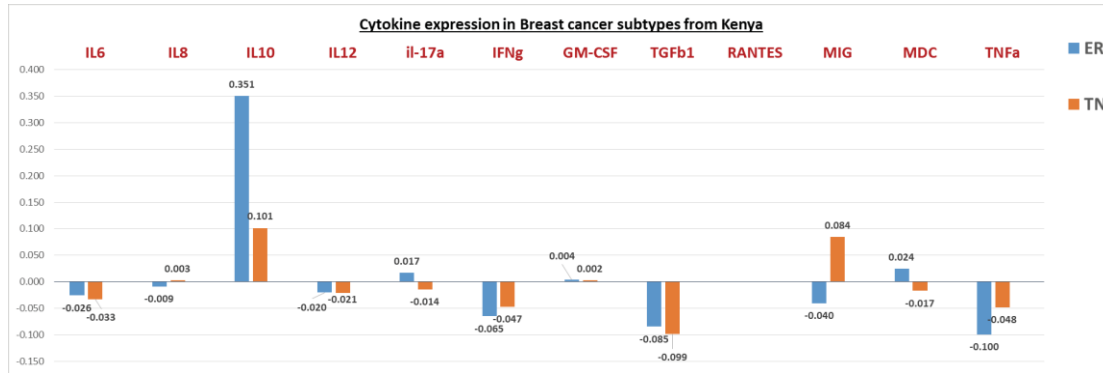


FIGURE 36: CYTOKINE EXPRESSION BETWEEN ER POSITIVE AND TRIPLE NEGATIVE BC

When we compared the ELISA results with the gene expression levels both TGF β 1 and MDC were downregulated in TNBC when compared with ER positive BC at the gene and protein level. Additionally, there was an upregulation of GM-CSF and IL8 at both the gene expression and the protein level in TNBC (**Figure 37**)

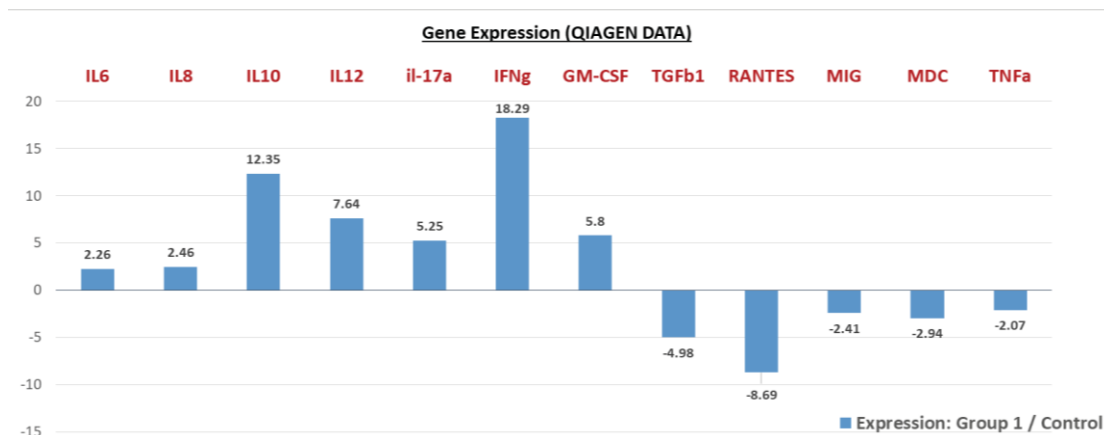


FIGURE 37: GENE EXPRESSION LEVELS (TNBC TEST GROUP/ ER POSITIVE CONTROL GROUP)

When we compared the median values of the cytokines between ER positive and TNBCs we found that TGF β 1 was downregulated more in TNBC than in ER positive BC and this was statistically significant (**Table 28**).

TABLE 27: DIFFERENCES IN MEDIAN VALUES OF CYTOKINES BETWEEN TNBC AND ER POSITIVE BREAST CANCERS

	<i>Tumor Type</i>		N	<i>TN</i>		N	P Value
	<i>ER</i>						
	Median	(IQR)		Median	(IQR)		
IL6	-.028	(-0.028, -.022)	5	-.036	(-0.038, -0.026)	7	0.122
IL8	0.000	(-0.014, 0.001)	39	0.000	(-0.006, 0.005)	31	0.263
IL10	.351	(0.35, 0.351)	3	-.024	(-0.024, 0.351)	9	0.296
IL12	-.016	(-0.03, -0.006)	39	-.015	(-0.024, -0.011)	31	0.453
Il-17a	-.005	(-0.024, 0)	39	-.015	(-0.035, -0.003)	31	0.105
IFNg	-.026	(-0.04, -0.017)	39	-.026	(-0.052, -0.015)	31	0.808
GM-CSF	-.020	(-0.026, -0.002)	35	-.023	(-0.034, 0.008)	30	0.585
TGF β1	-.093	(-0.104, -0.069)	39	-.109	(-0.117, -0.078)	31	0.031
MIG	-.042	(-0.1195, 0.039)	4	-.027	(-0.0315, -0.0245)	8	0.734
MDC	-.034	(-0.055, -0.002)	30	-.013	(-0.032, 0)	22	0.379
TNFa	-.052	(-0.168, -0.024)	9	-.031	(-0.188, 0.074)	3	0.518

4.12 DISCUSSION

Although not classically considered immunogenic, breast cancer initiation and progression are thought to be mediated through a dysregulated immune system in which cytokines play a key role in their function as mediators of intercellular communication. Immune cells and fibroblasts within the tumour environment produce a variety of cytokines that coordinate the inflammatory response resulting in anti-apoptotic and immunosuppressive factors enabling transformed cells to proliferate unchecked and facilitate invasion and metastases(303). Cytokines play a dual role in breast cancer with some promoting growth and invasion (IL-1, IL-6, IL-11, TGF β) while others (IL2, IL8, IFNs) are known to inhibit proliferation(252).

We describe the differences in expression of cytokines between Triple negative and ER positive breast cancer tissues from Kenya. We profiled genes that are mediators and effectors of the crosstalk between tumours and the immune system, and which influence the course of cancer progression and validated our results with a custom designed ELISA multi-analyte array.

We observed differential cytokine gene expression between triple negative and ER positive breast cancer tissues with enrichment of TNBC tissues for the immunosuppressive cytokine genes (NOS2, IL10, IDO, CXCL5, CSF2), pro-inflammatory cytokine genes (IL12A, IFN γ), growth factor genes (CSF2 (GM-CSF2), genes for cytokines involved in signal transduction (IL8) and transcription factors (HIF1 α). The fold change ranged from 70 for NOS2 and IL10 and approximately 10 for HIF1 α , IFN γ , CSF2 and IL17A. At the protein level we observed that both GM-CSF and IL8 were overexpressed in TNBC when compared with ER positive cancers while the other cytokines upregulated at the gene level were not significantly overexpressed at the protein level. Further work is needed to understand whether methodological issues or gene silencing through hyper methylation or other epigenetic events including translational defects could explain these findings.

Our results bear some similarity to those described by Chavey et al(304). In their study of 105 breast cancers they observed that out of a panel of 17 cytokines studied, IL-2, IL-6, IL-8, IL-10, IFN- γ , MCP-1, MIP-1 β , TNF α to be significantly overexpressed ($p < 0.001$) in ER-negative tumours (n=45) compared with ER positive BC (n=60), with the highest difference seen with IL8 which was four times the protein levels (in fg/ug) as that seen in ER positive cancers. Of note is that the study by Chavey et al was based on frozen tumour samples and cytokines were measured using a flow cytometry-based technique unlike the current study in which we used FFPE tissues, and the ELISA based platform. The role of IL8 in aggressive breast cancers is supported by another study investigating the mechanism by which Cox2 promotes invasion and metastases in breast cancer(305). It was shown that IL8 is an obligate partner for Cox 2 mediated invasiveness through the protein kinase C/IL-8/urokinase-type plasminogen activator pathway(305).

We also observed upregulation of the GM-CSF gene in TNBC tissues which was validated by the corresponding protein overexpression. The GM-CSF cytokine is secreted by immune cells and fibroblasts and plays a vital role in the differentiation of haematopoietic stem cells into macrophages or granulocyte series(306). Within the tumour microenvironment (TME) of breast cancer tissue, macrophages have been found to be the most abundant mononuclear cells(307). Known as tumour associated macrophages (TAMs), these cells are recruited from macrophages of the peripheral immune system into the TME(308).

The density of TAMs has been associated with poor prognosis in breast cancer(308). TNBCs have been shown to have a higher density of TAMs in the TME than ER positive BC(309). This ability of TNBC cells to recruit more TAMs has been attributed to an increase in the secretion of Macrophage Colony Stimulating Factor (M-CSF)(307).

Compared to ER positive breast cancers, we found that TNBC tumours showed a downregulation of the cytokine genes associated with inflammation and angiogenesis (VEGF), cytokines involved in immune and inflammatory responses (CCR9, CCL28, CXCL12), cytokines associated with signal transduction (STAT3) and Toll like receptor signalling (TLR3), and growth factors such as EGFR and TGF β 1 with a 2 to 15-fold difference between TNBC and ER positive cancers. Both TGF β 1 and MDC were downregulated in TNBC when compared with ER positive BC at both the gene and protein level.

The cytokine TGF- β is important for cellular communication and controls a wide range of activities that include cell proliferation, differentiation, and apoptosis. It has a dual role as a tumour suppressor and a tumour promoter(280). In the early stages it contributes to preventing the spread of cancer through its anti-proliferative effects and causes cell cycle arrest and has also been shown to promote the epithelial mesenchymal transition(303). In the later stages it has a counteractive effect - as tumour cells evade it and instead can proliferate further due to it(303).

Consistent with our findings, studies have shown triple-negative human breast cancer cell lines to escape the growth-inhibitory activity of TGF- β (310) thus increase our understanding of the aggressive nature of this breast cancer subtype. In contrast, a study of 180 non-metastatic TNBC showed overexpression of cytoplasmic TGF β protein on immunohistochemistry - this was significantly associated with higher tumour grade ($P < 0.001$) and more [axillary lymph node](#) metastases ($p = 0.029$) although high expression was not associated with overall survival (HR 0.747, 95% CI 0.367–1.522, $p = 0.422$)(309). The heterogeneity of the TNBC phenotype(311) and the dual role of TGF β may help to explain these conflicting findings. These results warrant further investigation into the role of TGF- β in TNBC subtypes.

Various factors could account for the differences we observed between the gene and the protein expression levels of cytokines in our BC cases. The use of FFPE tissue for cytokine analysis and the explorative use of the ELISA platform to validate our gene microarray results may have contributed to this, even though the ELISA method generally has a known high specificity and sensitivity(291).

However, it is to be noted that proteins are more stable than DNA/RNA extracted from formalin fixed tissues(312) and our protein extracts passed all quality checks. Furthermore, functional cytokine assays are superior since these permit differentiation from pathologic translational defects causing cytokine accumulation in the Golgi without its secretion (288). Also, cytokines by their nature are secreted in minute quantities, and have a very short half- life with varying kinetics depending on the cytokine gene (291). It may well be that for accurate characterization of cytokine production, multiple samples taken at different time points in animal models or cell cultures will be required(313). Furthermore, there is differential production in cytokines produced by various T cell subsets(314). As an example, naïve T- cells predominantly produce IL2 whereas memory T cells have a broader repertoire and can produce high levels of multiple cytokines(314). A further consideration is that a one-time measurement of cytokine protein concentrations may underestimate the actual level significantly due to the concurrent processes of cytokine secretion, uptake and degradation(315). Future studies using

various sample types and additional techniques such as reverse phase protein arrays, immunofluorescence studies and flow cytometry(298) may be necessary to confirm the cytokine concentrations and their differential expression in breast cancer subtypes.

The strength of this study was that we were able to demonstrate good protein yields from FFPE tissues. This is also the first study to explore the role of cytokines in BC from SSA, where antigenic exposure from various pathogens may regulate the immune response to tumours.

The limitations of our study are the small sample size and inability to quantify the cytokines since concentration for the standards used in the experiment were not provided in the kit. Also, the ELISA method is limited by the fact that individual identities of the cytokine producing cells cannot be determined.

In summary we found differential expression of inflammatory related cytokines (GM-CSF, IL8, TGF β 1 and MDC), at both the gene and the protein expression level between Triple negative and ER positive breast cancers. These differences can be targeted for therapeutic approaches as well be useful as biomarkers of diagnosis and prognosis.

Future work on the role of the tumour microenvironment including signalling molecules in breast cancer subtypes and the interaction between genetic, and geographic diversity with risk factors in influencing tumour progression and outcomes is recommended.

5.0 CONCLUSION

The immune landscape of breast cancer in patients from Sub Saharan Africa is understudied and findings are largely extrapolated from studies in Caucasians and African Americans. Furthermore, significant genetic and ancestral heterogeneity exists within Africa and across the African Diaspora in the prevalence of risk factors such as environmental/lifestyle factors including obesity, alcohol, and tobacco consumption as well as protective risk factors like reproductive history, and breastfeeding.

In our case series of 838 patients from Kenya we observed differences with regards to specific reproductive factors and linkage to molecular BC subtypes when compared with other populations from Africa. Specifically, we observed differences in the distribution of the molecular BC subtypes as defined by immunohistochemistry surrogates with proportions of 34.8% luminal A, 35.8%, luminal B, 10.7%, HER2-enriched, and 18.6%, triple negative (TN) breast cancer compared to 49% luminal A, 8% luminal B breast cancers, 28% TN and 15% HER2-enriched from a large case control study from Ghana. We also observed similar risk association with molecular subtypes in that women with higher BMI were more likely to develop luminal subtype of breast cancer while a higher parity was more likely observed with the triple negative subtype of BC. Although, our initial analysis suggested that prolonged breastfeeding duration was likely to be associated with the triple negative phenotype, this observation became insignificant when “Cumulative breastfeeding duration” and “Number of children” were analysed together.

An interesting finding was the suggestive association between alcohol consumption and the HER2 breast cancer subtype which warrants future studies. However, reproductive risk factors that included age at menarche, age at onset of menopause and first full-term pregnancy were not different among the molecular subtypes.

Further morphologic and immunohistochemical analysis of Tissue Microarrays from 226 definitive surgical specimens selected from the larger cohort of 838 cases, showed enrichment of Tumour Infiltrating Lymphocytes (TILs) in more aggressive breast cancers similar to what has been previously published. Increasing Tumour Infiltrating Lymphocytes were associated with high KI67 (>20%), high grade (Grade 3 BC) and HER2 status (positive HER2), luminal B subtype, and inversely associated with tumour size. The composition of the TILs was predominantly CD3 T cells followed by CD8 T cells and CD68 macrophages with a much smaller contribution of

CD4 T cells and CD20 B cells and FOxP3 T regulatory cells. Although we observed an unexpected association of CD3 T cells with worse breast cancer outcomes, this finding will need to be confirmed in a larger study. Overall, we did not observe any distinct patterns of TIL profiles in relation to clinical characteristics that are specific to this Kenyan population. However, we did observe higher TIL values for Leading edge TILs (LE TILs) compared to intra tumour stromal TILs (sTILs). We did not notice any pattern with regards to location of TILs between sTIL and LE TIL or combined average TIL in association with risk factors and tumour features. However, we did note that for some specific associations between a marker and risk factor/tumour feature, the sTIL component drove the significant differences; whereas in some immune cell parameters, it was the LE TIL part that contributed more. Further studies with a larger sample size are warranted to assess the clinical utility of assessing the stromal compartments separately.

To further understand the role of signalling molecules in molecular subtypes of breast cancer, we profiled cytokines both at the gene and protein expression and compared the expression in

ER positive and Triple negative breast cancers. We found that compared to ER positive BC, there was enrichment of TNBC tissues for the immunosuppressive cytokine genes (NOS2, IL10, IDO, CXCL5, CSF2), pro-inflammatory cytokine genes (IL12A, IFN γ), growth factor genes (CSF2, GM-CSF2), genes for cytokines involved in signal transduction (IL8) and transcription factors (HIF1 α). At the protein level we observed that both GM-CSF and IL8 were overexpressed in TNBC when compared with ER positive cancers while the other cytokines upregulated at the gene level were not significantly overexpressed at the protein level. As to whether this discrepancy between gene and protein expression can be explained by our methodological approach or other epigenetic events will require further exploration.

Our findings from this study highlight the similarities and differences in BC risk factors, the tumour immune microenvironment and molecular subtypes of BC in Kenyan patients and adds to the limited body of knowledge in this area within the African continent. Differential expression of signalling molecules across the molecular subtype of BC further underscores the potential role of these molecules as biomarkers of diagnosis, prognosis, and therapeutic approaches.

Large scale studies of the tumour microenvironment in BC subtypes among various racial and ethnic populations will supplement our findings and provide further insight into the contribution of the immune milieu in heterogeneity of breast cancer subtype distribution, prognosis, response to therapy and outcomes.

6.0 RECOMMENDATIONS FOR FUTURE WORK

There is clear evidence that non-biological factors such as socioeconomic status, access to health care, living environments, and nutrition contribute to racial disparity in terms of both early diagnosis of breast cancer as well as stage matched outcomes in terms of recurrence and overall survival. However, even after controlling for these factors, overall phenotypic differences in breast cancer pathology remain among different ethnic groups. In addition, pathways that are enriched for signals of adaptation tend to be involved in immune response and wound healing, which may partially explain the high prevalence in women of African ancestry of TNBC in some regions of Africa, suggesting that African breast cancer may have distinct immune profiles. Furthermore, Africans are a heterogeneous group; there are great variations in genetic structure and exposures across different African populations. Despite this, the distribution of molecular subtypes and immune phenotypes in African breast cancer remains largely unknown.

Future work on the role of the tumour microenvironment including signalling molecules in breast cancer subtypes and the interaction between genetic and geographic diversity with risk factors in influencing tumour progression and outcomes in a broader African population is warranted. Furthermore, studies to understand how TIL quantity, TIL subtype and their spatial distribution contributes to breast cancer prognosis are recommended. The clinical significance of stromal cellular density (SCD) of TILs using machine learning algorithms compared to the visual TIL estimation methods require further investigation. Special techniques such as multi-spectral immunofluorescence (qMIF) to define the spatial localization of TIL subtypes in relation to tumour cells and the clinical relevance of the co-localisation of TIL immune cells may help to inform breast cancer prognosis and outcomes.

Studies of gene expression profiling analysis of breast tumours for characterization of molecularly defined subtypes including intrinsic, immune and tumour microenvironment (TME) subtypes in breast cancers diagnosed in Kenya and comparing to those in diverse populations will provide further insight into the role of the TME in BC subtype heterogeneity. Further, the assessment of molecular and genomic features of breast tumours and the TME and its associations with breast cancer risk factors and clinical outcomes (including survival) may help provide a clear understanding of the relationship between tumour biology, host genetics and environmental modifications of established breast cancer risks. The interactions between the immune cells in the TME and the pro-inflammatory cytokines in breast cancer subtypes warrants further studies as potential targets for therapeutic approaches and biomarkers of diagnosis and prognosis.

REFERENCES

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*. 2021;71(3):209-49.
2. Korir A, Okerosi N, Ronoh V, Mutuma G, Parkin M. Incidence of cancer in Nairobi , Kenya (2004 – 2008). *International Journal of Cancer*. 2015;2059(137):2053-9.
3. Sayed S, Moloo Z, Ngugi A, Allidina A, Ndumia R, Mutui A, et al. Breast Camps for Awareness and Early Diagnosis of Breast Cancer in Countries With Limited Resources: A Multidisciplinary Model From Kenya. *Oncologist*. 2016;21(9):1138-48.
4. Chlebowski RT, Chen Z, Anderson GL, Rohan T, Aragaki A, Lane D, et al. Ethnicity and breast cancer: factors influencing differences in incidence and outcome. *J Natl Cancer Inst*. 2005;97(6):439-48.
5. Knaul FM, Bhadelia A, Gralow J, Arreola-Ornelas H, Langer A, Frenk J. Meeting the emerging challenge of breast and cervical cancer in low- and middle-income countries. *International journal of gynaecology and obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics*. 2012;119 Suppl 1:S85-8.
6. Lu H, Ouyang W, Huang C. Inflammation, a key event in cancer development. *Molecular cancer research : MCR*. 2006;4:221-33.
7. Mittal D, Gubin MM, Schreiber RD, Smyth MJ. New insights into cancer immunoediting and its three component phases--elimination, equilibrium and escape. *Curr Opin Immunol*. 2014;27:16-25.
8. Demaria S, Pikarsky E, Karin M, Coussens LM, Chen Y-C, El-Omar EM, et al. Cancer and inflammation: promise for biologic therapy 2010. 335-51 p.
9. Korkaya H, Liu S, Wicha MS. Review series Breast cancer stem cells , cytokine networks , and the tumor microenvironment. *The Journal of clinical investigation*. 2011;121(10):3804-9.
10. Coussens LM, Pollard JW. Leukocytes in mammary development and cancer. *Cold Spring Harbor perspectives in biology*. 2011;3(3):1-22.
11. Pierce BL, Ballard-Barbash R, Bernstein L, Baumgartner RN, Neuhaus ML, Wener MH, et al. Elevated biomarkers of inflammation are associated with reduced survival among breast cancer patients. *Journal of Clinical Oncology*. 2009;27(21):3437-44.
12. Coussens LM, Werb Z. Inflammation and cancer. *Nature*. 2002;420(6917):860-7.
13. Pandith AA, Bhat I, Mansoor S, Koul A, Manzoor U, Anwar I, et al. Cytokine Gene Polymorphism and Cancer Risk: A Promising Tool for Individual Susceptibility and Prognostic Implications. *Genetic Polymorphisms-New Insights: IntechOpen*; 2021.
14. Giordani L, Bruzzi P, Lasalandra C, Quaranta M, Schittulli F, Della Ragione F, et al. Association of breast cancer and polymorphisms of interleukin-10 and tumor necrosis factor- α genes. *Clinical chemistry*. 2003;49(10):1664-7.
15. Karakus N, Kara N, Ulusoy AN, Özaslan C, Bek Y. Tumor necrosis factor alpha and beta and interferon gamma gene polymorphisms in Turkish breast cancer patients. *DNA and Cell Biology*. 2011;30(6):371-7.
16. Tie Y, Tang F, Wei Y-q, Wei X-w. Immunosuppressive cells in cancer: mechanisms and potential therapeutic targets. *Journal of Hematology & Oncology*. 2022;15(1):1-33.
17. Gulić T, Grebić D, Starčević A, Avirović M, Valković Zujić P, Veljković Vujaklija D, et al. Clinical relevance of natural killer cells in breast cancer. *Medicina Fluminensis*. 2022;58(2):113-20.
18. Hanamura T, Kitano S, Kagamu H, Yamashita M, Terao M, Tsuda B, et al. Immunological profiles of the breast cancer microenvironment represented by tumor-infiltrating lymphocytes and PD-L1 expression. *Scientific Reports*. 2022;12(1):1-10.

19. Zheng H, Tan Z, Kan C, Minqiong S, Yang F, Wong M. Mapping breast cancer microenvironment through single cell omics. *Frontiers in Immunology*. 2022;1439.
20. Franzén AS, Raftery MJ, Pecher G. Implications for Immunotherapy of Breast Cancer by Understanding the Microenvironment of a Solid Tumor. *Cancers*. 2022;14(13):3178.
21. Salgado R, Denkert C, Demaria S, Sirtaine N, Klauschen F, Pruneri G, et al. The evaluation of tumor-infiltrating lymphocytes (TILs) in breast cancer: recommendations by an International TILs Working Group 2014. *Ann Oncol*. 2015;26(2):259-71.
22. Loi S, Michiels S, Adams S, Loibl S, Budczies J, Denkert C, et al. The journey of tumor-infiltrating lymphocytes as a biomarker in breast cancer: clinical utility in an era of checkpoint inhibition. *Ann Oncol*. 2021;32(10):1236-44.
23. Savas P, Salgado R, Denkert C, Sotiriou C, Darcy PK, Smyth MJ, et al. Clinical relevance of host immunity in breast cancer: from TILs to the clinic. *Nat Rev Clin Oncol*. 2016;13(4):228-41.
24. Luen S, Virassamy B, Savas P, Salgado R, Loi S. The genomic landscape of breast cancer and its interaction with host immunity. *Breast*. 2016;29:241-50.
25. Der EM, Gyasi RK, Tettey Y, Edusei L, Bayor MT, Jiagge E, et al. Triple-Negative Breast Cancer in Ghanaian Women: The Korle Bu Teaching Hospital Experience. 2015;21(6):627-33.
26. Rambau PF, Chalya PL, Manyama MM, Jackson KJ. Pathological features of Breast Cancer seen in Northwestern Tanzania: a nine years retrospective study. *BMC research notes*. 2011;4(1):1-6.
27. Abdulrahman GO, Rahman GA. Epidemiology of breast cancer in Europe and Africa. *Journal of cancer epidemiology*. 2012;2012.
28. Osei-Twum J-A, Gedleh S, Lofters A, Nnorom O. Differences in Breast Cancer Presentation at Time of Diagnosis for Black and White Women in High Resource Settings. *Journal of Immigrant and Minority Health*. 2021;23(6):1305-42.
29. Jedy-Agba E, McCormack V, Adebamowo C, dos-Santos-Silva I. Stage at diagnosis of breast cancer in sub-Saharan Africa: a systematic review and meta-analysis. *The Lancet Global Health*. 2016;4(12):e923-e35.
30. Yao S, Cheng T-YD, Elkhanany A, Yan L, Omilian A, Abrams SI, et al. Breast tumor microenvironment in black women: a distinct signature of CD8+ T-cell exhaustion. *JNCI: Journal of the National Cancer Institute*. 2021;113(8):1036-43.
31. Lee MC, Newman LA. Management of Patients with Locally Advanced Breast Cancer. *Surgical Clinics of North America*. 2007;87(2):379-98.
32. Galon J, Franck P, Marincola FM, Angell HK, Thurin M, Lugli A, et al. Cancer classification using the Immunoscore: a worldwide task force. *Journal of Translational Medicine*. 2012;10(1):205-.
33. Galon J, Costes A, Sanchez-Cabo F, Kirilovsky A, Mlecnik B, Lagorce-Pagès C, et al. Type, Density, and Location of Immune Cells Within Human Colorectal Tumors Predict Clinical Outcome. *Science*. 2006;313:1960-4.
34. Pagès F, Kirilovsky A, Mlecnik B, Asslaber M, Tosolini M, Bindea G, et al. In situ cytotoxic and memory T cells predict outcome in patients with early-stage colorectal cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2009;27(35):5944-51.
35. Messerschmidt JL, Prendergast GC, Messerschmidt GL. How Cancers Escape Immune Destruction and Mechanisms of Action for the New Significantly Active Immune Therapies: Helping Nonimmunologists Decipher Recent Advances. *The oncologist*. 2016;21(2):233-43.
36. Cyprian Fs SAZGSV. Targeted immunotherapy with a checkpoint inhibitor in combination with chemotherapy: A new clinical paradigm in the treatment of triple-negative breast cancer. *Bosnian Journal of Basic Medical Sciences*. 2019.

37. Schmid P, Cortés J, Dent R, Pusztai L, McArthur H, Kuemmel S, et al. KEYNOTE-522: phase 3 study of pembrolizumab+ chemotherapy vs placebo+ chemotherapy as neoadjuvant treatment, followed by pembrolizumab vs placebo as adjuvant treatment for early-stage high-risk triple-negative breast cancer. *Ann Oncol*. 2019;30.
38. Szymiczek A, Lone A, Akbari MR. Molecular intrinsic versus clinical subtyping in breast cancer: A comprehensive review. *Clinical genetics*. 2021;99(5):613-37.
39. Yang XR, Sherman ME, Rimm DL, Lissowska J, Brinton LA, Peplonska B, et al. Differences in risk factors for breast cancer molecular subtypes in a population-based study. *Cancer Epidemiology Biomarkers and Prevention*. 2007;16(3):439-43.
40. Dietze EC, Chavez TA, Seewaldt VL. Obesity and triple-negative breast cancer: disparities, controversies, and biology. *The American journal of pathology*. 2018;188(2):280-90.
41. Figueroa JD, Davis Lynn BC, Edusei L, Titiloye N, Adjei E, Clegg-Lampsey JN, et al. Reproductive factors and risk of breast cancer by tumor subtypes among Ghanaian women: A population-based case-control study. *Int J Cancer*. 2020;147(6):1535-47.
42. Sayed S, Fan S, Moloo Z, Wasike R, Bird P, Saleh M, et al. Breast cancer risk factors in relation to molecular subtypes in breast cancer patients from Kenya. *Breast Cancer Res*. 2021;23(1):68.
43. Sayed S, Moloo Z, Wasike R, Bird P, Oigara R, Govender D, et al. Is breast cancer from Sub Saharan Africa truly receptor poor? Prevalence of ER/PR/HER2 in breast cancer from Kenya. *The Breast*. 2014;23(5):591-6.
44. Report MW. National Epilepsy Awareness Month — November 2012 Epilepsy in Adults and Access to Care — United States , 2010. 2012;61(45):2005-9.
45. Parise CA, Bauer KR, Caggiano V. Variation in breast cancer subtypes with age and race/ethnicity. *Critical reviews in oncology/hematology*. 2010;76(1):44-52.
46. Forouzanfar MH, Foreman KJ, Delossantos AM, Lozano R, Lopez AD, Murray CJ, et al. Breast and cervical cancer in 187 countries between 1980 and 2010: a systematic analysis. *Lancet (London, England)*. 2011;378(9801):1461-84.
47. Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin D, Piñeros M, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *International journal of cancer*. 2019;144(8):1941-53.
48. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. 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):394-424.
49. Lund MJ, Trivers KF, Porter PL, Coates RJ, Leyland-Jones B, Brawley OW, et al. Race and triple negative threats to breast cancer survival: a population-based study in Atlanta, GA. *Breast Cancer Res Treat*. 2009;113(2):357-70.
50. Bauer KR, Brown M, Cress RD, Parise CA, Caggiano V. Descriptive analysis of estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and HER2-negative invasive breast cancer, the so-called triple-negative phenotype: a population-based study from the California cancer Registry. *Cancer*. 2007;109(9):1721-8.
51. Benefield HC, Allott EH, Reeder-Hayes KE, Perou CM, Carey LA, Geradts J, et al. Borderline Estrogen Receptor-Positive Breast Cancers in Black and White Women. *JNCI: Journal of the National Cancer Institute*. 2020;112(7):728-36.
52. Pitt JJ, Riester M, Zheng Y, Yoshimatsu TF, Sanni A, Oluwasola O, et al. Characterization of Nigerian breast cancer reveals prevalent homologous recombination deficiency and aggressive molecular features. *Nature Communications*. 2018;9(1):4181-.
53. Byun JS, Singhal SK, Park S, Yi DI, Yan T, Caban A, et al. Racial differences in the association between luminal master regulator gene expression levels and breast cancer survival. *Clinical Cancer Research*. 2020;26(8):1905-14.

54. Gong Z, Wang J, Wang D, Buas MF, Ren X, Freudenheim JL, et al. Differences in microRNA expression in breast cancer between women of African and European ancestry. *Carcinogenesis*. 2019;40(1):61-9.
55. Dietze EC, Sistrunk C, Miranda-Carboni G, O'regan R, Seewaldt VL. Triple-negative breast cancer in African-American women: disparities versus biology. *Nature Reviews Cancer*. 2015;15(4):248-54.
56. Lara OD, Wang Y, Asare A, Xu T, Chiu HS, Liu Y, et al. Pan-cancer clinical and molecular analysis of racial disparities. *Cancer*. 2020;126(4):800-7.
57. Palmer JR, Ruiz-Narvaez EA, Rotimi CN, Cupples LA, Cozier YC, Adams-Campbell LL, et al. Genetic susceptibility loci for subtypes of breast cancer in an African American population. *Cancer Epidemiology and Prevention Biomarkers*. 2013;22(1):127-34.
58. Julie R. Palmer DABLAWCBALL, Adams C, Lynn R. Parity and lactation in relation to estrogen receptor negative breast cancer in African American women. *Cancer Epidemiol Biomarkers Prev* 2011 September ; 20(9): 1883–1891. 2011;20(9):1883-91.
59. Kwan ML, Kushi LH, Weltzien E, Maring B, Kutner SE, Fulton RS, et al. Epidemiology of breast cancer subtypes in two prospective cohort studies of breast cancer survivors. *Breast Cancer Research*. 2009;11(3):1-13.
60. Yang XR, Figueroa JD, Falk RT, Zhang H, Pfeiffer RM, Hewitt SM, et al. Analysis of terminal duct lobular unit involution in luminal A and basal breast cancers. *Breast Cancer Research*. 2012.
61. Smith RA, Brooks D, Cokkinides V, Saslow D, Brawley OW. Cancer Screening in the United States , A Review of Current American Cancer Society Guidelines , Current Issues in Cancer Screening and New Guidance on Cervical Cancer Screening and Lung Cancer Screening. CA: A Cancer Journal for Clinicians. 2013;63(2):87-105.
62. Ohio Department of H. Non-Modifiable, Modifiable and Environmental Risk Factors for Cancer. 2015(April):1-9.
63. Bertwistle D, Ashworth A. The pathology of familial breast cancer. How do the functions of BRCA1 and BRCA2 relate to breast tumour pathology? *Breast Cancer Research*. 1999;1(1):41-7.
64. Kuchenbaecker K, Hopper J, Barnes D. Phillips. KA, Mooij. TM, and Roos-Blom. M.(2017). Risks of Breast, Ovarian, and Contralateral Breast Cancer for BRCA1 and BRCA2 Mutation Carriers *Jama*.317(23):2402-16.
65. Hall MJ, Reid JE, Burbidge LA, Pruss D, Deffenbaugh AM, Frye C, et al. BRCA1 and BRCA2 mutations in women of different ethnicities undergoing testing for hereditary breast-ovarian cancer. *Cancer*. 2009;115(10):2222-33.
66. Kurian AW. BRCA1 and BRCA2 mutations across race and ethnicity: distribution and clinical implications. *Current Opinion in Obstetrics and Gynecology*. 2010;22(1):72-8.
67. Rebbeck TR, Friebel TM, Friedman E, Hamann U, Huo D, Kwong A, et al. Mutational spectrum in a worldwide study of 29,700 families with BRCA1 or BRCA2 mutations. *Human mutation*. 2018;39(5):593-620.
68. Nelson HD, Fu R, Goddard K, Mitchell JP, Okinaka-Hu L, Pappas M, et al. Risk assessment, genetic counseling, and genetic testing for BRCA-related cancer: systematic review to update the US Preventive Services Task Force Recommendation [Internet]. 2013.
69. Jones ME, Schoemaker MJ, Wright LB, Ashworth A, Swerdlow AJ. Smoking and risk of breast cancer in the Generations Study cohort. *Breast Cancer Research*. 2017;19(1):118-.
70. Bisgin A, Sag SO, Dogan ME, Yildirim MS, Gumus AA, Akkus N, et al. Germline landscape of BRCA1 by 7-site collaborations as a BRCA consortium in Turkey. *The Breast*. 2022;65:15-22.

71. Van der Merwe NC, Combrink HM, Ntaita KS, Oosthuizen J. Prevalence of Clinically Relevant Germline BRCA Variants in a Large Unselected South African Breast and Ovarian Cancer Cohort: A Public Sector Experience. *Frontiers in genetics*. 2022;13:834265.
72. Mighri N, Hamdi Y, Boujemaa M, Othman H, Ben Nasr S, El Benna H, et al. Identification of novel BRCA1 and RAD50 mutations associated with breast cancer predisposition in Tunisian patients. *Frontiers in genetics*. 2020;13:97.
73. Ndiaye R, Diop JPD, Bourdon-Huguenin V, Dem A, Diouf D, Dieng MM, et al. Evidence for an ancient BRCA1 pathogenic variant in inherited breast cancer patients from Senegal. *NPJ genomic medicine*. 2020;5(1):1-6.
74. Martins FC, De S, Almendro V, Gönen M, Park SY, Blum JL, et al. Evolutionary Pathways in BRCA1-Associated Breast Tumors. *BRCA1-Associated Breast Tumorigenesis. Cancer discovery*. 2012;2(6):503-11.
75. Romagnolo A, Romagnolo D, Selmin O. BRCA1 as Target for Breast Cancer Prevention and Therapy. *Anti-Cancer Agents in Medicinal Chemistry*. 2014;15(1):4-14.
76. Birgisdottir V, Stefansson OA, Bodvarsdottir SK, Hilmarsdottir H, Jonasson JG, Eyfjord JE. Epigenetic silencing and deletion of the BRCA1 gene in sporadic breast cancer. *Breast cancer research*. 2006;8(4):1-10.
77. Ha SM, Chae EY, Cha JH, Kim HH, Shin HJ, Choi WJ. Association of BRCA mutation types, imaging features, and pathologic findings in patients with breast cancer with BRCA1 and BRCA2 Mutations. *American Journal of Roentgenology*. 2017;209(4):920-8.
78. Holm J, Eriksson L, Ploner A, Eriksson M, Rantalainen M, Li J, et al. Assessment of Breast Cancer Risk Factors Reveals Subtype Heterogeneity. *Cancer Res*. 2017;77(13):3708-17.
79. Angeli D, Salvi S, Tedaldi G. Genetic predisposition to breast and ovarian cancers: how many and which genes to test? *International journal of molecular sciences*. 2020;21(3):1128.
80. Sun Y-SS, Zhao Z, Yang Z-NN, Xu F, Lu H-JJ, Zhu Z-YY, et al. Risk factors and preventions of breast cancer. *International Journal of Biological Sciences*. 2017;13(11):1387-97.
81. SEER Cancer Statistics Review 1975-2006 National Cancer Institute SEER Cancer Statistics Review 1975-2006 National Cancer Institute. *Cancer*. 2008;1975-2006.
82. Shumway DA, Sabolch A, Jagsi R. Breast Cancer: Facts and figures 2019-2020. *medical radiology*. 2020;1-43.
83. Sayed S, Moloo Z, Wasike R, Bird P, Oigara R, Njoroge FW, et al. Ethnicity and breast cancer characteristics in Kenya. *Breast cancer research and treatment*. 2018;167(2):425-37.
84. Benz CC. Impact of aging on the biology of breast cancer. *Critical reviews in oncology/hematology*. 2008;66(1):65-74.
85. Radisky DC, Visscher DW, Frank RD, Vierkant RA, Winham S, Stallings-Mann M, et al. Natural history of age-related lobular involution and impact on breast cancer risk. *Breast Cancer Res Treat*. 2016;155(3):423-30.
86. Brewer HR, Jones ME, Schoemaker MJ, Ashworth A, Swerdlow AJ. Family history and risk of breast cancer: an analysis accounting for family structure. *Breast Cancer Research and Treatment*. 2017;165(1):193-200.
87. Nindrea RD, Aryandono T, Lazuardi L. Breast Cancer Risk From Modifiable and Non-Modifiable Risk Factors among Women in Southeast Asia: A Meta-Analysis. *Asian Pacific journal of cancer prevention : APJCP*. 2017;18(12):3201-6.
88. Tan MM, Ho WK, Yoon SY, Mariapun S, Hasan SN, Lee DS, et al. A case-control study of breast cancer risk factors in 7,663 women in Malaysia. *PLoS One*. 2018;13(9):e0203469.
89. Britt K, Short R. The plight of nuns: hazards of nulliparity. *The Lancet*. 2012;379(9834):2322-3.

90. Ghoncheh M, Pournamdar Z, Salehiniya H. Incidence and Mortality and Epidemiology of Breast Cancer in the World. *Asian Pacific journal of cancer prevention : APJCP*. 2016;17(S3):43-6.
91. Badr LK, Bourdeanu L, Alatrash M, Bekarian G. Breast Cancer Risk Factors: a Cross-Cultural Comparison between the West and the East. *Asian Pacific journal of cancer prevention : APJCP*. 2018;19(8):2109-16.
92. Subramani R, Nandy SB, Pedroza DA, Lakshmanaswamy R. Role of growth hormone in breast cancer. *Endocrinology*. 2017;158(6):1543-55.
93. Endogenous H, Breast Cancer Collaborative G. Sex hormones and breast cancer risk in premenopausal women: collaborative reanalysis of seven prospective studies. *Lancet Oncology*. 2013;14(10):1009-19.
94. Iyengar NM, Arthur R, Manson JE, Chlebowski RT, Kroenke CH, Peterson L, et al. Association of Body Fat and Risk of Breast Cancer in Postmenopausal Women With Normal Body Mass Index: A Secondary Analysis of a Randomized Clinical Trial and Observational Study. *JAMA oncology*. 2019;5(2):155-63.
95. van den Brandt PA, Spiegelman D, Yaun SS, Adami HO, Beeson L, Folsom AR, et al. Pooled analysis of prospective cohort studies on height, weight, and breast cancer risk. *Am J Epidemiol*. 2000;152(6):514-27.
96. Neilson HK, Farris MS, Stone CR, Vaska MM, Brenner DR, Friedenreich CM. Moderate-vigorous recreational physical activity and breast cancer risk, stratified by menopause status: A systematic review and meta-analysis. *Menopause*. 2016;24(3):322-44.
97. Niehoff NM, Nichols HB, Zhao S, White AJ, Sandler DP. Adult Physical Activity and Breast Cancer Risk in Women with a Family History of Breast Cancer. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 2019;28(1):51-8.
98. Petrakis NL, Gruenke LD, Beelen TC, Castagnoli N, Craig JC. Nicotine in breast fluid of nonlactating women. *Science*. 1978;199(4326):303-5.
99. Weigelt B, Geyer FC, Reis-Filho JS. Histological types of breast cancer: how special are they? *Molecular oncology*. 2010;4(3):192-208.
100. Burstein HJ, Polyak K. Ductal carcinoma in situ of the breast. *New England Journal* 2004;350:1430-41.
101. Tan PH, Ellis I, Allison K, Brogi E, Fox SB, Lakhani S, et al. The 2019 World Health Organization classification of tumours of the breast. *Histopathology*. 2020;77(2):181-5.
102. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology*. 1991;19(5):403-10.
103. Schnitt SJ. Classification and prognosis of invasive breast cancer: from morphology to molecular taxonomy. *Modern Pathology*. 2010;23(S2):S60-S4.
104. Goldhirsch A, Ingle JN, Gelber RD, Coates AS, Thurlimann B, Senn HJ. Thresholds for therapies: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2009. *Annals of Oncology*. 2009;20(8):1319-29.
105. Eifel P, Axelson JA, Costa J, Crowley J, Curran WJ, Jr., Deshler A, et al. National Institutes of Health Consensus Development Conference Statement: adjuvant therapy for breast cancer, November 1-3, 2000. *J Natl Cancer Inst*. 2001;93(13):979-89.
106. Osborne CK. Tamoxifen in the Treatment of Breast Cancer. *Drug Therapy*. 2001;344(4):276-85.
107. Perou CM, Sørli T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. *Nature*. 2000;406(6797):747-52.
108. Sørli T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications.

- Proceedings of the National Academy of Sciences of the United States of America. 2001;98(19):10869-74.
109. Prat A, Parker JS, Karginova O, Fan C, Livasy C, Herschkowitz JI, et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res.* 2010;12(5):R68.
 110. Herschkowitz JI, Simin K, Weigman VJ, Mikaelian I, Usary J, Hu Z, et al. Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome biology.* 2007;8(5):R76-R.
 111. Prat A, Perou CM. Deconstructing the molecular portraits of breast cancer. *Mol Oncol.* 2011;5(1):5-23.
 112. Hennessy BT, Gonzalez-Angulo AM, Stemke-Hale K, Gilcrease MZ, Krishnamurthy S, Lee JS, et al. Characterization of a naturally occurring breast cancer subset enriched in epithelial-to-mesenchymal transition and stem cell characteristics. *Cancer Res.* 2009;69(10):4116-24.
 113. Cheang MC, Voduc D, Bajdik C, Leung S, McKinney S, Chia SK, et al. Basal-like breast cancer defined by five biomarkers has superior prognostic value than triple-negative phenotype. *Clin Cancer Res.* 2008;14(5):1368-76.
 114. Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thürlimann B, Senn HJ. Strategies for subtypes-dealing with the diversity of breast cancer: Highlights of the St Gallen international expert consensus on the primary therapy of early breast cancer 2011. *Annals of Oncology.* 2011;22(8):1736-47.
 115. Maisonneuve P, Disalvatore D, Rotmensz N, Curigliano G, Colleoni M, Dellapasqua S, et al. Proposed new clinicopathological surrogate definitions of luminal A and luminal B (HER2-negative) intrinsic breast cancer subtypes. *Breast Cancer Research.* 2014;16(3):R65-R.
 116. Sayed S, Moloo Z, Wasike R, Bird P, Oigara R, Njoroge FW, et al. Ethnicity and breast cancer characteristics in Kenya. *Breast Cancer Res Treat.* 2018;167(2):425-37.
 117. Harbeck N, Thomssen C, Gnant M. St. Gallen 2013: Brief preliminary summary of the consensus discussion. *Breast Care.* 2013;8(2):102-9.
 118. Rakha E, Ellis IO, Histologic Grade, in O'Malley, FP, et al, . *Breast Cancer Grading.* breast Pathology 2011.
 119. Lokuhetty D, White V, Watanabe R, Cree I. WHO Classification of Tumours Editorial Board. *Breast Tumours. WHO Classification of Tumours Series, 5th ed; International Agency for Research on Cancer: Lyon, France.* 2019;2:88-97.
 120. Cree IA, Tan PH, Travis WD, Wesseling P, Yagi Y, White VA, et al. Counting mitoses: SI (ze) matters! *Modern Pathology.* 2021;34(9):1651-7.
 121. Pereira L, Mutesa L, Tindana P, Ramsay M. African genetic diversity and adaptation inform a precision medicine agenda. *Nature Reviews Genetics.* 2021;22(5):284-306.
 122. Mremi A, Broadwater G, Jackson K, Amsi P, Mbulwa C, Hyslop T, et al. Breast cancer in Tanzanian, black American, and white American women: An assessment of prognostic and predictive features, including tumor infiltrating lymphocytes. *PLOS ONE.* 2019;14(11):e0224760-e.
 123. Traoré B, Koulibaly M, Diallo A, Bah M. Molecular profile of breast cancers in Guinean oncological settings. *The Pan African medical journal.* 2019;33:22.
 124. Rayne S, Schnippel K, Grover Md S, Fearnhead K, Kruger D, Benn C, et al. Unraveling the South African Breast Cancer Story: The Relationship of Patients, Delay to Diagnosis, and Tumor Biology With Stage at Presentation in an Urban Setting. *Journal of Surgical Research.* 2019;235:181-9.
 125. Palmer JR, Viscidi E, Troester MA, Hong CC, Schedin P, Bethea TN, et al. Parity, lactation, and breast cancer subtypes in African American Women: Results from the AMBER Consortium. *Journal of the National Cancer Institute.* 2014;106(10).

126. Abubakar M, Sung H, Bcr D, Guida J, Tang TS, Pfeiffer RM, et al. Breast cancer risk factors, survival and recurrence, and tumor molecular subtype: analysis of 3012 women from an indigenous Asian population. *Breast Cancer Res.* 2018;20(1):114.
127. Hadgu E, Seifu D, Tigneh W, Bokretsion Y, Bekele A, Abebe M, et al. Breast cancer in Ethiopia: evidence for geographic difference in the distribution of molecular subtypes in Africa. *BMC Women's Health.* 2018;18(1):40-.
128. Li J, Huang Y, Zhang B-N, Fan J-H, Huang R, Zhang P, et al. Body Mass Index and Breast Cancer Defined by Biological Receptor Status in Pre-Menopausal and Post-Menopausal Women: A Multicenter Study in China. *PLoS ONE.* 2014;9(1):e87224-e.
129. Bandera EV, Chandran U, Hong CC, Troester MA, Bethea TN, Adams-Campbell LL, et al. Obesity, body fat distribution, and risk of breast cancer subtypes in African American women participating in the AMBER Consortium. *Breast Cancer Res Treat.* 2015;150(3):655-66.
130. Yang XR, Chang-Claude J, Goode EL, Couch FJ, Nevanlinna H, Milne RL, et al. Associations of breast cancer risk factors with tumor subtypes: a pooled analysis from the Breast Cancer Association Consortium studies. *J Natl Cancer Inst.* 2011;103(3):250-63.
131. Bandera EV, Maskarinec G, Romieu I, John EM. Racial and ethnic disparities in the impact of obesity on breast cancer risk and survival: a global perspective. *Advances in nutrition (Bethesda, Md).* 2015;6(6):803-19.
132. Varela-Rey M, Woodhoo A, Martinez-Chantar ML, Mato JM, Lu SC. Alcohol, DNA methylation, and cancer. 2012. p. 25-35.
133. Baglia ML, Cook LS, Mei-Tzu C, Wiggins C, Hill D, Porter P, et al. Alcohol, smoking, and risk of H er2-overexpressing and triple-negative breast cancer relative to estrogen receptor-positive breast cancer. *International journal of cancer.* 2018;143(8):1849-57.
134. Wong AW, Dunlap SM, Holcomb VB, Nunez NP. Alcohol promotes mammary tumor development via the estrogen pathway in estrogen receptor alpha-negative HER2/neu mice. *Alcoholism: Clinical and Experimental Research.* 2012;36(4):577-87.
135. Zhang L, Huang Y, Feng Z, Wang X, Li H, Song F, et al. Comparison of breast cancer risk factors among molecular subtypes: A case-only study. *Cancer medicine.* 2019;8(4):1882-92.
136. Anderson WF, Pfeiffer RM, Wohlfahrt J, Ejlersen B, Jensen MB, Kroman N. Associations of parity-related reproductive histories with ER $\pm$ and HER2 $\pm$ receptor-specific breast cancer aetiology. *International journal of epidemiology.* 2017;46(1):86-95.
137. Fortner RT, Sisti J, Chai B, Collins LC, Rosner B, Hankinson SE, et al. Parity, breastfeeding, and breast cancer risk by hormone receptor status and molecular phenotype: results from the Nurses' Health Studies. *Breast Cancer Research.* 2019;21(1):40-.
138. Gaudet MM, Gierach GL, Carter BD, Luo J, Milne RL, Weiderpass E, et al. Pooled Analysis of Nine Cohorts Reveals Breast Cancer Risk Factors by Tumor Molecular Subtype. *Cancer Res.* 2018;78(20):6011-21.
139. Chen L, Li CI, Tang MTC, Porter P, Hill DA, Wiggins CL, et al. Reproductive factors and risk of luminal, HER2-overexpressing, and triple-negative breast cancer among multiethnic women. *Cancer Epidemiology Biomarkers & Prevention.* 2016;25(9):1297-304.
140. Ma H, Bernstein L, Pike MC, Ursin G. Reproductive factors and breast cancer risk according to joint estrogen and progesterone receptor status: a meta-analysis of epidemiological studies. *Breast Cancer Res.* 2006;8(4):R43.
141. Islami F, Liu Y, Jemal A, Zhou J, Weiderpass E, Colditz G, et al. Breastfeeding and breast cancer risk by receptor status—a systematic review and meta-analysis. *Annals of Oncology.* 2015;26(12):2398-407.

142. Brouckaert O, Rudolph A, Laenen A, Keeman R, Bolla MK, Wang Q, et al. Reproductive profiles and risk of breast cancer subtypes: a multi-center case-only study. *Breast Cancer Research*. 2017;19(1):119-.
143. Martinez ME, Wertheim BC, Natarajan L, Schwab R, Bondy M, Daneri-navarro A, et al. Reproductive factors, heterogeneity, and breast tumor subtypes in women of Mexican descent. 2013;1853-62.
144. Hsu A, Mustoe TA. The principles of wound healing. *Plastic surgery secrets plus*. 2010;2:3-7.
145. Mitchell R. Tissue repair: cellular growth, fibrosis and wound healing. *Pathologic basis of disease*. 1999:89-112.
146. Kim R. Cancer immunoediting: from immune surveillance to immune escape. *Cancer Immunotherapy*. 2007:9-27.
147. Fishbein A, Hammock BD, Serhan CN, Panigrahy D. Carcinogenesis: Failure of resolution of inflammation? *Pharmacology & Therapeutics*. 2020:107670.
148. Finak G, Bertos N, Pepin F, Sadekova S, Souleimanova M, Zhao H, et al. Stromal gene expression predicts clinical outcome in breast cancer. *Nature medicine*. 2008;14(5):518-27.
149. Yang C-R, Hsieh S-L, Ho F-M, Lin W-W. Decoy receptor 3 increases monocyte adhesion to endothelial cells via NF- κ B-dependent up-regulation of intercellular adhesion molecule-1, VCAM-1, and IL-8 expression. *The journal of immunology*. 2005;174(3):1647-56.
150. F R. The Journal of The American Society of Hematology. *Mol Biol*. 1990;76(4):655–63. . *Molecular Biology*. 1990;76(4):655-63.
151. Ames BN, Gold LS, Willett WC. The causes and prevention of cancer. *Proceedings of the National Academy of Sciences*. 1995;92(12):5258-65.
152. Müller A, Homey B, Soto H, Ge N, Catron D, Buchanan ME, et al. Involvement of chemokine receptors in breast cancer metastasis. *Nature*. 2001;410(6824):50-6.
153. Nakamoto Y, Guidotti LG, Kuhlen CV, Fowler P, Chisari FV. Immune pathogenesis of hepatocellular carcinoma. *J Exp Med*. 1998;188(2):341-50.
154. Jacoby RF, Seibert K, Cole CE, Kelloff G, Lubet RA. The cyclooxygenase-2 inhibitor celecoxib is a potent preventive and therapeutic agent in the min mouse model of adenomatous polyposis. *Cancer Res*. 2000;60(18):5040-4.
155. Mantovani A, Allavena P, Sica A, Balkwill F. Cancer-related inflammation. *Nature*. 2008;454(7203):436-44.
156. Sparmann A, Bar-Sagi D. Ras-induced interleukin-8 expression plays a critical role in tumor growth and angiogenesis. *Cancer Cell*. 2004;6(5):447-58.
157. Zeh HJ, Lotze MT. Addicted to death: invasive cancer and the immune response to unscheduled cell death. *Journal of immunotherapy (Hagerstown, Md : 1997)*. 2005;28(1):1-9.
158. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011;144(5):646-74.
159. Naito Y, Saito K, Shiiba K, Ohuchi A, Saigenji K, Nagura H, et al. CD8+ T cells infiltrated within cancer cell nests as a prognostic factor in human colorectal cancer. *Cancer Res*. 1998;58(16):3491-4.
160. Nakano O, Sato M, Naito Y, Suzuki K, Orikasa S, Aizawa M, et al. Proliferative activity of intratumoral CD8(+) T-lymphocytes as a prognostic factor in human renal cell carcinoma: clinicopathologic demonstration of antitumor immunity. *Cancer Res*. 2001;61(13):5132-6.
161. Dranoff G. Tumour immunology Immune recognition and tumor protection Editorial overview. *Current Opinion in Immunology*. 2002;14:161-4.
162. Elgert KD, Alleva DG, Mullins DW. Tumor-induced immune dysfunction: the macrophage connection. *Journal of leukocyte biology*. 1998;64(3):275-90.
163. Coussens LM, Zitvogel L, Palucka AK. Neutralizing Tumor-Promoting Chronic Inflammation : A Magic Bullet ? *Science*. 2013;286(6117):286-91.

164. Fernandes A, Bianchi G, Feltri AP, Pérez M, Correnti M. Presence of human papillomavirus in breast cancer and its association with prognostic factors. *Ecancermedicalscience*. 2015;9.
165. Purrahan D, Avarvand AY, Paradowska-Gorycka A, Saki N, Karimpourian H, Jodat H, et al. Association of human papillomavirus with breast cancer: a new perspective on an old debate. *Future Oncology*. 2022;18(22):2483-94.
166. Tishkoff SA, Reed FA, Friedlaender FR, Ehret C, Ranciaro A, Froment A, et al. The genetic structure and history of Africans and African Americans. *Science*. 2009;324(5930):1035-44.
167. Eng A, McCormack V, dos-Santos-Silva I. Receptor-Defined Subtypes of Breast Cancer in Indigenous Populations in Africa: A Systematic Review and Meta-Analysis. *PLoS Medicine*. 2014;11(9):e1001720-e.
168. Daub JT, Hofer T, Cutivet E, Dupanloup I, Quintana-Murci L, Robinson-Rechavi M, et al. Evidence for polygenic adaptation to pathogens in the human genome. *Molecular biology and evolution*. 2013;30(7):1544-58.
169. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Yang T-HO, et al. The immune landscape of cancer. *Immunity*. 2019;51(2):411-2.
170. Dushyanthen S, Beavis PA, Savas P, Teo ZL, Zhou C, Mansour M, et al. Relevance of tumor-infiltrating lymphocytes in breast cancer. *BMC medicine*. 2015;13(1):202-.
171. Loi S, Michiels S, Salgado R, Sirtaine N, Jose V, Fumagalli D, et al. Tumor infiltrating lymphocytes are prognostic in triple negative breast cancer and predictive for trastuzumab benefit in early breast cancer: Results from the FinHER trial. *Annals of Oncology*. 2014;25(8):1544-50.
172. Zhu B, Tse LA, Wang D, Koka H, Zhang T, Abubakar M, et al. Immune gene expression profiling reveals heterogeneity in luminal breast tumors. *Breast Cancer Research*. 2019;21(1):1-11.
173. Mahmoud SMA, Paish EC, Powe DG, Macmillan RD, Grainge MJ, Lee AHS, et al. Tumor-Infiltrating CD8+ Lymphocytes Predict Clinical Outcome in Breast Cancer. *Journal of Clinical Oncology*. 2011;29(15):1949-55.
174. Das Roy L, Pathangey LB, Tinder TL, Schettini JL, Gruber HE, Mukherjee P. Breast-cancer-associated metastasis is significantly increased in a model of autoimmune arthritis. *Breast cancer research : BCR*. 2009;11(4):R56-R.
175. Baghban R, Roshangar L, Jahanban-Esfahlan R, Seidi K, Ebrahimi-Kalan A, Jaymand M, et al. Tumor microenvironment complexity and therapeutic implications at a glance. *Cell Commun Signal*. 2020;18(1):59.
176. Balkwill FR, Capasso M, Hagemann T. The tumor microenvironment at a glance. *J Cell Sci*. 2012;125(Pt 23):5591-6.
177. Galluzzi L, Buqué A, Kepp O, Zitvogel L, Kroemer G. Immunological Effects of Conventional Chemotherapy and Targeted Anticancer Agents. *Cancer Cell*. 2015;28(6):690-714.
178. Cohen IJ, Blasberg R. Impact of the Tumor Microenvironment on Tumor-Infiltrating Lymphocytes: Focus on Breast Cancer. *Breast Cancer: Basic and Clinical Research*. 2017;11:117822341773156-.
179. Gonzalez H, Hagerling C, Werb Z. Roles of the immune system in cancer: From tumor initiation to metastatic progression. *Cold Spring Harbor Laboratory Press*; 2018. p. 1267-84.
180. Salgado R DC, Demaria S, Sirtaine N, Klauschen F, Pruneri G, et al. The evaluation of tumour-infiltrating lymphocytes (TILs) in breast cancer: recommendations by an International TILs Working Group 2014. *Annals of Oncology*. 2015.
181. Badr Nahla M BF, Shaaban Abeer M. The Immune Microenvironment in Breast Carcinoma: Predictive and Prognostic Role in the Neoadjuvant Setting. . *Pathobiology*. 2020.

182. Tower H, Ruppert M, Britt K. The Immune Microenvironment of Cancer. *Cancers*. 2019;11(1375):1-15.
183. Denkert C, Loibl S, Noske A, Roller M, Müller BM, Komor M, et al. Tumor-associated lymphocytes as an independent predictor of response to neoadjuvant chemotherapy in breast cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2010;28(1):105-13.
184. Johansson M, DeNardo DG, Coussens LM. Polarized immune responses differentially regulate cancer development. *Immunol Rev*. 2008;222(4):145-54.
185. Althobiti M, Aleskandarany MA, Joseph C, Toss M, Mongan N, Diez-Rodriguez M, et al. Heterogeneity of tumour-infiltrating lymphocytes in breast cancer and its prognostic significance. *Histopathology*. 2018;73(6):887-96.
186. Anz D, Eiber S, Scholz C, Endres S, Kirchner T, Bourquin C, et al. In breast cancer, a high ratio of tumour-infiltrating intraepithelial CD8+ to FoxP3+ cells is characteristic for the medullary subtype. *Histopathology*. 2011;59(5):965-74.
187. Verma R, Hanby AM, Horgan K, Verghese ET, Volpato M, Carter CR, et al. Levels of different subtypes of tumour-infiltrating lymphocytes correlate with each other, with matched circulating lymphocytes, and with survival in breast cancer. *Breast Cancer Res Treat*. 2020;183(1):49-59.
188. Stanton SE, Disis ML. Clinical significance of tumor-infiltrating lymphocytes in breast cancer. *Journal for ImmunoTherapy of Cancer*. 2016;4(1):59-.
189. West NR, Kost SE, Martin SD, Milne K, J R, Nelson BH, et al. Tumour-infiltrating FOXP3⁺ lymphocytes are associated with cytotoxic immune responses and good clinical outcome in oestrogen receptor-negative breast cancer. *British Journal of Cancer*. 2012;108(1):155-62.
190. Luen SJ, Savas P, Fox SB, Salgado R, Loi S. Tumour-infiltrating lymphocytes and the emerging role of immunotherapy in breast cancer. *Pathology*. 2017;49(2):141-55.
191. Cimino-Mathews A. Tumor-infiltrating lymphocytes and PD-L1 in breast cancer (and, what happened to medullary carcinoma?). *Diagnostic Histopathology*. 2021;27(4):148-54.
192. Wei Y, Yu Y, Ji Y, Zhong Y, Min N, Hu H, et al. Surgical management in phyllodes tumors of the breast: a systematic review and meta-analysis. *Gland Surgery*. 2022;11(3):513.
193. Brady NJ, Chuntova P, Schwertfeger KL. Macrophages : Regulators of the Inflammatory Microenvironment during Mammary Gland Development and Breast Cancer. *Mediators of inflammation*. 2016:1-13.
194. Campbell MJ, Tonlaar NY, Garwood ER, Huo D, Moore DH, Khramtsov AI, et al. Proliferating macrophages associated with high grade, hormone receptor negative breast cancer and poor clinical outcome. *Breast Cancer Research and Treatment*. 2011;128(3):703-11.
195. Zhang X, Tian W, Cai X, Wang X, Dang W, Tang H, et al. Hydrazinocurcumin Encapsulated nanoparticles "re-educate" tumor-associated macrophages and exhibit anti-tumor effects on breast cancer following STAT3 suppression. *PLoS One*. 2013;8(6):e65896.
196. Huang Y, Goel S, Duda DG, Fukumura D, Jain RK. Vascular normalization as an emerging strategy to enhance cancer immunotherapy. *Cancer Res*. 2013;73(10):2943-8.
197. Shevach EM. Mechanisms of foxp3+ T regulatory cell-mediated suppression. *Immunity*. 2009;30(5):636-45.
198. Oleinika K, Nibbs RJ, Graham GJ, Fraser AR. Suppression, subversion and escape: The role of regulatory T cells in cancer progression. *Clinical and Experimental Immunology*. 2013;171(1):36-45.
199. Bohling SD, Allison KH. Immunosuppressive regulatory T cells are associated with aggressive breast cancer phenotypes: a potential therapeutic target. *Mod Pathol*. 2008;21(12):1527-32.

200. Gobert M, Treilleux I, Bendriss-Vermare N, Bachelot T, Goddard-Leon S, Arfi V, et al. Regulatory T cells recruited through CCL22/CCR4 are selectively activated in lymphoid infiltrates surrounding primary breast tumors and lead to an adverse clinical outcome. *Cancer Res.* 2009;69(5):2000-9.
201. Yeong J, Lim JCT, Lee B, Li H, Chia N, Ong CCH, et al. High densities of tumor-associated plasma cells predict improved prognosis in triple negative breast cancer. *Frontiers in Immunology.* 2018;9(MAY):2-11.
202. Gabrilovich DI, Nagaraj S. Myeloid-derived suppressor cells as regulators of the immune system. *Nat Rev Immunol.* 2009;9(3):162-74.
203. Markowitz J, Wesolowski R, Papenfuss T, Brooks TR, Carson WE, 3rd. Myeloid-derived suppressor cells in breast cancer. *Breast Cancer Res Treat.* 2013;140(1):13-21.
204. Zou W, Restifo NP. T(H)17 cells in tumour immunity and immunotherapy. *Nat Rev Immunol.* 2010;10(4):248-56.
205. Hong CC, Yao S, McCann SE, Dolnick RY, Wallace PK, Gong Z, et al. Pretreatment levels of circulating Th1 and Th2 cytokines, and their ratios, are associated with ER-negative and triple negative breast cancers. *Breast Cancer Research and Treatment.* 2013;139(2):477-88.
206. Welsh TJ, Green RH, Richardson D, Waller DA, O'Byrne KJ, Bradding P. Macrophage and mast-cell invasion of tumor cell islets confers a marked survival advantage in non-small-cell lung cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology.* 2005;23(35):8959-67.
207. Dabiri S, Huntsman D, Makretsov N, Cheang M, Gilks B, Bajdik C, et al. The presence of stromal mast cells identifies a subset of invasive breast cancers with a favorable prognosis. *Mod Pathol.* 2004;17(6):690-5.
208. Ma Y, Shurin GV, Peiyuan Z, Shurin MR. Dendritic cells in the cancer microenvironment. *J Cancer.* 2013;4(1):36-44.
209. Mbongue J, Nicholas D, Firek A, Langridge W. The role of dendritic cells in tissue-specific autoimmunity. *Journal of immunology research.* 2014;2014:857143-.
210. Lotze MT. Getting to the source: dendritic cells as therapeutic reagents for the treatment of patients with cancer. *Annals of surgery.* 1997;226(1):1-5.
211. Wein L, Savas P, Luen SJ, Virassamy B, Salgado R, Loi S. Clinical Validity and Utility of Tumor-Infiltrating Lymphocytes in Routine Clinical Practice for Breast Cancer Patients: Current and Future Directions. *Front Oncol.* 2017;7:156.
212. Solinas C, Carbognin L, De Silva P, Criscitiello C, Lambertini M. Tumor-infiltrating lymphocytes in breast cancer according to tumor subtype: Current state of the art. *Breast.* 2017;35:142-50.
213. Luen SJ, Salgado R, Fox S, Savas P, Eng-Wong J, Clark E, et al. Tumour-infiltrating lymphocytes in advanced HER2-positive breast cancer treated with pertuzumab or placebo in addition to trastuzumab and docetaxel: a retrospective analysis of the CLEOPATRA study. *The Lancet Oncology.* 2017;18(1):52-62.
214. Yu X, Zhang Z. Prognostic and predictive value of tumor-infiltrating lymphocytes in breast cancer : a systematic review and meta-analysis. 2016:497-506.
215. Salgado R, Denkert C, Campbell C, Savas P, Nuciforo P, Aura C, et al. Tumor-infiltrating lymphocytes and associations with pathological complete response and event-free survival in HER2-positive early-stage breast cancer treated with lapatinib and trastuzumab: a secondary analysis of the NeoALTTO trial. *JAMA oncology.* 2015;1(4):448-55.
216. Loi S, Drubay D, Adams S, Pruneri G, Francis PA, Lacroix-Triki M, et al. Tumor-infiltrating lymphocytes and prognosis: a pooled individual patient analysis of early-stage triple-negative breast cancers. *Journal of clinical oncology.* 2019;37(7):559.

217. Wang K, Shen T, Siegal GP, Wei S. The CD4/CD8 ratio of tumor-infiltrating lymphocytes at the tumor-host interface has prognostic value in triple-negative breast cancer. *Hum Pathol.* 2017;69:110-7.
218. Denkert C, von Minckwitz G, Darb-Esfahani S, Lederer B, Heppner BI, Weber KE, et al. Tumour-infiltrating lymphocytes and prognosis in different subtypes of breast cancer: a pooled analysis of 3771 patients treated with neoadjuvant therapy. *The Lancet Oncology.* 2018;19(1):40-50.
219. Criscitiello C, Esposito A, Trapani D, Curigliano G. Prognostic and predictive value of tumor infiltrating lymphocytes in early breast cancer. *Cancer treatment reviews.* 2016;50:205-7.
220. Gao G, Wang Z, Qu X, Zhang Z. Prognostic value of tumor-infiltrating lymphocytes in patients with triple-negative breast cancer: a systematic review and meta-analysis. *BMC Cancer.* 2020;20(1):179.
221. Mao Y, Qu Q, Chen X, Huang O, Wu J, Shen K. The Prognostic Value of Tumor-Infiltrating Lymphocytes in Breast Cancer : A Systematic Review and Meta-Analysis. 2016:1-13.
222. Gao Z, Xu Y, Sun C, Wang X, Guo Y, Qiu S, et al. A systematic review of asymptomatic infections with COVID-19. *Journal of Microbiology, Immunology and Infection.* 2020(January).
223. Afghahi A, Purington N, Han SS, Desai M, Pierson E, Mathur MB, et al. Higher Absolute Lymphocyte Counts Predict Lower Mortality from Early-Stage Triple-Negative Breast Cancer. *Clin Cancer Res.* 2018;24(12):2851-8.
224. Choudhury A, Aron S, Sengupta D, Hazelhurst S, Ramsay M. African genetic diversity provides novel insights into evolutionary history and local adaptations. *Human molecular genetics.* 2018;27(R2):R209-R18.
225. Sawe RT, Kerper M, Badve S, Li J, Sandoval-cooper M, Xie J, et al. Aggressive breast cancer in western Kenya has early onset , high proliferation , and immune cell infiltration. *BMC Cancer.* 2016:1-15.
226. Sawe RT, Mining SK, Ofulla AV, Patel K, Guyah B, Chumba D, et al. Tumor infiltrating leukocyte density is independent of tumor grade and molecular subtype in aggressive breast cancer of Western Kenya. *Tropical medicine and health.* 2017;45(1):1-11.
227. O'Meara T, Marczyk M, Qing T, Yaghoobi V, Blenman K, Cole K, et al. Immunological Differences Between Immune-Rich Estrogen Receptor-Positive and Immune-Rich Triple-Negative Breast Cancers. *JCO Precision Oncology.* 2020;3:767-79.
228. Ahmed Elkhanany EKAKT. Abstract P2-10-03: Detrimental impact of T-regulatory cells on outcome of breast cancer in African American population. 2020;80(4).
229. Abdou Y, Attwood K, Cheng T-yD, Yao S, Bandera EV, Zirpoli GR, et al. Racial differences in CD8 + T cell infiltration in breast tumors from Black and White women. 2020:1-10.
230. Chen CH, Lu YS, Cheng AL, Huang CS, Kuo WH, Wang MY, et al. Disparity in tumor immune microenvironment of breast cancer and prognostic impact: Asian versus Western populations. *The oncologist.* 2020;25(1):e16.
231. Jin YW, Hu P. Tumor-Infiltrating CD8 T Cells Predict Clinical Breast Cancer Outcomes in Young Women. *Cancers (Basel).* 2020;12(5).
232. Floris G, Richard F, Hamy A-S, Jongen L, Wildiers H, Ardui J, et al. Body mass index and tumor infiltrating lymphocytes in triple-negative breast cancer. *JNCI: Journal of the National Cancer Institute.* 2020.
233. Simon V, Laot L, Laas E, Rozette S, Guerin J, Balezeau T, et al. No Impact of Smoking Status on Breast Cancer Tumor Infiltrating Lymphocytes, Response to Neoadjuvant Chemotherapy and Prognosis. *Cancers (Basel).* 2020;12(10).

234. Salgado R, Denkert C, Demaria S, Sirtaine N, Klauschen F, Pruner G, et al. The evaluation of tumor-infiltrating lymphocytes (TILs) in breast cancer: recommendations by an International TILs Working Group 2014. *Annals of Oncology*. 2015;26(2):259-71.
235. Abubakar M, Zhang J, Ahearn TU, Koka H, Guo C, Lawrence SM, et al. Tumor-Associated Stromal Cellular Density as a Predictor of Recurrence and Mortality in Breast Cancer: Results from Ethnically Diverse Study Populations. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 2021;30(7):1397-407.
236. Hendry S, Salgado R, Gevaert T, Russell PA, John T, Thapa B, et al. Assessing tumor infiltrating lymphocytes in solid tumors: a practical review for pathologists and proposal for a standardized method from the International Immuno-Oncology Biomarkers Working Group: Part 2: TILs in melanoma, gastrointestinal tract carcinom. *Advances In Anatomic Pathology*. 2017;24(6):311-35.
237. Solinas C, Ceppi M, Lambertini M, Scartozzi M, Buisseret L, Garaud S, et al. Tumor-infiltrating lymphocytes in patients with HER2-positive breast cancer treated with neoadjuvant chemotherapy plus trastuzumab, lapatinib or their combination: A meta-analysis of randomized controlled trials. *Cancer treatment reviews*. 2017;57:8-15.
238. Salgado R, Loi S. Tumour infiltrating lymphocytes in breast cancer: increasing clinical relevance. *The Lancet Oncology*. 2018;19(1):3-5.
239. Leon-Ferre RA, Hieken TJ, Boughhey JC. The Landmark Series: Neoadjuvant Chemotherapy for Triple-Negative and HER2-Positive Breast Cancer. *Ann Surg Oncol*. 2021;28(4):2111-9.
240. Ochi T, Bianchini G, Ando M, Nozaki F, Kobayashi D, Criscitiello C, et al. Predictive and prognostic value of stromal tumour-infiltrating lymphocytes before and after neoadjuvant therapy in triple negative and HER2-positive breast cancer. *Eur J Cancer*. 2019;118:41-8.
241. Krishnamurti U, Wetherilt CS, Yang J, Peng L, Li X. Tumor-infiltrating lymphocytes are significantly associated with better overall survival and disease-free survival in triple-negative but not estrogen receptor-positive breast cancers. *Hum Pathol*. 2017;64:7-12.
242. Burstein HJ, Curigliano G, Loibl S, Dubsy P, Gnant M, Poortmans P, et al. Estimating the benefits of therapy for early-stage breast cancer: the St. Gallen International Consensus Guidelines for the primary therapy of early breast cancer 2019. *Ann Oncol*. 2019;30(10):1541-57.
243. He L, Wang Y, Wu Q, Song Y, Ma X, Zhang B, et al. Association between levels of tumor-infiltrating lymphocytes in different subtypes of primary breast tumors and prognostic outcomes: a meta-analysis. *BMC Womens Health*. 2020;20(1):194.
244. Garg SK, Delaney C, Toubai T, Ghosh A, Reddy P, Banerjee R, et al. Aging is associated with increased regulatory T-cell function. *Aging cell*. 2014;13(3):441-8.
245. Shou J, Zhang Z, Lai Y, Chen Z, Huang J. Worse outcome in breast cancer with higher tumor-infiltrating FOXP3+ Tregs : a systematic review and meta-analysis. *BMC Cancer*. 2016;16(1):687.
246. Sato E, Olson SH, Ahn J, Bundy B, Nishikawa H, Qian F, et al. Intraepithelial CD8+ tumor-infiltrating lymphocytes and a high CD8+/regulatory T cell ratio are associated with favorable prognosis in ovarian cancer. *Proc Natl Acad Sci U S A*. 2005;102(51):18538-43.
247. Miyashita M, Sasano H, Tamaki K, Hirakawa H, Takahashi Y, Nakagawa S, et al. Prognostic significance of tumor-infiltrating CD8+ and FOXP3+ lymphocytes in residual tumors and alterations in these parameters after neoadjuvant chemotherapy in triple-negative breast cancer: A retrospective multicenter study. *Breast Cancer Research*. 2015;17(1):1-13.
248. Peng G-l, Li L, Guo Y-w, Yu P, Yin X-j, Wang S, et al. CD8+ cytotoxic and FoxP3+ regulatory T lymphocytes serve as prognostic factors in breast cancer. 2019;11(8):5039-53.

249. Meng S, Li L, Zhou M, Jiang W, Niu H, Yang K. Distribution and prognostic value of tumor-infiltrating T cells in breast cancer. *Mol Med Rep.* 2018;18(5):4247-58.
250. Ni C, Yang L, Xu Q, Yuan H, Wang W, Xia W, et al. CD68- and CD163-positive tumor infiltrating macrophages in non-metastatic breast cancer: a retrospective study and meta-analysis. *J Cancer.* 2019;10(19):4463-72.
251. Burugu S, Gao D, Leung S, Chia S, Nielsen T. LAG-3+ tumor infiltrating lymphocytes in breast cancer: clinical correlates and association with PD-1/PD-L1+ tumors. *Annals of Oncology.* 2017;28(12):2977-84.
252. Nicolini A, Carpi A, Rossi G. Cytokines in breast cancer. *Cytokine and Growth Factor Reviews.* 2006;17(5):325-37.
253. Yao M, Brummer G, Acevedo D, Cheng N. Cytokine Regulation of Metastasis and Tumorigenicity. *Advances in cancer research.* 2016;132:265-367.
254. Turner MD, Nedjai B, Hurst T, Pennington DJ. Cytokines and chemokines: At the crossroads of cell signalling and inflammatory disease. *Biochimica et Biophysica Acta - Molecular Cell Research.* 2014;1843(11):2563-82.
255. Ayari J. Prognostic Value of Circulating Cytokines in Breast Cancer. *Cancer Med J.* 2020;3(1):1-9.
256. Julie AS, Andreas JP. Bioanalytical Chemistry of Cytokines-A Review Julie. *Anal Chim Acta.* 2015:95-115.
257. Hamed EA, Zakhary MM, Maximous DW. Apoptosis, angiogenesis, inflammation, and oxidative stress: Basic interactions in patients with early and metastatic breast cancer. *Journal of Cancer Research and Clinical Oncology.* 2012;138(6):999-1009.
258. Chen F, Long Q, Fu D, Zhu D, Ji Y, Han L, et al. Targeting SPINK1 in the damaged tumour microenvironment alleviates therapeutic resistance. *Nature communications.* 2018;9(1):1-19.
259. Contreras LB, Cid RC, Lorca LA, Ribeiro IL. Effectiveness of Resistance Training on Fatigue in Patients Undergoing Cancer Treatment: A Meta-Analysis of Randomized Clinical Trials. *International journal of breast cancer.* 2022;2022.
260. Esquivel-Velázquez M, Ostoa-Saloma P, Palacios-Arreola MI, Nava-Castro KE, Castro JI, Morales-Montor J. The role of cytokines in breast cancer development and progression. *Journal of Interferon & Cytokine Research.* 2015;35(1):1-16.
261. Setrerrahmane S, Xu H. Tumor-related interleukins: Old validated targets for new anti-cancer drug development. *BioMed Central Ltd.;* 2017. p. 1-17.
262. Liu T, Zhang L, Joo D, Sun SC. NF- κ B signaling in inflammation. *Signal transduction and targeted therapy.* 2017;2:17023-.
263. Xie R, Zhang Y, Zhao N, Zhou S, Wang X, Han W, et al. Pharmacokinetics and safety of recombinant human interleukin-1 receptor antagonist GR007 in healthy Chinese subjects. *European Journal of Drug Metabolism and Pharmacokinetics.* 2019;44(3):353-60.
264. Nava-Castro KE, Ramírez-Nieto R, Méndez-García LA, Girón-Pérez MI, Segovia-Mendoza M, Navidad-Murrieta MS, et al. Environmental Pollution as a Risk Factor in Testicular Tumour Development: Focus on the Interaction between Bisphenol A and the Associated Immune Response. *International journal of environmental research and public health.* 2019;16(21).
265. Morse MA, Cox GJ, Van Kessel AG, Campos M, Babiuk LA. Interleukin-2. *Cytokines in Animal Health and Disease: CRC Press;* 2020. p. 89-120.
266. Rosenberg SA, Yang JC, Sherry RM, Kammula US, Hughes MS, Phan GQ, et al. Durable complete responses in heavily pretreated patients with metastatic melanoma using T-cell transfer immunotherapy. *Clin Cancer Res.* 2011;17(13):4550-7.
267. Iwaszko M, Biały S, Bogunia-Kubik K. Significance of Interleukin (IL)-4 and IL-13 in Inflammatory Arthritis. *Cells.* 2021;10(11).

268. Li Z, Chen L, Qin Z. Paradoxical roles of IL-4 in tumor immunity. *Cell Mol Immunol*; 2009. p. 415-22.
269. Venmar KT, Carter KJ, Hwang DG, Dozier EA, Fingleton B. IL4 receptor ILR4 α regulates metastatic colonization by mammary tumors through multiple signaling pathways. *Cancer Research*. 2014;74(16):4329-40.
270. Gyamfi J, Lee YH, Eom M, Choi J. Interleukin-6/STAT3 signalling regulates adipocyte induced epithelial-mesenchymal transition in breast cancer cells. *Sci Rep*. 2018;8(1):8859.
271. Masjedi A, Hashemi V, Hojjat-Farsangi M, Ghalamfarsa G, Azizi G, Yousefi M, et al. The significant role of interleukin-6 and its signaling pathway in the immunopathogenesis and treatment of breast cancer. *Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie*. 2018;108:1415-24.
272. Dethlefsen C, Højfeldt G. The role of intratumoral and systemic IL-6 in breast cancer. 2013;657-64.
273. Alfaro C, Sanmamed MF, Rodríguez-Ruiz ME, Teijeira Á, Oñate C, González Á, et al. Interleukin-8 in cancer pathogenesis, treatment and follow-up. *Cancer treatment reviews*. 2017;60:24-31.
274. Brat DJ, Bellail AC, Van Meir EG. The role of interleukin-8 and its receptors in gliomagenesis and tumoral angiogenesis. *Neuro Oncol*. 2005;7(2):122-33.
275. Ma Y, Ren Y, Dai ZJ, Wu CJ, Ji YH, Xu J. IL-6, IL-8 and TNF- α levels correlate with disease stage in breast cancer patients. *Advances in clinical and experimental medicine : official organ Wroclaw Medical University*. 2017;26(3):421-6.
276. Hamidullah, Changkija B, Konwar R. Role of interleukin-10 in breast cancer. *Breast Cancer Res Treat*. 2012;133(1):11-21.
277. Hamidullah, Changkija B, Konwar R. Role of interleukin-10 in breast cancer. *Breast Cancer Research and Treatment*. 2012;133(1):11-21.
278. Mohamed MM, El-Ghonaimy EA, Nouh MA, Schneider RJ, Sloane BF, El-Shinawi M. Cytokines secreted by macrophages isolated from tumor microenvironment of inflammatory breast cancer patients possess chemotactic properties. *The international journal of biochemistry & cell biology*. 2014;46:138-47.
279. Mercogliano MF, Bruni S, Elizalde PV, Schillaci R. Tumor Necrosis Factor α Blockade: An Opportunity to Tackle Breast Cancer. *Frontiers Media S.A.*; 2020. p. 584-.
280. Imamura T, Hikita A, Inoue Y. The roles of TGF- β signaling in carcinogenesis and breast cancer metastasis. *Breast Cancer*. 2012;19(2):118-24.
281. Pasqualini JR, Gelly C, Nguyen BL, Vella C. Importance of estrogen sulfates in breast cancer. *J Steroid Biochem*. 1989;34(1-6):155-63.
282. Casneuf T, Axel AE, King P, Alvarez JD, Werbeck JL, Verhulst T, et al. Interleukin-6 is a potential therapeutic target in interleukin-6 dependent, estrogen receptor- α -positive breast cancer. *Breast Cancer (Dove Med Press)*. 2016;8:13-27.
283. Salgado R, Junius S, Benoy I, Van Dam P, Vermeulen P, Van Marck E, et al. Circulating interleukin-6 predicts survival in patients with metastatic breast cancer. *Int J Cancer*. 2003;103(5):642-6.
284. Yu KD, Di GH, Fan L, Chen AX, Yang C, Shao ZM. Lack of an association between a functional polymorphism in the interleukin-6 gene promoter and breast cancer risk: a meta-analysis involving 25,703 subjects. *Breast Cancer Res Treat*. 2010;122(2):483-8.
285. Ma YF, Chen C, Li D, Liu M, Lv ZW, Ji Y, et al. Targeting of interleukin (IL)-17A inhibits PDL1 expression in tumor cells and induces anticancer immunity in an estrogen receptor-negative murine model of breast cancer. *Oncotarget*. 2017;8(5):7614-24.
286. Matsumoto H, Koo SL, Dent R, Tan PH, Iqbal J. Role of inflammatory infiltrates in triple negative breast cancer. *Journal of Clinical Pathology*. 2015;68(7):506-10.

287. Hartman ZC, Poage GM, Den Hollander P, Tsimelzon A, Hill J, Panupinthu N, et al. Growth of triple-negative breast cancer cells relies upon coordinate autocrine expression of the proinflammatory cytokines IL-6 and IL-8. *Cancer research*. 2013;73(11):3470-80.
288. Todorović-Raković N, Milovanović J. Interleukin-8 in breast cancer progression. *Journal of Interferon & Cytokine Research*. 2013;33(10):563-70.
289. Charafe-jauffret E, Ginestier C, Iovino F, Wicinski J, Cervera N, Finetti P, et al. Breast cancer cell lines contain functional cancer stem cells with metastatic capacity and distinct molecular signature. *Cancer Research*. 2009;69(4):1302-13.
290. Liu Z-j, Semenza GL, Zhang H-f. Hypoxia-inducible factor 1 and breast cancer metastasis. *Journal of Zhejiang University-SCIENCE B*. 2015;16(1):32-43.
291. Whiteside TL. Cytokine assays. *BioTechniques*. 2002;33(4 SUPPL.).
292. Ribeiro-Silva A, Zhang H, Jeffrey SS. RNA extraction from ten year old formalin-fixed paraffin-embedded breast cancer samples: a comparison of column purification and magnetic bead-based technologies. *BMC Mol Biol*. 2007;8:118.
293. Donczo B, Guttman A. Biomedical analysis of formalin-fixed, paraffin-embedded tissue samples: The Holy Grail for molecular diagnostics. *J Pharm Biomed Anal*. 2018;155:125-34.
294. Specht K, Richter T, Müller U, Walch A, Werner M, Höfler H. Quantitative gene expression analysis in microdissected archival formalin-fixed and paraffin-embedded tumor tissue. *Am J Pathol*. 2001;158(2):419-29.
295. Yin Y, Mitson-Salazar A, Prussin C. Detection of intracellular cytokines by flow cytometry. *Current Protocols in Immunology*. 2015;2015(August 2015):6.24.1-6.18.
296. Earley MC, Vogt Jr RF, Shapiro HM, Mandy FF, Kellar KL, Bellisario R, et al. Report from a workshop on multianalyte microsphere assays. *Wiley Online Library*; 2002.
297. Schutte RJ, Xie L, Klitzman B, Reichert WM. In vivo cytokine-associated responses to biomaterials. *Biomaterials*. 2009;30(2):160-8.
298. Schweitzer B, Roberts S, Grimwade B, Shao W, Wang M, Fu Q, et al. Multiplexed protein profiling on microarrays by rolling-circle amplification. *Nature biotechnology*. 2002;20(4):359-65.
299. Weller MG. Immunochromatographic techniques--a critical review. *Fresenius J Anal Chem*. 2000;366(6-7):635-45.
300. Abrahamsen HN, Steiniche T, Nexø E, Hamilton-Dutoit SJ, Sørensen BS. Towards quantitative mRNA analysis in paraffin-embedded tissues using real-time reverse transcriptase-polymerase chain reaction: a methodological study on lymph nodes from melanoma patients. *J Mol Diagn*. 2003;5(1):34-41.
301. Addis MF, Tanca A, Pagnozzi D, Crobu S, Fanciulli G, Cossu-Rocca P, et al. Generation of high-quality protein extracts from formalin-fixed, paraffin-embedded tissues. *Proteomics*. 2009;9(15):3815-23.
302. Newman LA, Kaljee LM. Health Disparities and Triple-Negative Breast Cancer in African American Women: A Review. *JAMA Surg*. 2017;152(5):485-93.
303. Esquivel-Velázquez M, Ostoa-Saloma P, Palacios-Arreola MI, Nava-Castro KE, Castro JI, Morales-Montor J. The role of cytokines in breast cancer development and progression. *Journal of interferon & cytokine research : the official journal of the International Society for Interferon and Cytokine Research*. 2015;35(1):1-16.
304. Chavey C, Bibeau F, Gourgou-Bourgade S, Burlinckon S, Boissière F, Laune D, et al. Oestrogen receptor negative breast cancers exhibit high cytokine content. *Breast Cancer Research*. 2007;9(1):R15-R.
305. Simeone AM, Nieves-Alicea R, McMurtry VC, Colella S, Krahe R, Tari AM. Cyclooxygenase-2 uses the protein kinase C/ interleukin-8/urokinase-type plasminogen activator pathway to increase the invasiveness of breast cancer cells. *International journal of oncology*. 2007;30(4):785-92.

306. Bolli E, Movahedi K, Laoui D, Van Ginderachter JA. Novel insights in the regulation and function of macrophages in the tumor microenvironment. *Current opinion in oncology*. 2017;29(1):55-61.
307. Lu X, Yang R, Zhang L, Xi Y, Zhao J, Wang F, et al. Macrophage Colony-stimulating Factor Mediates the Recruitment of Macrophages in Triple negative Breast Cancer. *Int J Biol Sci*. 2019;15(13):2859-71.
308. Qiu SQ, Waaijer SJH, Zwager MC, de Vries EGE, van der Vegt B, Schröder CP. Tumor-associated macrophages in breast cancer: Innocent bystander or important player? *Cancer treatment reviews*. 2018;70:178-89.
309. Hollmén M, Karaman S, Schwager S, Lisibach A, Christiansen AJ, Maksimow M, et al. G-CSF regulates macrophage phenotype and associates with poor overall survival in human triple-negative breast cancer. *OncoImmunology*. 2016;5(3):e1115177-e.
310. Lin S, Yang J, Elkahoun AG, Bandyopadhyay A, Wang L, Cornell JE, et al. Attenuation of TGF- β signaling suppresses premature senescence in a p21-dependent manner and promotes oncogenic Ras-mediated metastatic transformation in human mammary epithelial cells. *Molecular biology of the cell*. 2012;23(8):1569-81.
311. Zhang M, Wu J, Mao K, Deng H, Yang Y, Zhou E, et al. Role of transforming growth factor- β 1 in triple negative breast cancer patients. *International Journal of Surgery*. 2017;45:72-6.
312. Amsen D, de Visser KE, Town T. Approaches to determine expression of inflammatory cytokines. *Inflammation and Cancer: Springer*; 2009. p. 107-42.
313. Skalnikova H, Motlik J, Gadher SJ, Kovarova H. Mapping of the secretome of primary isolates of mammalian cells, stem cells and derived cell lines. *Proteomics*. 2011;11(4):691-708.
314. Sachdeva N, Asthana D. Cytokine quantitation: technologies and applications. *Front Biosci*. 2007;12:4682-95.
315. Favre N, Bordmann G, Rudin W. Comparison of cytokine measurements using ELISA, ELISPOT and semi-quantitative RT-PCR. *Journal of immunological methods*. 1997;204(1):57-66.

APPENDICES

APPENDIX I: Hospital Facilities in study.

Facility	Frequency	Percentage	Facility group
Aga Khan Hospital – Kisumu, Kenya	57	6.8	AKU
Aga Khan Hospital – Mombasa, Kenya	16	1.91	AKU
Aga Khan University Hospital, Nairobi	277	33.05	AKU
Garissa PGH	3	0.36	Other
Jaramogi Oginga Odinga Teaching and Referral Hospital	46	5.49	Other
Kenyatta National Hospital, Nairobi	85	10.14	Other
Kijabe Mission Hospital, AIC	105	12.53	Kijabe
Kisii General Hospital	4	0.48	Other
Moi Teaching and Referral Hospital, Eldoret	5	0.6	Other
Nyeri PGH	110	13.13	Nyeri
St Mary's Mission Hospital, Nairobi	122	14.56	St Mary's
Tenwek Mission Hospital	8	0.95	Other

Facility group	Frequency	Percentage
AKU (private)	350	41.77
Kijabe (faith based)	105	12.53
Nyeri (public)	110	13.13
St Mary's (faith based)	122	14.56
Other	151	18.02
Total	838	100

APPENDIX II: Associations of key risk factors with oestrogen-receptor negative (ER-) tumours, by hospitals (N=834)

Key risk factors	AKU (n=350, 42%)						Kijabe (n=105, 13%)					
	ER+		ER-		ER- vs. ER+		ER+		ER-		ER- vs. ER+	
	N	%	N	%	OR (95% CI) †	P†	N	%	N	%	OR (95% CI) †	
BMI												
Normal (<25.0)	42	18.9	25	26.9	1.00 (Ref)		27	40.9	8	38.1	1.00 (Ref)	
Overweight (25.0 - 29.9)	96	43.2	30	32.3	0.47 (0.25, 0.91)	0.026	25	37.9	8	38.1	1.20 (0.38, 3.82)	
Obese (≥30.0)	84	37.8	38	40.9	0.67 (0.35, 1.27)	0.22	14	21.2	5	23.8	1.07 (0.28, 4.02)	
Global P-value					0.08						0.95	
Age at menarche/year												
≤14 (9-14)	125	50.2	55	54.5	1.00 (Ref)		29	42.0	12	57.1	1.00 (Ref)	
≥15 (15-20)	124	49.8	46	45.5	0.90 (0.55, 1.49)	0.69	40	58.0	9	42.9	0.47 (0.15, 1.50)	
Age at first pregnancy/year												
Nulliparous ^a or ≥30	44	17.7	12	11.9	0.53 (0.25, 1.13)	0.10	7	8.9	4	17.4	1.57 (0.33, 7.53)	
<25	138	55.4	66	65.3	1.00 (Ref)		53	67.1	17	73.9	1.00 (Ref)	
≥25-29	67	26.9	23	22.8	0.60 (0.33, 1.08)	0.09	19	24.1	2	8.7	0.54 (0.11, 2.80)	
Global P-value					0.10						0.60	
Parity												
Nulliparous ^a	23	9.2	3	3.0	(Results not shown)		4	5.0	2	8.0	(Results not shown)	
Parous	226	90.8	98	97.0	1.00 (Ref)		76	95.0	23	92.0	1.00 (Ref)	
Number of children												
1-3	145	64.2	54	55.1	1.00 (Ref)		38	50.0	6	26.1	1.00 (Ref)	
4 or 5	81	35.8	44	44.9	1.55 (0.90, 2.67)	0.11	38	50.0	17	73.9	3.30 (0.93, 11.76)	
Cumulative breastfeeding duration /month												
T1: 1 - <48	81	36.3	22	22.7	1.00 (Ref)		32	43.2	4	20.0	1.00 (Ref)	
T2: 48 - <81	95	42.6	47	48.5	1.77 (0.95, 3.31)	0.07	26	35.1	6	30.0	1.66 (0.34, 8.00)	
T3: 81 - 720	47	21.1	28	28.9	2.00 (0.96, 4.18)	0.06	16	21.6	10	50.0	4.54 (0.97, 21.21)	

	AKU (n=350, 42%)						Kijabe (n=105, 13%)					
	ER+		ER-		ER- vs. ER+		ER+		ER-		ER- vs. ER+	
Key risk factors	N	%	N	%	OR (95% CI) †	P†	N	%	N	%	OR (95% CI) †	
Global P-value	0.12						0.12					
P-trend							2.19 (1.00, 4.78)					
Mean breastfeeding duration per child/month												
<24	142	63.7	64	66.0	1.00 (Ref)		55	74.3	14	70.0	1.00 (Ref)	
≥ 24	81	36.3	33	34.0	1.03 (0.60, 1.76)	0.92	19	25.7	6	30.0	0.91 (0.25, 3.27)	
Age at first pregnancy & Number of children												
Age 25+ yr, 1-3 births	74	32.7	24	24.5	1.00 (Ref)		14	18.7	2	9.1	1.00 (Ref)	
Age <25 yr, 1-3 births	71	31.4	30	30.6	1.50 (0.77, 2.96)	0.24	23	30.7	4	18.2	0.46 (0.06, 3.86)	
Age 25+ yr, 4+ births	15	6.6	8	8.2	1.78 (0.65, 4.92)	0.27	9	12.0	3	13.6	1.35 (0.15, 12.46)	
Age <25 yr, 4+ births	66	29.2	36	36.7	1.93 (0.99, 3.80)	0.06	29	38.7	13	59.1	2.17 (0.37, 12.56)	
Global P-value	0.28						0.31					

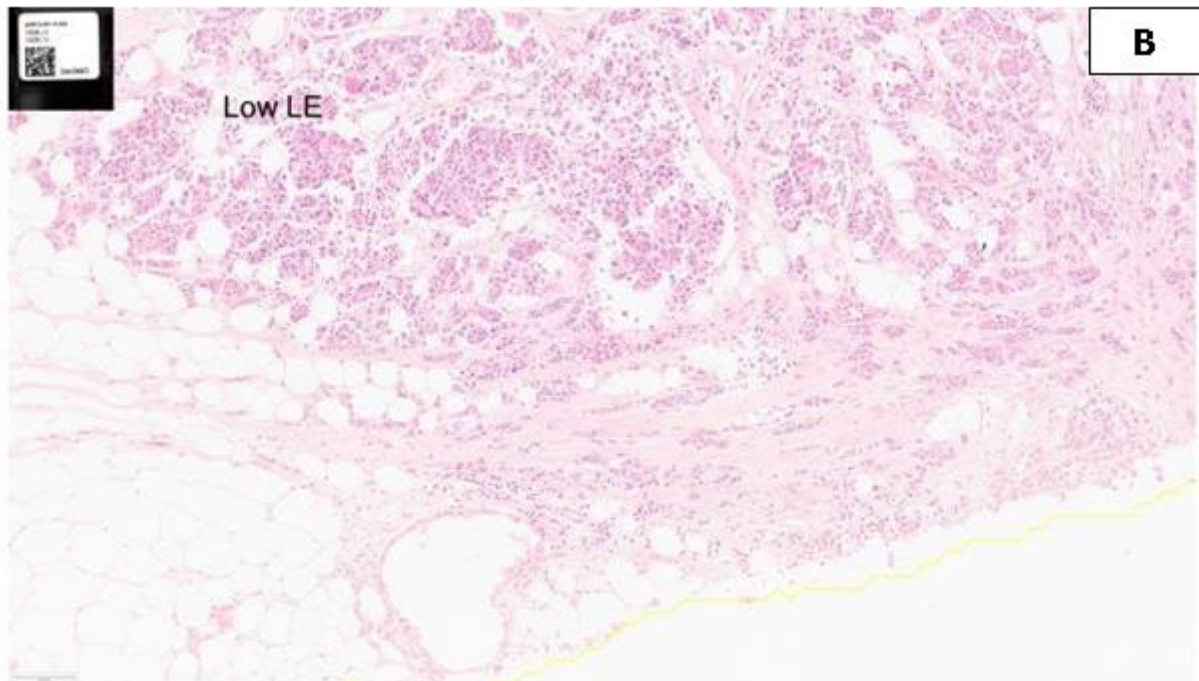
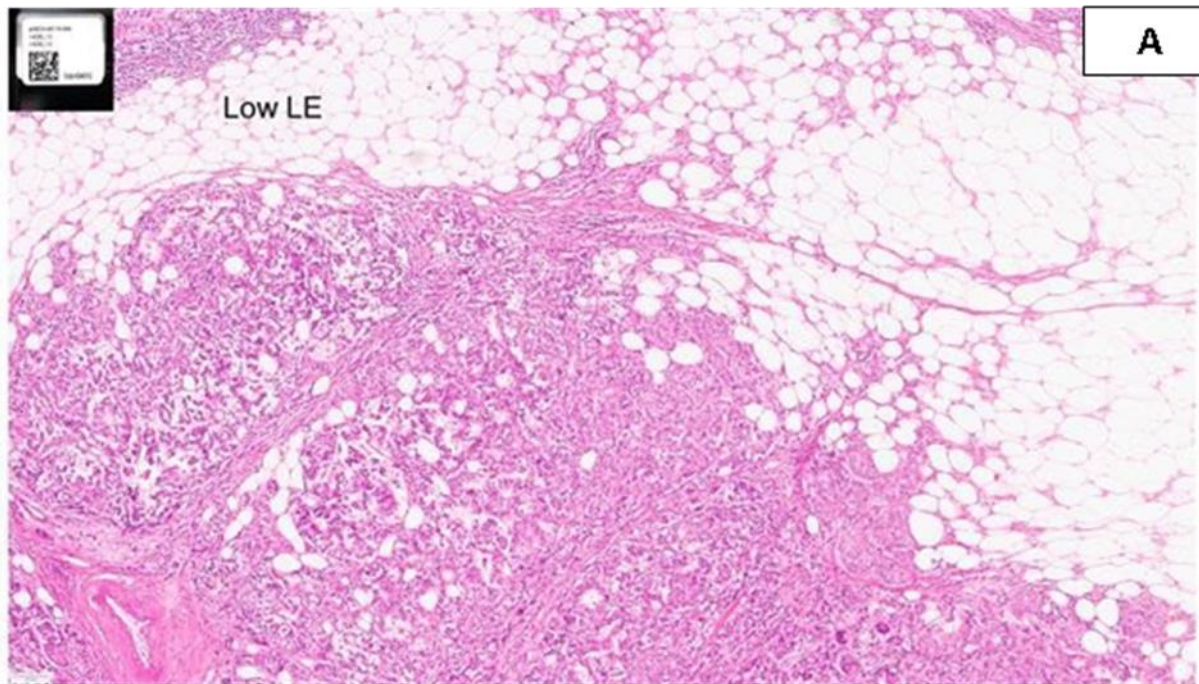
Key risk factors	Nyeri (n=110, 13%)						St Mary's (n=122, 15%)					
	ER+		ER-		ER- vs. ER+		ER+		ER-		ER- vs. ER+	
	N	%	N	%	OR (95% CI) †	P†	N	%	N	%	OR (95% CI) †	P†
BMI												
Normal (<25.0)	12	36.4	13	43.3	1.00 (Ref)		28	43.1	8	33.3	1.00 (Ref)	
Overweight (25.0 - 29.9)	15	45.5	13	43.3	0.86 (0.27, 2.68)	0.79	25	38.5	12	50.0	1.74 (0.61, 5.00)	0.30
Obese (≥30.0)	6	18.2	4	13.3	0.71 (0.15, 3.40)	0.67	12	18.5	4	16.7	1.16 (0.29, 4.68)	0.84
Global P-value					0.91						0.56	
Age at menarche/year												
≤14 (9-14)	27	45.0	19	45.2	1.00 (Ref)		52	59.1	16	50.0	1.00 (Ref)	
≥15 (15-20)	33	55.0	23	54.8	2.68 (0.75, 9.59)	0.13	36	40.9	16	50.0	2.96 (1.08, 8.12)	0.036
Age at first pregnancy/year												
Nulliparous ^a or ≥30	4	6.5	1	2.2	1.43 (0.07, 28.47)	0.82	4	4.9	3	10.0	2.75 (0.46, 16.55)	0.27
<25	50	80.6	38	84.4	1.00 (Ref)		64	79.0	26	86.7	1.00 (Ref)	
≥25-29	8	12.9	6	13.3	5.51 (0.50, 60.64)	0.16	13	16.0	1	3.3	0.32 (0.04, 2.80)	0.30
Global P-value					0.37						0.30	
Parity												
Nulliparous ^a	2	3.1	0	0.0	(Results not shown)		1	1.1	1	3.1	(Results not shown)	
Parous	63	96.9	45	100.0	1.00 (Ref)		89	98.9	31	96.9	1.00 (Ref)	
Number of children												
1-3	30	47.6	23	51.1	1.00 (Ref)		41	46.1	9	29.0	1.00 (Ref)	
4 or 5	33	52.4	22	48.9	0.82 (0.28, 2.43)	0.72	48	53.9	22	71.0	2.88 (0.86, 9.64)	0.09
Cumulative breastfeeding duration /month												

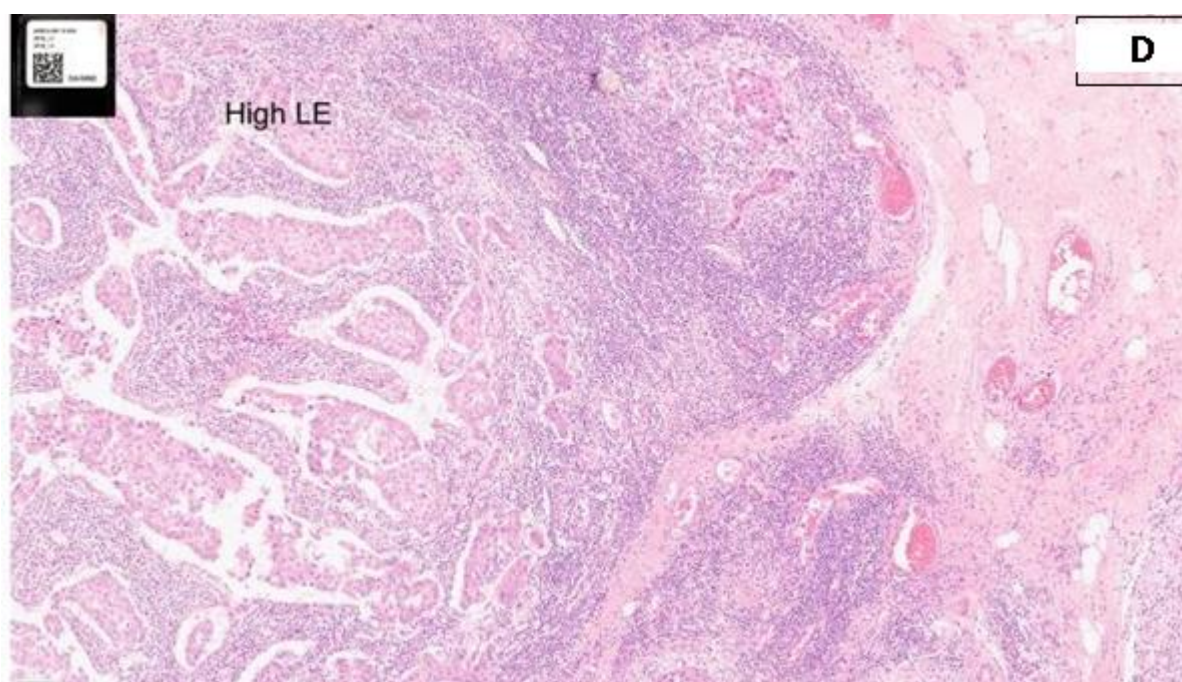
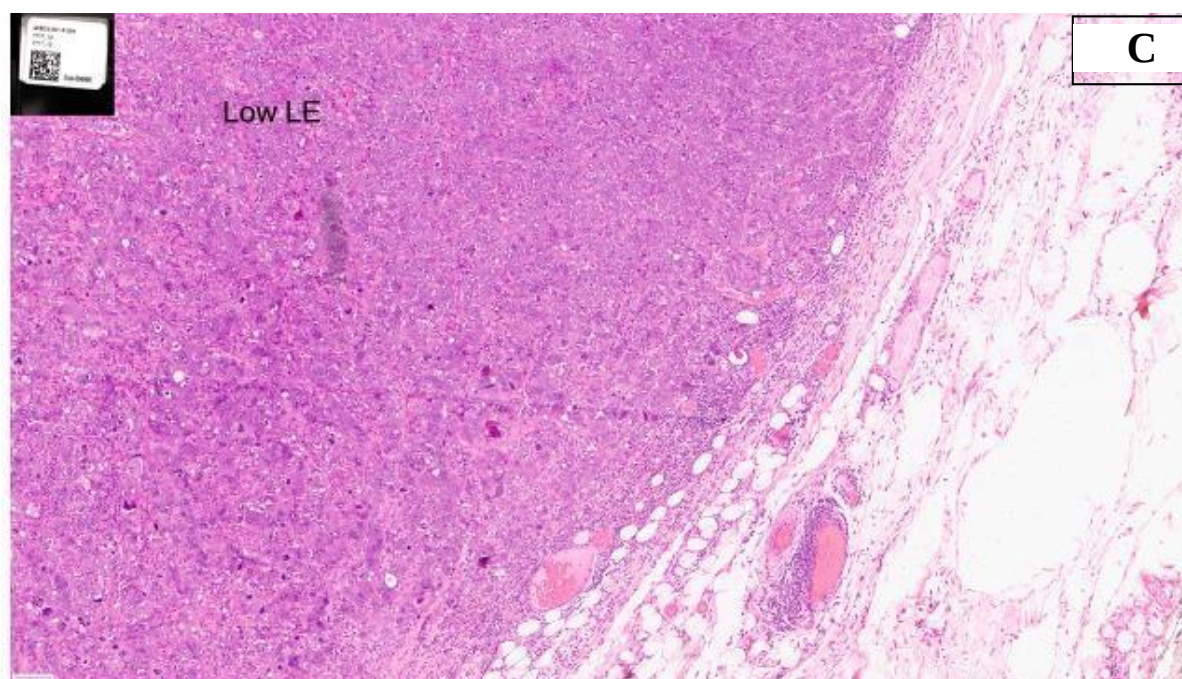
	Nyeri (n=110, 13%)						St Mary's (n=122, 15%)					
	ER+		ER-		ER- vs. ER+		ER+		ER-		ER- vs. ER+	
Key risk factors	N	%	N	%	OR (95% CI) †	P†	N	%	N	%	OR (95% CI) †	P†
T1: 1 - <48	16	26.2	13	28.9	1.00 (Ref)		31	37.8	6	19.4	1.00 (Ref)	
T2: 48 - <81	20	32.8	16	35.6	1.32 (0.34, 5.18)	0.69	19	23.2	7	22.6	1.80 (0.43, 7.63)	0.42
T3: 81 - 720	25	41.0	16	35.6	1.17 (0.28, 4.93)	0.83	32	39.0	18	58.1	2.60 (0.73, 9.35)	0.14
Global P-value						0.92						0.34
P-trend												
Mean breastfeeding duration per child/month												
<24	45	73.8	34	75.6	1.00 (Ref)		60	73.2	22	71.0	1.00 (Ref)	
≥ 24	16	26.2	11	24.4	1.67 (0.41, 6.77)	0.47	22	26.8	9	29.0	0.91 (0.26, 3.20)	0.88
Age at first pregnancy & Number of children												
Age 25+ yr, 1-3 births	8	13.3	5	11.1	1.00 (Ref)		12	15.0	1	3.4	1.00 (Ref)	
Age <25 yr, 1-3 births	20	33.3	18	40.0	0.27 (0.02, 3.05)	0.29	29	36.3	8	27.6	2.38 (0.24, 23.83)	0.46
Age 25+ yr, 4+ births	2	3.3	2	4.4	--		4	5.0	2	6.9	12.02 (0.64, 226.73)	0.10
Age <25 yr, 4+ births	30	50.0	20	44.4	0.23 (0.02, 2.59)	0.23	35	43.8	18	62.1	5.18 (0.55, 48.72)	0.15
Global P-value						0.70						0.28

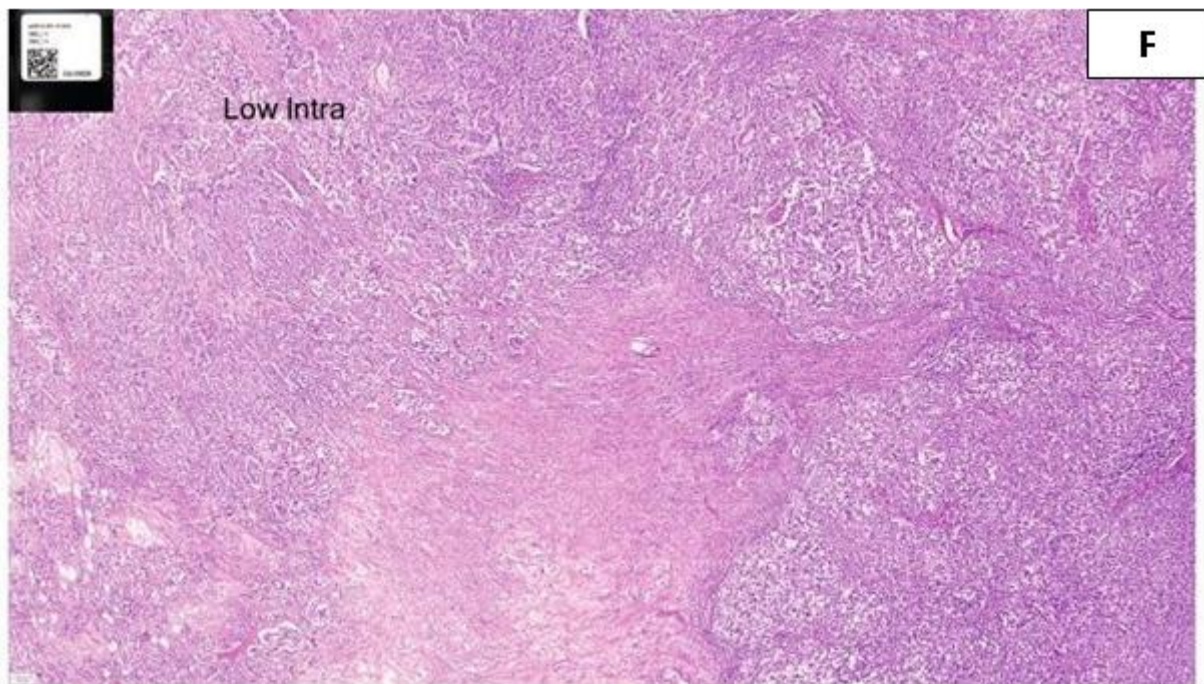
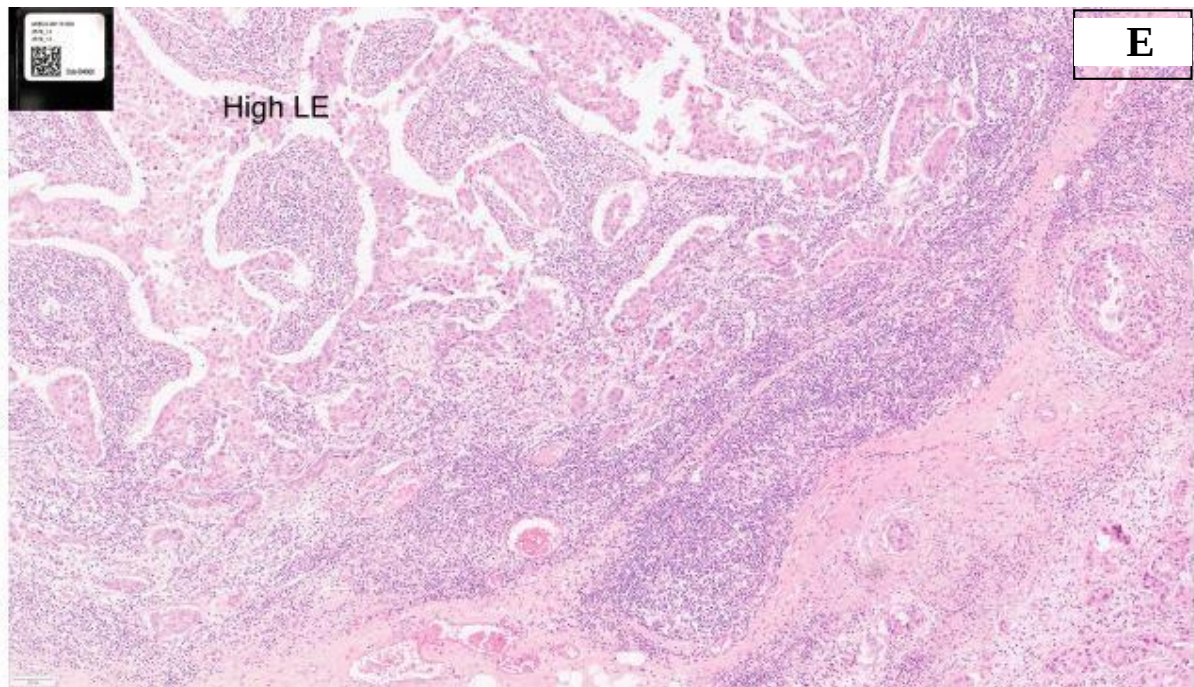
	Others (n=151, 18%)					
	ER+		ER-		ER- vs. ER+	
Key risk factors	N	%	N	%	OR (95% CI) †	P†
BMI						
Normal (<25.0)	27	38.0	20	43.5	1.00 (Ref)	
Overweight (25.0 - 29.9)	24	33.8	16	34.8	0.91 (0.37, 2.24)	0.84
Obese (≥30.0)	20	28.2	10	21.7	0.68 (0.26, 1.82)	0.44
Global P-value					0.74	
Age at menarche/year						
≤14 (9-14)	47	52.2	26	52.0	1.00 (Ref)	
≥15 (15-20)	43	47.8	24	48.0	1.02 (0.47, 2.22)	0.96
Age at first pregnancy/year						
Nulliparous ^a or ≥30	11	11.3	3	5.7	0.42 (0.11, 1.63)	0.21
<25	78	80.4	47	88.7	1.00 (Ref)	
≥25-29	8	8.2	3	5.7	0.87 (0.19, 3.90)	0.85
Global P-value						
Parity						
Nulliparous ^a	7	7.1	1	1.9	(Results not shown)	
Parous	91	92.9	52	98.1	1.00 (Ref)	
Number of children						
1-3	39	42.9	16	30.8	1.00 (Ref)	
4 or 5	52	57.1	36	69.2	1.59 (0.69, 3.67)	0.28
Cumulative breastfeeding duration /month						
T1: 1 - <48	27	31.0	8	16.7	1.00 (Ref)	
T2: 48 - <81	25	28.7	12	25.0	1.35 (0.41, 4.40)	0.62
T3: 81 - 720	35	40.2	28	58.3	2.93 (0.97, 8.83)	0.056
Global P-value					0.45	
P-trend					1.76 (1.01, 3.04)	0.044

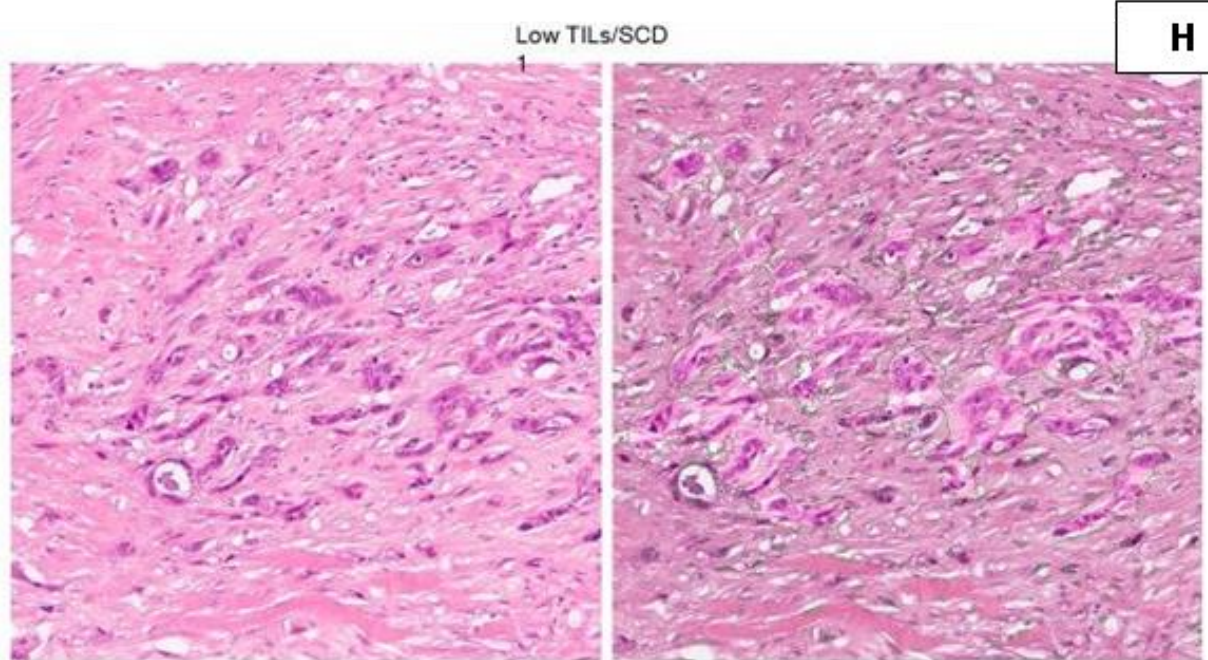
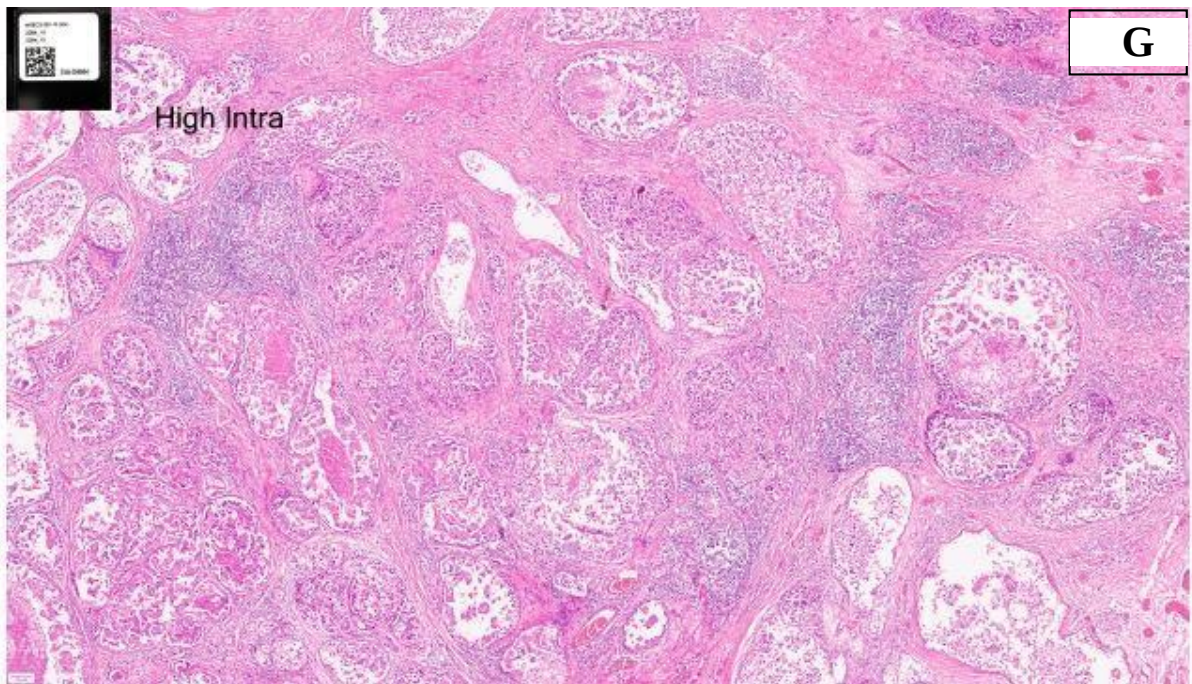
	Others (n=151, 18%)					
	ER+		ER-		ER- vs. ER+	
Key risk factors	N	%	N	%	OR (95% CI) †	P†
Mean breastfeeding duration per child/month						
<24	59	67.8	32	66.7	1.00 (Ref)	
≥ 24	28	32.2	16	33.3	1.24 (0.50, 3.08)	0.64
Age at first pregnancy & Number of children						
Age 25+ yr, 1-3 births	9	10.0	5	9.6	1.00 (Ref)	
Age <25 yr, 1-3 births	30	33.3	11	21.2	0.63 (0.15, 2.56)	0.52
Age 25+ yr, 4+ births	4	4.4	0	0	--	--
Age <25 yr, 4+ births	47	52.2	36	69.2	1.22 (0.35, 4.26)	0.76
Global P-value	0.61					

APPENDIX III: TIL H &E-Stained Images (Leading Edge, Intratumoural)

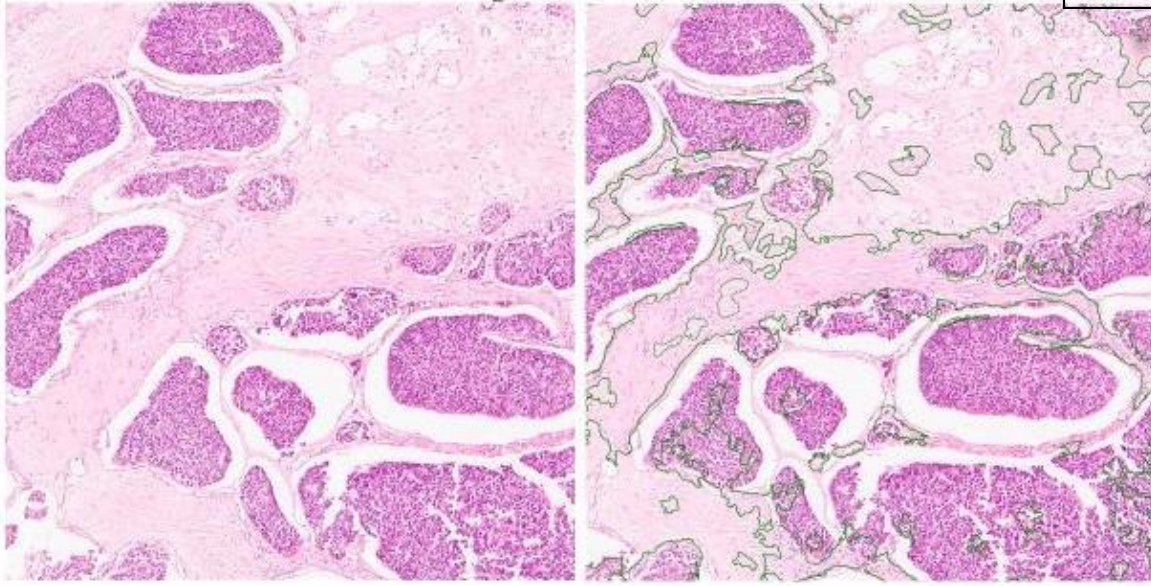






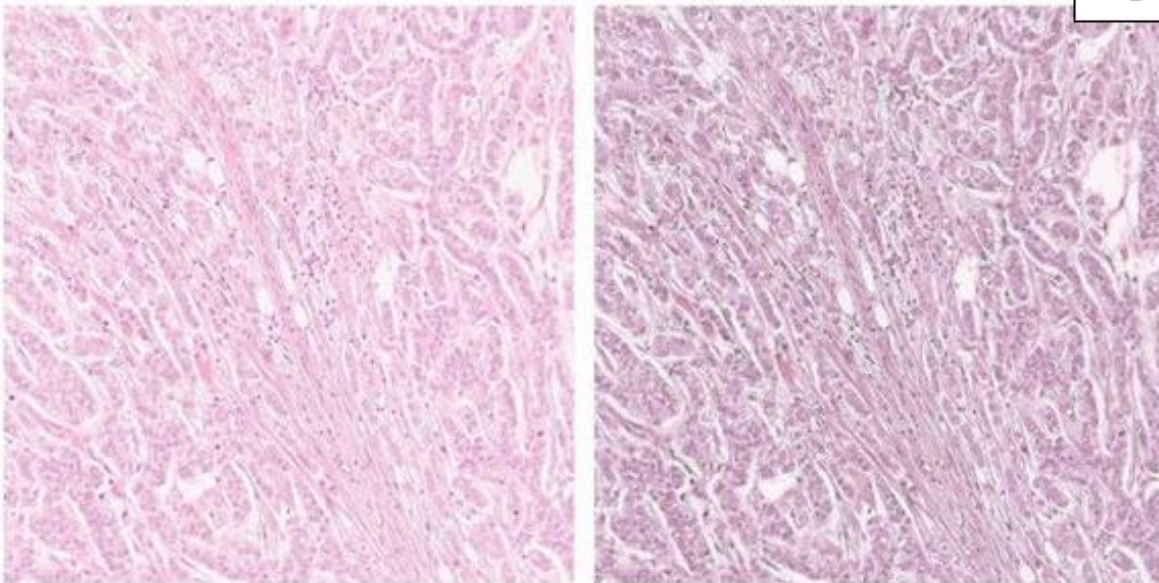


Low TILs/SCD
2



I

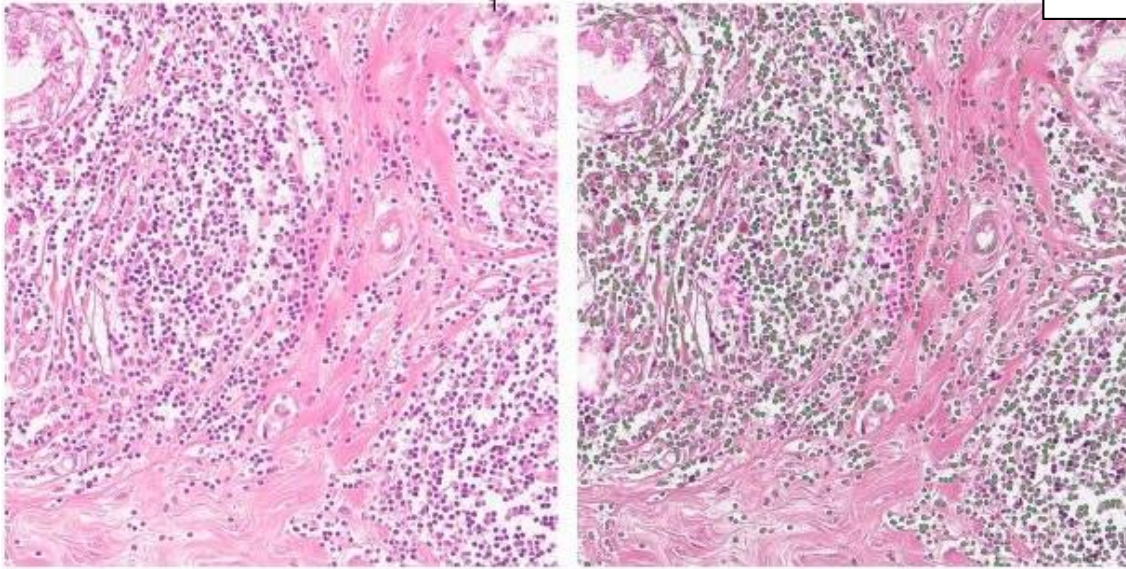
Low intra TILs/SCD 3



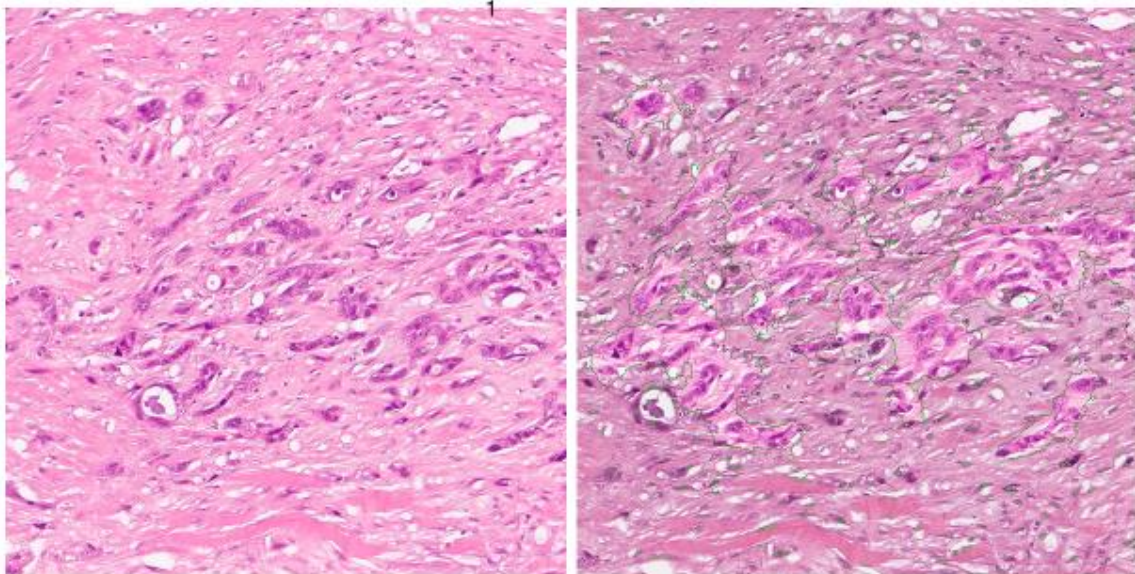
J

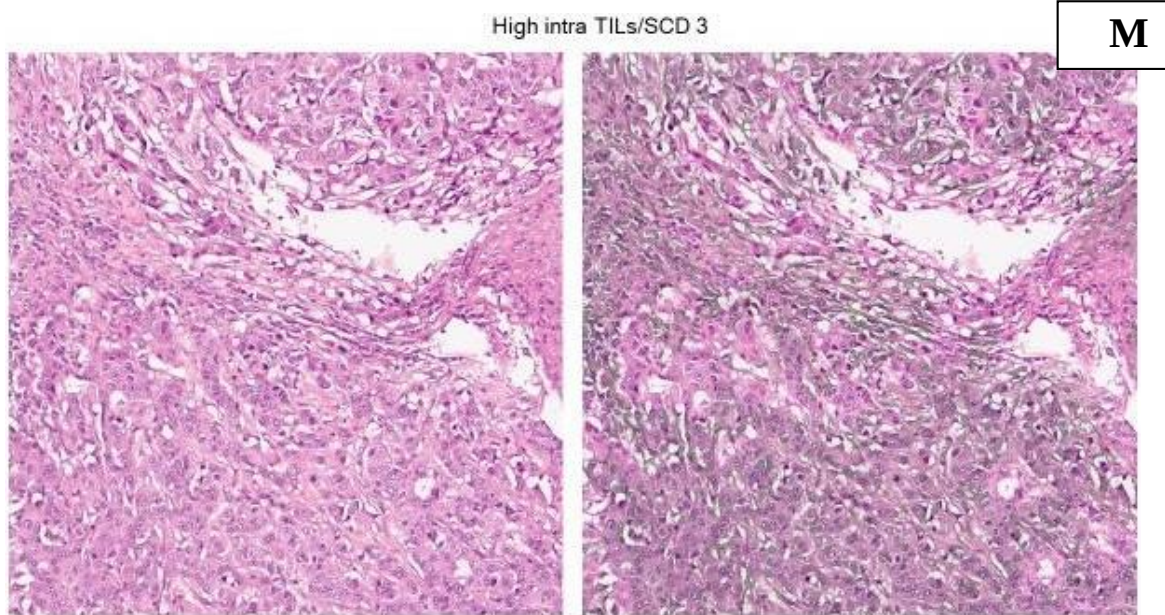
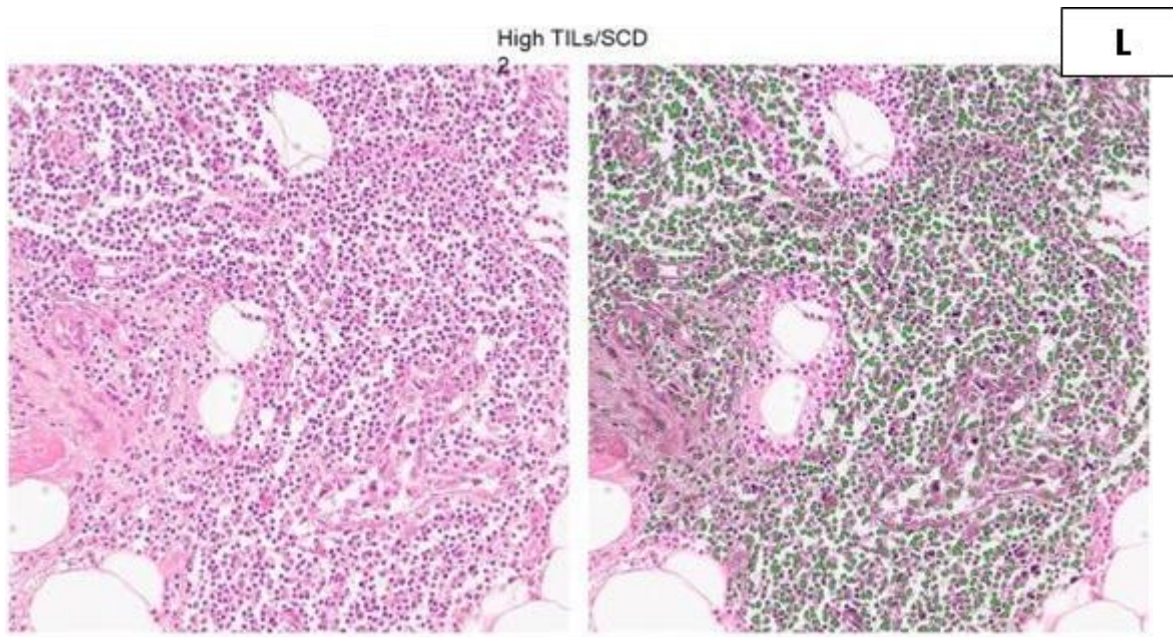
K

High TILs/SCD



Low TILs/SCD





Images A-G: Visual TIL estimation in breast cancer H&E stained whole tissue sections showing Low and High Leading edge TIL(A-C); High LE (D-E); low intratumoural TIL(F); High intratumoural TIL(G)

Images H-M: Supervised machine learning for tissue segmentation and cell detection in breast cancer H&E-stained whole-tissue sections showing Low intratumoural TIL/SCD (H-J) and High LE TILs (K-M)

APPENDIX IV: Tissue Array Map Template

	1 mm TMA		AKU 1-1													
Y X	1	2	3	4	5	6	7	8	9	10	11	12	14	15	16	17
1	1	1	1	1	2	2	2	2	3	3	3	3				
2	4	4	4	4	5	5	5	5	6	6	6	6				
3	8	8	8	8	9	9	9	9	10	10	10	10				
4	11	11	11	11	12	12	12	12	13	13	13	13				
5	14	14	14	14	15	15	15	15	16	16	16	16				
6	17	17	17	17	20	20	20	20	21	21	21	21				
7	25	25	25	25	C LIVER	C LIVER	C KIDNEY	C KIDNEY								
8																
9																
10																
11																
12																
Date:12/2/2016			For: Dr. Shahina Syed and Dr. Rosen- Double spaced.													
Built By: Shahida Salar																

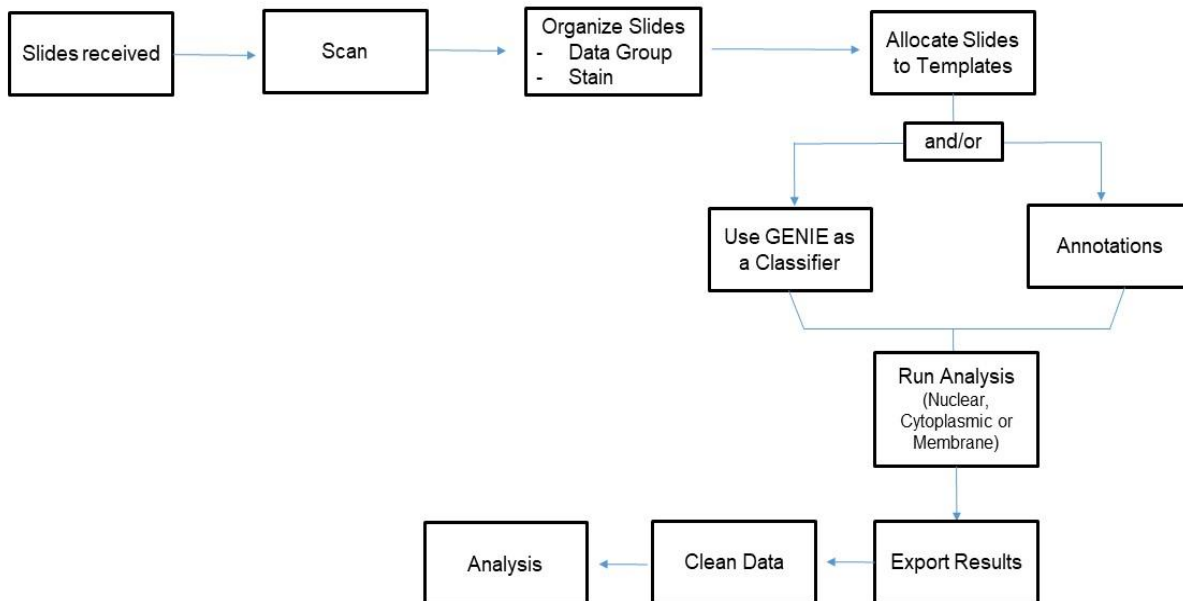
KEY	
Numbers in black	Represent 2 cores from intra-tumoural stroma
Numbers in blue	Represent 2 cores from the leading-edge stroma
Cores labelled C	Represent liver and kidney tissue controls.

TMA MAP

	1 mm TMA		AKU 1-1									
Y X	1	2	3	4	5	6	7	8	9	10	11	12
1	1	1	1	1	2	2	2	2	3	3	3	3
2	4	4	4	4	5	5	5	5	6	6	6	6
3	8	8	8	8	9	9	9	9	10	10	10	10
4	11	11	11	11	12	12	12	12	13	13	13	13
5	14	14	14	14	15	15	15	15	16	16	16	16
6	17	17	17	17	20	20	20	20	21	21	21	21
7	25	25	25	25	C LIVER	C LIVER	C KIDNEY	C KIDNEY				

APPENDIX V: Aperio-Image Scope Digital Analysis Process

Overview



Scanner: Aperio Versa AT2

<https://www.leicabiosystems.com/digital-pathology/scan/aperio-at2>

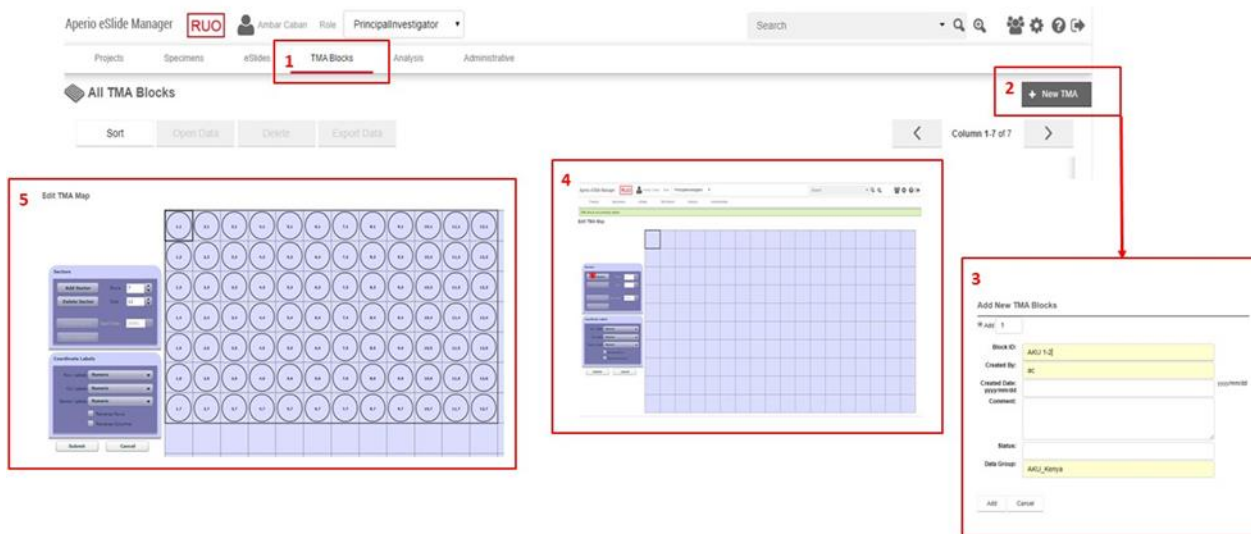
Software: Aperio eSlide Manager

https://drp8p5tqcb2p5.cloudfront.net/fileadmin/biosystems/PDF/Aperio_eSlide_Manager_-_RUO.pdf

Application: Aperio Image Scope

https://www.leicabiosystems.com/fileadmin/biosystems/PDF/Aperio_ImageScope_User_Guide.pdf

Creating a Template



Click on “TMA Blocks” > select “All TMA Blocks (As List)”

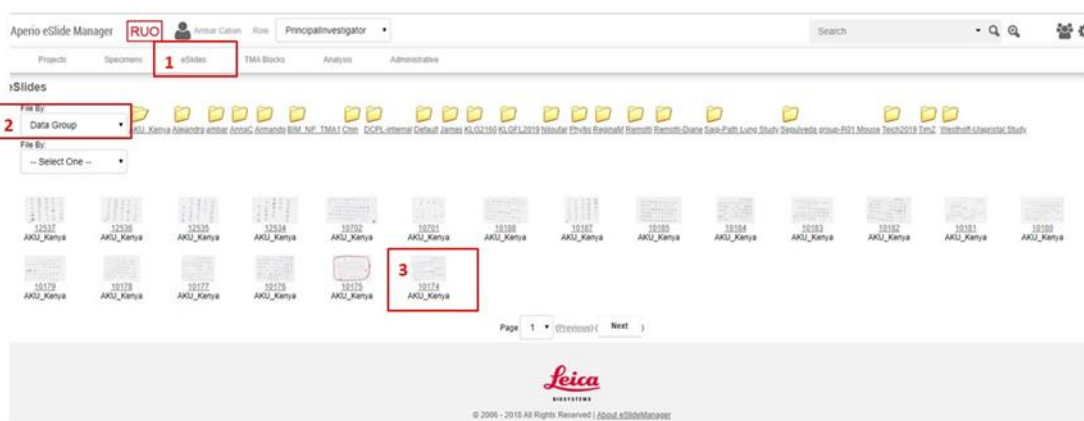
Click on “New TMA”

Add information and save it by clicking “Add”

What you will see after clicking “Add” on the previous step (#3), to start creating the template press “Add sector”, then, you will be able to add columns and rows by increasing the number.

Example AKU 1-2

Allocate slide to template

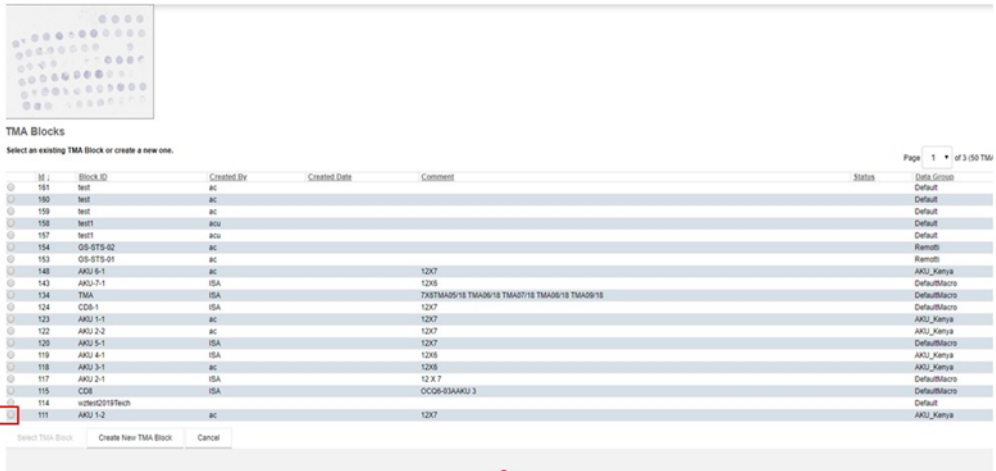


Click on “eSlides” > “All eSlides (As Folders)”

Filter by: Data Group, select “AKU_Kenya”

Select slide of interest (ex. 10174)

Allocate slide to template



TMA Blocks
Select an existing TMA Block or create a new one.

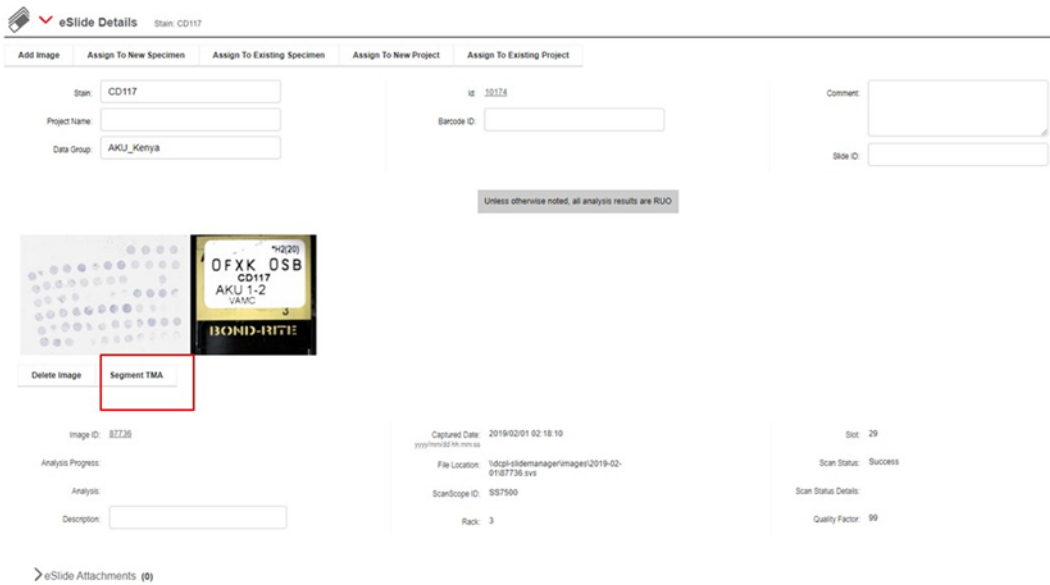
Page: 1 of 3 (50 TMA)

ID	Block ID	Created By	Created Date	Comment	Status	Data Group
151	test	AC				Default
150	test	AC				Default
159	test	AC				Default
158	test1	ACU				Default
157	test1	ACU				Default
154	OS-815-02	AC				Remotti
153	OS-815-01	AC				Remotti
148	AKU 5-1	AC	12X7			AKU_Kenya
143	AKU 7-1	ISA	12X8			DefaultMacro
134	TMA	ISA	7X8TMA05/18 TMA06/18 TMA07/18 TMA08/18 TMA09/18			DefaultMacro
124	CD8-1	ISA	12X7			DefaultMacro
123	AKU 1-1	AC	12X7			AKU_Kenya
122	AKU 2-2	AC	12X7			AKU_Kenya
120	AKU 5-1	ISA	12X7			DefaultMacro
119	AKU 4-1	ISA	12X8			DefaultMacro
118	AKU 3-1	AC	12X8			AKU_Kenya
117	AKU 2-1	ISA	12 X 7			DefaultMacro
115	CD8	ISA	OC08-03AAKU 3			DefaultMacro
114	cd8cd191Tech					Default
111	AKU 1-2	AC	12X7			AKU_Kenya

Select TMA Block Create New TMA Block Cancel

Click on “eSlides” > “All eSlides (As Folders)”
Filter by: Data Group, select “AKU_Kenya”
Select slide of interest (ex. 10174)

Allocate slide to template



eSlide Details Stain: CD117

Assign To New Specimen Assign To Existing Specimen Assign To New Project Assign To Existing Project

Stain: CD117
Project Name:
Data Group: AKU_Kenya

Id: 10174
Barcode ID:
Comment:
Slide ID:

Unless otherwise noted, all analysis results are RUQ

Image ID: 87738
Analysis Progress:
Analysis:
Description:

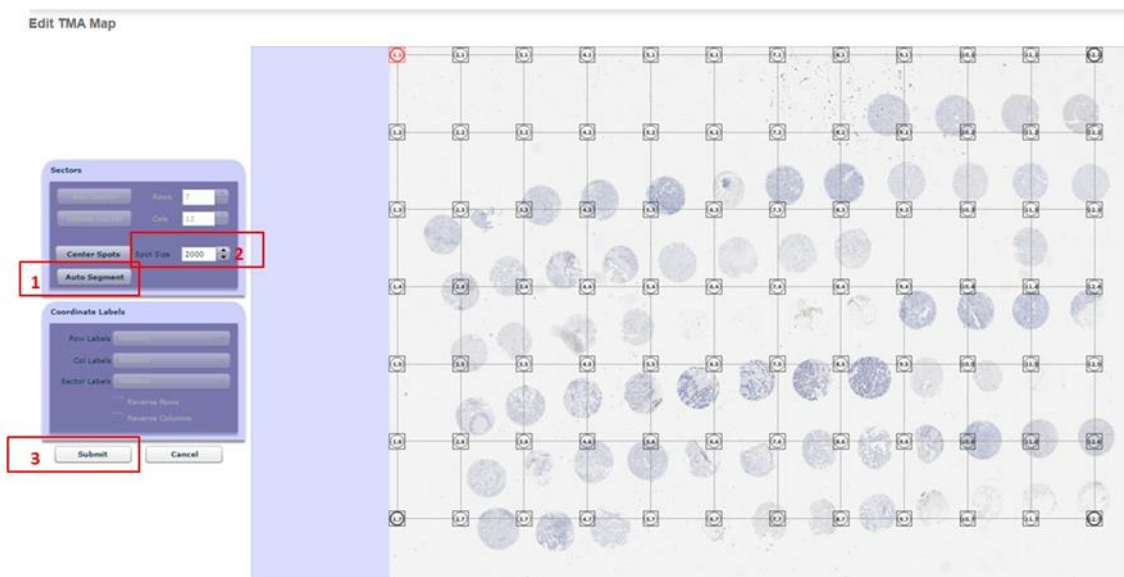
Captured Date: 2019/02/01 02:18:10
File Location: /tmp/slidesmanager/images/2019-02-01/87738.png
ScanScope ID: SS7599
Rack: 3

Slot: 29
Scan Status: Success
Scan Status Details:
Quality Factor: 99

> eSlide Attachments (0)

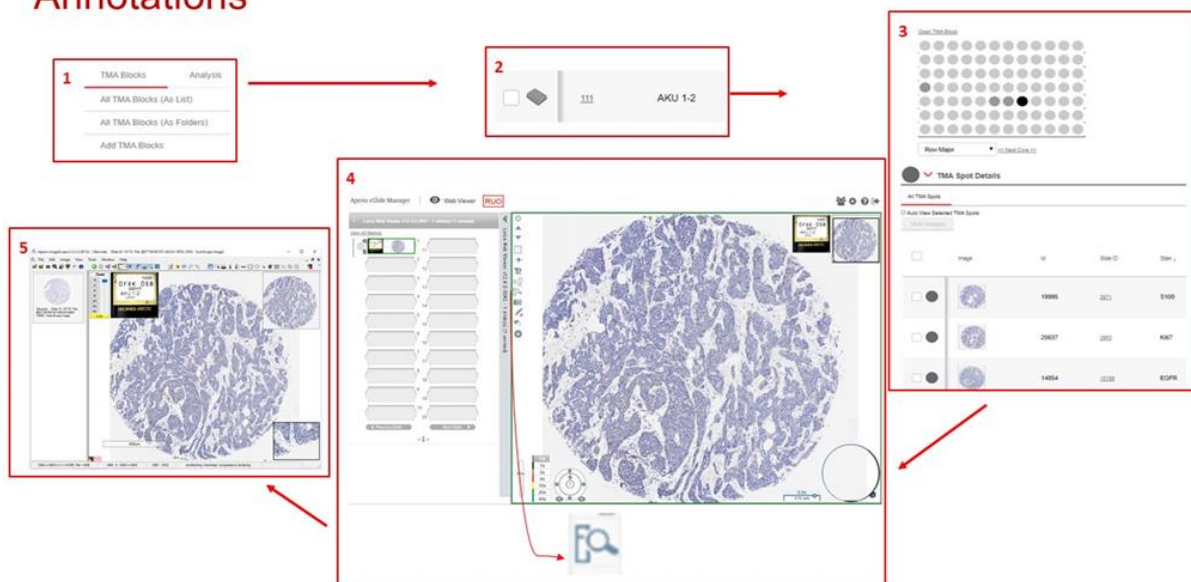
Select TMA Block of interest, by clicking on the circle close to ID
If you don't have a TMA block you can create a new one by clicking on “Create New TMA Block”

Slide Segmentation



1. Click “Auto Segment”, then, adjust (by clicking on the spot and moving it).
2. You can increase or decrease the spot in the “Spot Size” section.
3. When finishing click “Submit”

Annotations



In the main menu select: TMA Blocks > All TMA Blocks (As List)
 Select block of interest by clicking on the number (ex "111" for AKU 1-2) or the block icon.
 Click any spot and you will be able to see the cores related to that specific spot. To see the image, click on the core of interest in the image section.
 The image will open on Web Viewer, click on the icon that has a slide with a magnifier.
 Image opened on ImageScope, here is where you are going to start doing the annotations.

Annotations

1. Open the image on Image Scope
2. Start selecting region of interest by using any of the following icons  (to exclude an area )
3. Save annotations by clicking on the Annotations icon,  and then save it 



4. Now you can go back and select another spot

Run Analysis

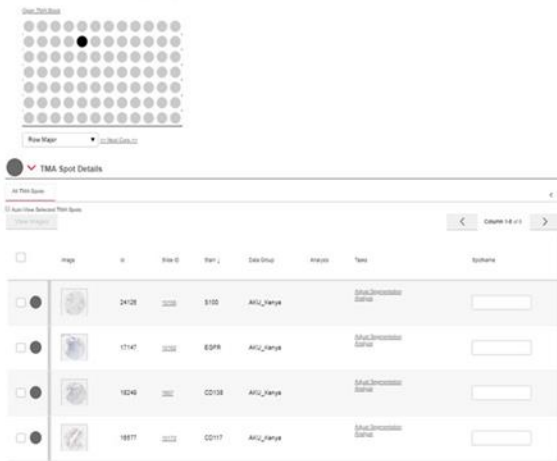
Before running the analysis you have to adjust the algorithm parameters, by tuning.

1. Open the image or core on ImageScope
2. On the toolbar click, View> Analysis
3. Choose Algorithm> Nuclear V9> Default
4. Click on the section of interest (ex. Nuclear identification), and adjust settings
5. Analyze> Tune Algorithm
6. Repeat steps 4 and 5 until you get the results you want
7. Save settings by clicking on save icon
 - a. Select between local or remote (recommended-remote)
 - b. Select data group (AKU_Kenya)

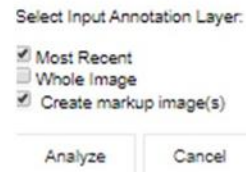


Run Analysis

1. TMA Blocks > All TMA Blocks (As List)
2. Select TMA of interest
3. Click on any spot

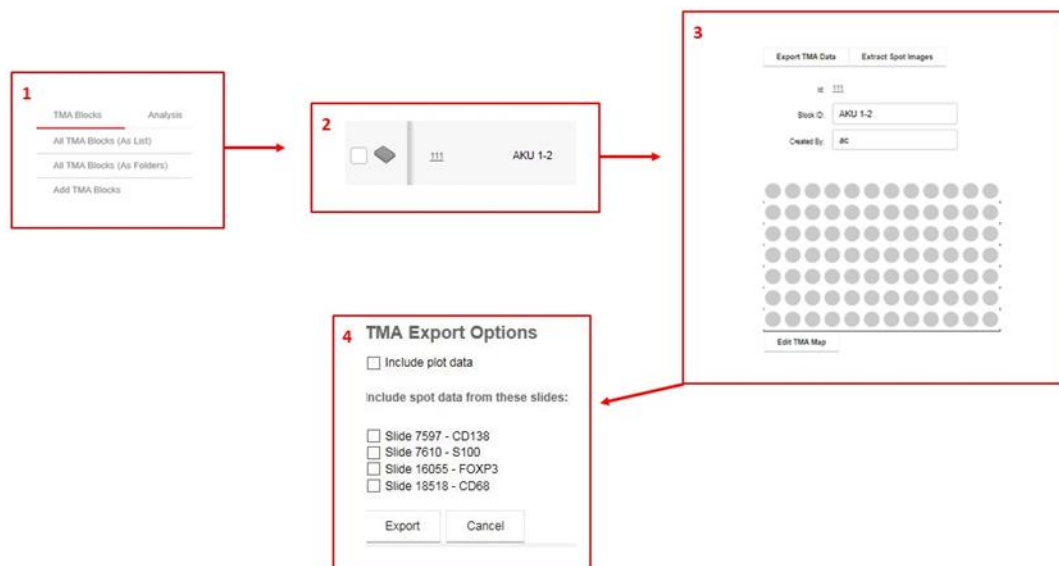


4. If there are several images, search for the one that you are interested in analyzing. Click on "Analyze" and select "This Spot" or "All Spots on Slide"
5. Search for your Macro (ex. Nuclear_Shahin) and select it
6. Click on Analyze



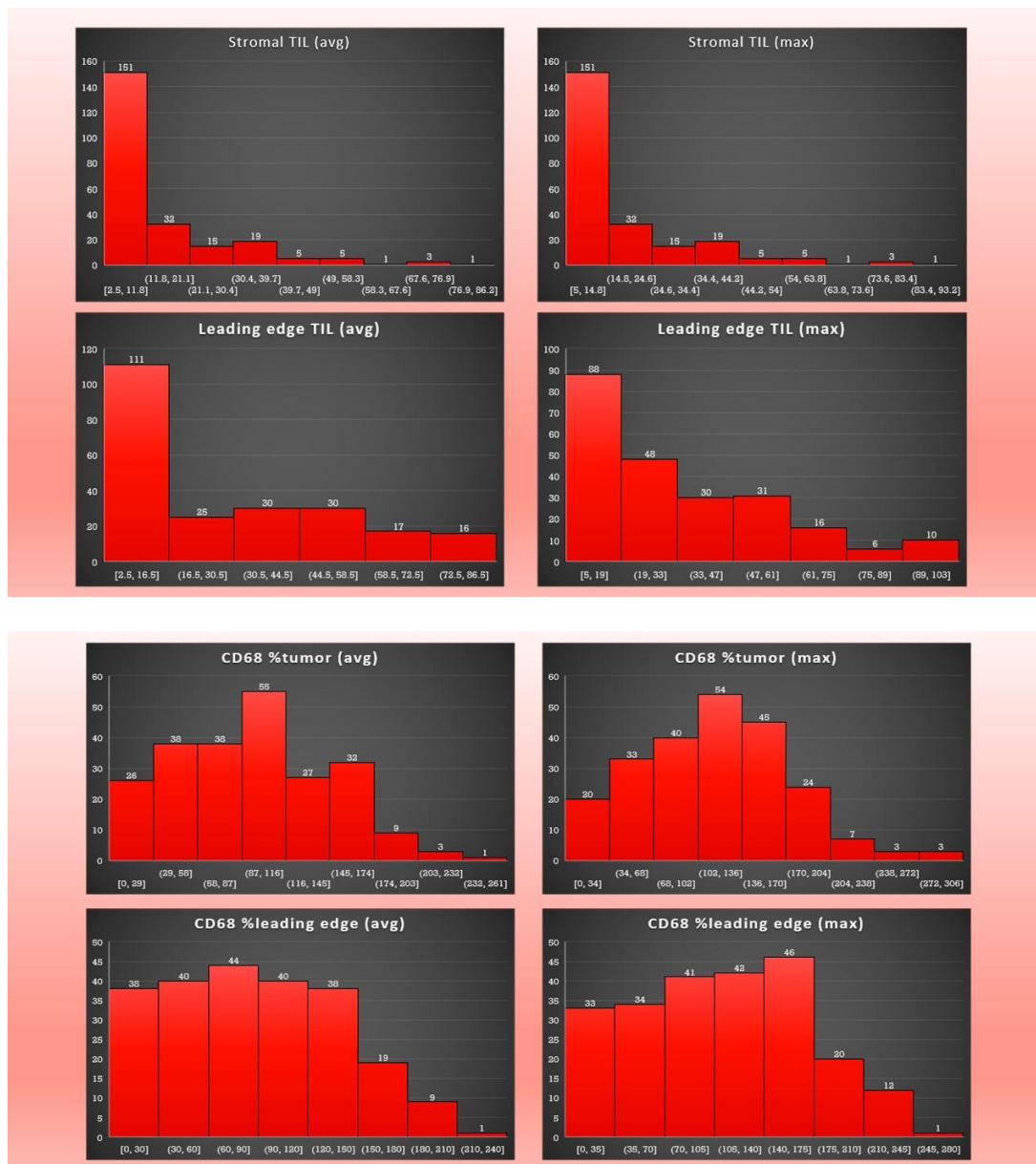
You can see the status of your analysis by going to "Analysis" (on the main menu) and selecting "Jobs"

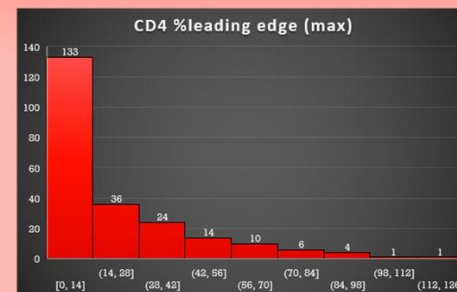
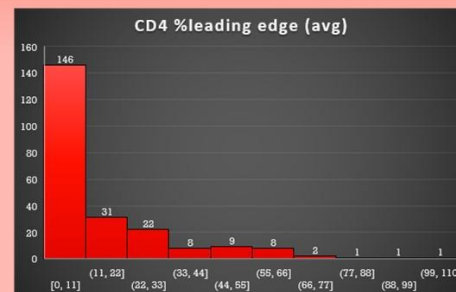
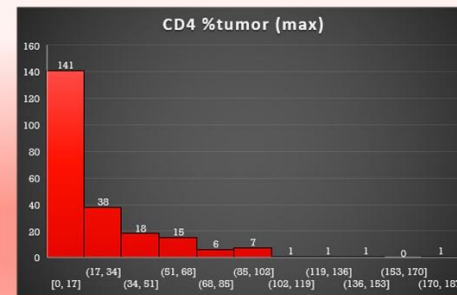
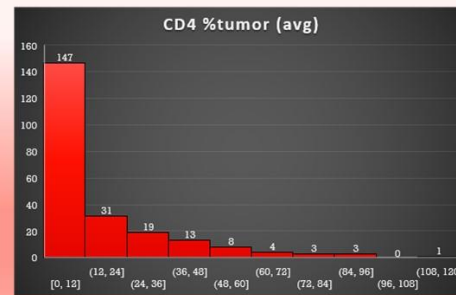
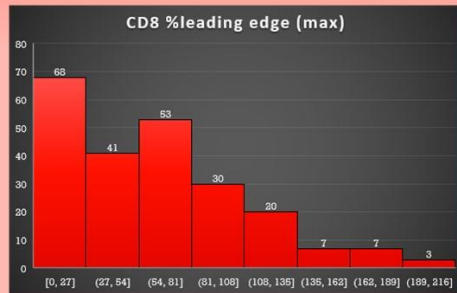
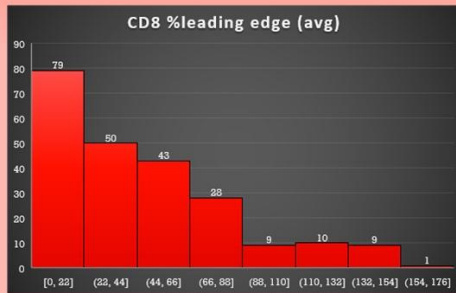
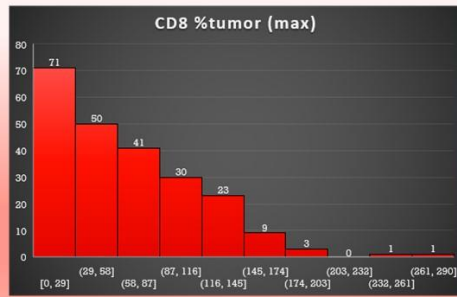
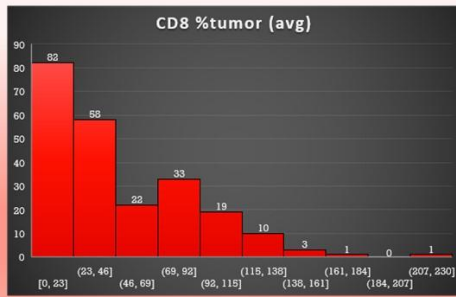
Export Results

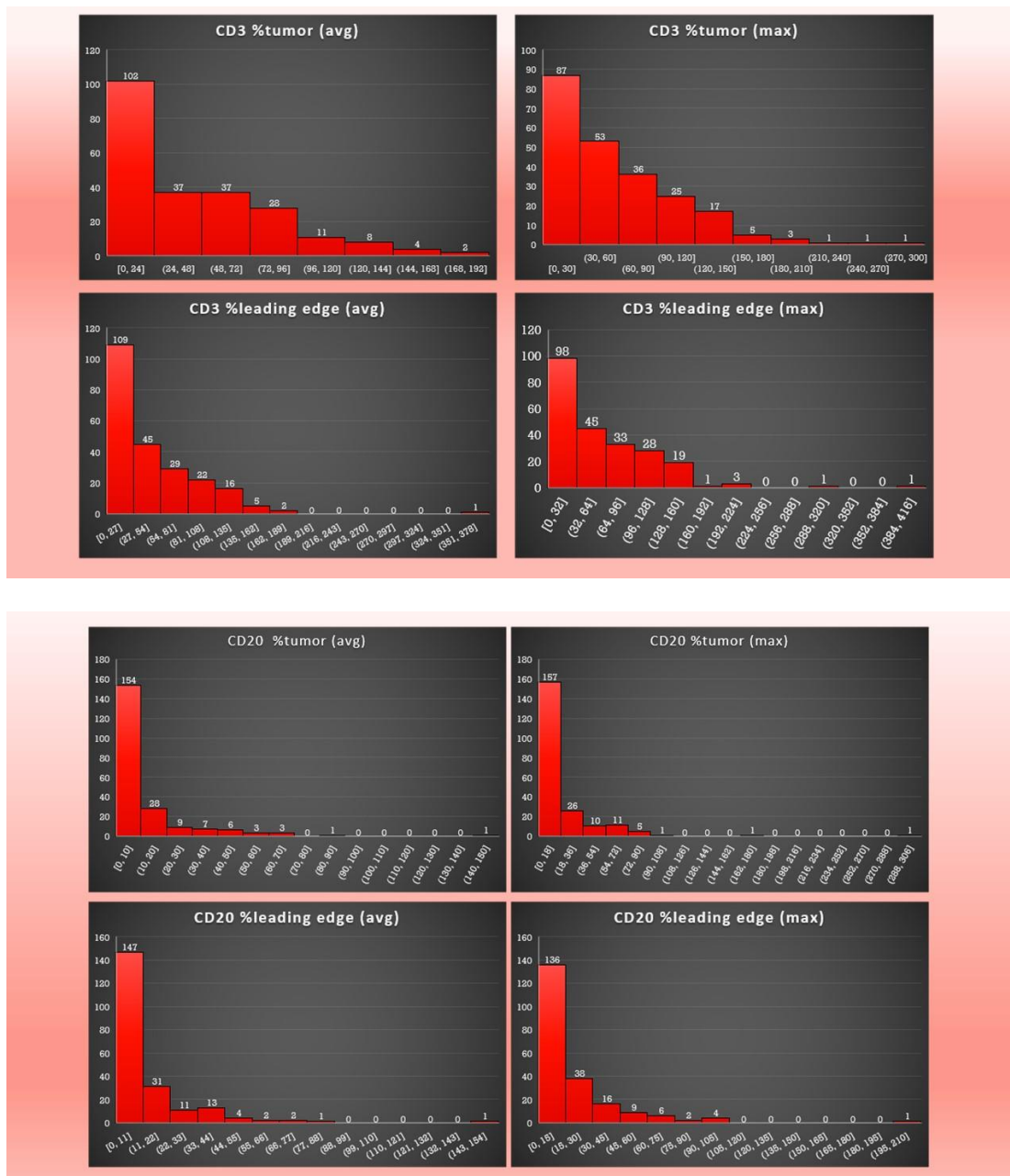


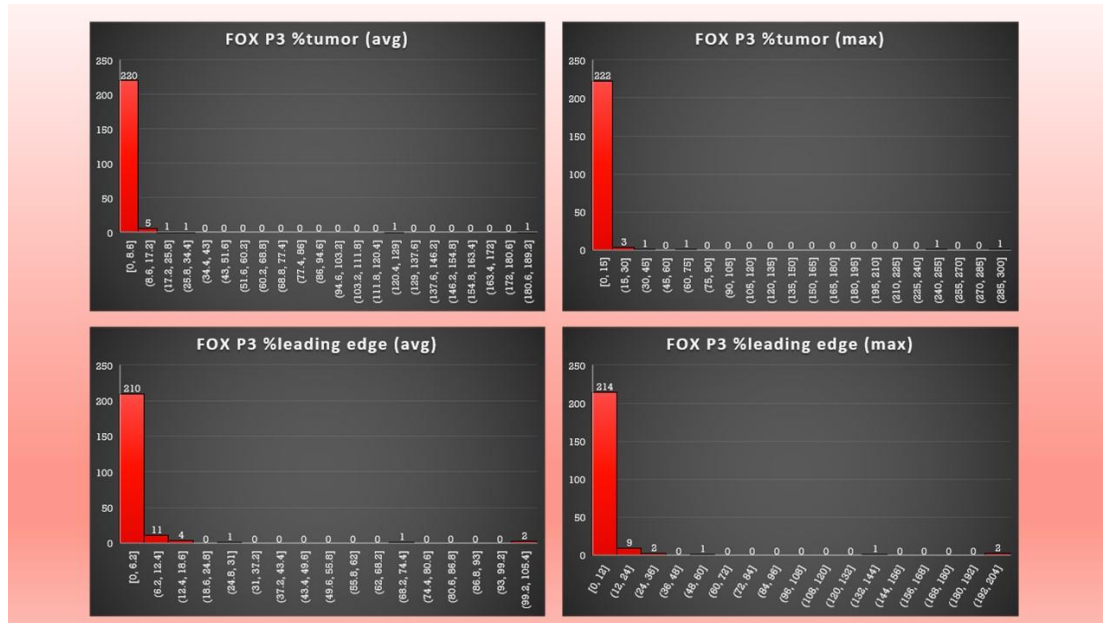
Main Menu on TMA Blocks, select "All TMA Blocks (As List)"
 Select TMA of interest and open it by clicking on the template icon or the number (ex "111")
 Click on "Export TMA Data"
 Select the slide from which you want to export the results, then, click Export

APPENDIX VI: Correlation Graphs of Intratumoral and Leading-Edge Stroma

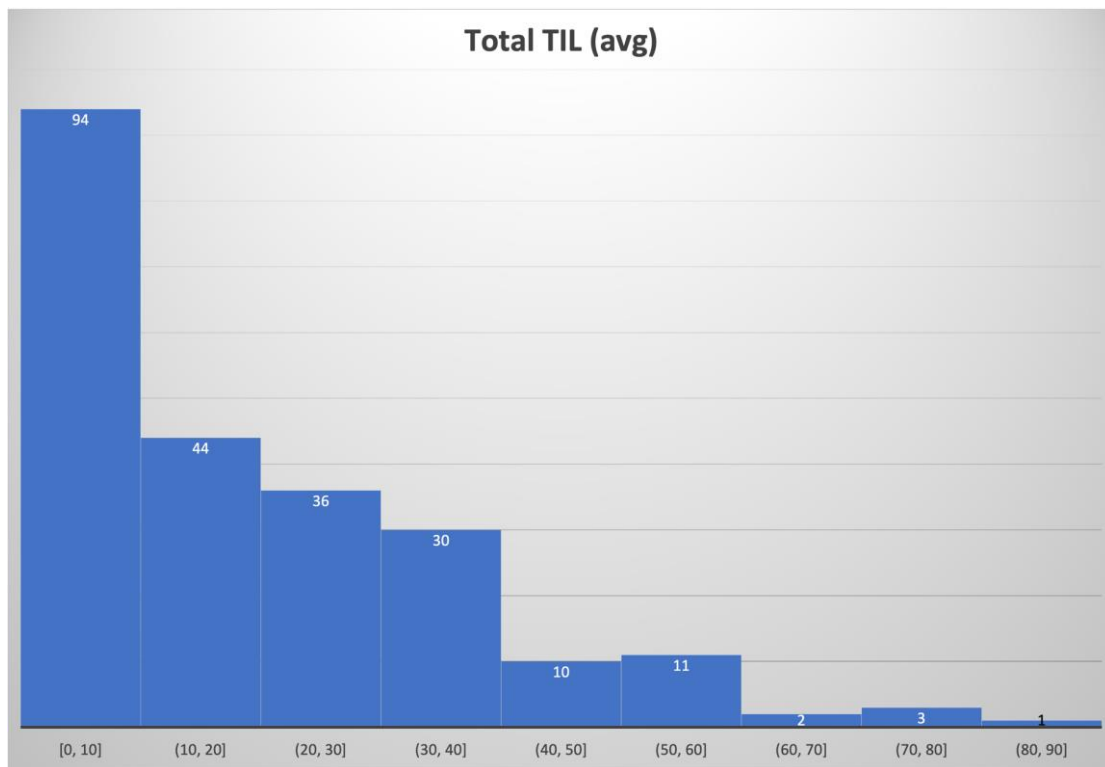








APPENDIX VII- Quartile Decoding for individual TILs (figure in Excel Sheet)



APPENDIX VIII: RNA Extraction Protocol

Extraction of RNA was done according to the manufacturer's

<https://www.qiagen.com/ke/products/discovery-and-translational-research/dna-rna-purification/rna-purification/total-rna/rneasy-ffpe-kit/>)

The scrapped tumour sections were immediately placed in individual 2 ml Eppendorf tube and lid closed. 240ul of Proteinase K Digest Buffer (PKD) was added to each sample tube separately and mixed by vortexing.

- The tubes were centrifuged for 1 min at 11,000g (10,000 rpm)
- 10 µL of proteinase K were added to the lower clear phase of each sample tube and mixed gently by pipetting up and down
- The tubes were incubated at 56°C for 15 minutes and 80°C for a further 15 minutes to ensure almost complete digestion of the tissue by proteinase K and reversal of formaldehyde crosslinks and optimal RNA performance in downstream applications such as real time RT-PCR.
- The tubes were mixed by briefly vortexing every 5 minutes during the 30-minute heating period
- The lower uncoloured phase in the tubes were transferred into individual new 2 ml micro centrifuge tubes.
- The tubes were then incubated on ice for 3 minutes and then centrifuged for 15 minutes at 20,000g (13,500 rpm)
- The supernatant was transferred to a new 2 ml Eppendorf tube taking care not to disturb the pellet (The pellet contains insoluble tissue debris including cross-linked DNA).
- DNase booster buffer of 25ul (a tenth of the total sample volume) and 10ul DNase 1 stock solution were added. This was mixed by inverting the tube followed by a brief centrifugation to collect residual liquid from the sides of the tube.
- The tubes were incubated at room temperature for 15 min.
- 500ul of buffer RBC was added to each tube to adjust binding conditions and mix the lysate thoroughly.
- 1200ul of ethanol (100%) was added to the lysate and mixed well by pipetting
- Immediately 700ul of the sample (including any precipitate formed) was transferred to a RNeasy MinElute spin column placed in a 2 ml collection tube which was re-used for the next three steps. The lids of the tubes were closed gently and centrifuged for 15 seconds at 8000xg (10,000 rpm). The flow through was discarded.
- The step above was repeated until all the sample had passed through the RNeasy MinElute spin column.
- 500 µL of Buffer RPE was then added to the RNeasy MinElute spin column. The lid was closed gently and centrifuged for 15 seconds at 8000xg (10,000 rpm). The flow through was discarded.
- 500 µL of Buffer RPE was then added to the RNeasy MinElute spin column. The lid was closed gently and centrifuged for 2 minutes at 8000xg (10,000 rpm) to wash the spin column membrane. The RNeasy MinElute spin column was carefully removed from the collection tube to avoid contact with the flow through and ethanol carryover. The collection tube with the flow through was then discarded.

- The RNeasy MinElute spin column was then placed in a new 2 ml collection tube supplied with the kit. The lid of the spin column was opened, and the column centrifuged at full speed for 5 min. to ensure that no ethanol is carried over during RNA elution the collection tube with flow through was discarded.
- The RNeasy MinElute spin column was placed in a new 1.5ml collection tube (supplied with kit) and 20ul RNase free water added directly to the spin column membrane. The lid was closed gently and centrifuged for 1 minute at full speed to elute the RNA.
- The dead volume of the RNeasy MinElute spin column is 2ul: elution with 20ul of RNase free water resulted in an 18ul of RNA elute
- Elutes were then stored at -20°C until the next part of the experiment.

APPENDIX IX: RNA concentration and purity

1. Concentration of RNA can be determined by measuring the absorbance at 260nm (A260nm) in a spectrophotometer.
2. Buffer was diluted in neutral pH and buffer solution was used to calibrate the spectrophotometer.

Quantification of RNA: This was performed on the Nanodrop OneC (Thermo Scientific). 1ul of RNA elute was pipetted from each collection tube (placed on a box of crushed ice to avoid denaturation of RNA). The Nanodrop sampling area was cleaned with tissue made wet with buffer.

Purity of RNA: The ratio of the absorbance between A260nm and 280nm provides an estimate of RNA purity with respect to protein and other contaminants that absorb in the UV spectrum. For accurate values, absorbance is measured in 10mM Tris-Cl, pH 7.5. Pure RNA has an absorbance ratio at 260/280 of 1.9-2.1 in 10mM TrisCl, pH 7.5. Calibration of spectrophotometer must always be done with same solution used for dilution.

DNA contamination: An absorbance reading of 1.0 at 260 nm in a 1 cm detection path corresponds to an RNA concentration of 40ug/ml. A260/A230 ratio should be greater than 1.7. A260/A280 ratio should be 1.8 to 2.0.

Concentration determined by A260 should be greater than 40ug/ml (Ref: Qiagen RT2 PreAMP cDNA Synthesis Handbook- March 2011- page 14).

Results of Quantification of RNA

Sample Name	Nucleic Acid (ng/ μ L)	A260/A280	A260/A230	A260	A280	Nucleic Acid Factor	Baseline Correction (nm)	Baseline Absorbance
Sample 327	25.821	1.838	1.744	0.646	0.351	40	340	-0.006
Sample 346	56.994	1.913	2.048	1.425	0.745	40	340	-0.023
Sample 304	87.275	1.933	2.107	2.182	1.129	40	340	-0.014
Sample 341	161.599	1.936	1.943	4.04	2.087	40	340	0.02
Sample 322	25.283	1.822	1.537	0.632	0.347	40	340	0.087
Sample 133	88.074	1.941	2.159	2.202	1.134	40	340	-0.165
Sample 38	152.878	1.95	2.075	3.822	1.96	40	340	-0.011
Sample 311	23.639	1.717	1.149	0.591	0.344	40	340	0.264
Sample 331	35.132	1.928	2.038	0.878	0.456	40	340	-0.004
Sample 342	201.453	1.955	2.164	5.036	2.576	40	340	-0.012
Sample 308	91.435	1.964	2.177	2.286	1.164	40	340	-0.009
Sample 129	1.87	1.135	0.497	0.047	0.041	40	340	0.005
Sample 129	1.887	1.444	0.528	0.047	0.033	40	340	0.032
Sample 37	269.65	1.963	2.126	6.741	3.435	40	340	0.026
Sample 123	277.23	1.959	2.112	6.931	3.537	40	340	0.046
Sample 272	64.464	1.91	2.006	1.612	0.844	40	340	0.031
Sample 42	-0.426	1.084	0.423	-0.011	-0.01	40	340	0.119
Sample 324	51.85	1.861	1.853	1.296	0.697	40	340	0.086
Sample 35	36.749	1.95	2.126	0.919	0.471	40	340	0.011
Sample 234	82.087	1.936	2.085	2.052	1.06	40	340	0.014

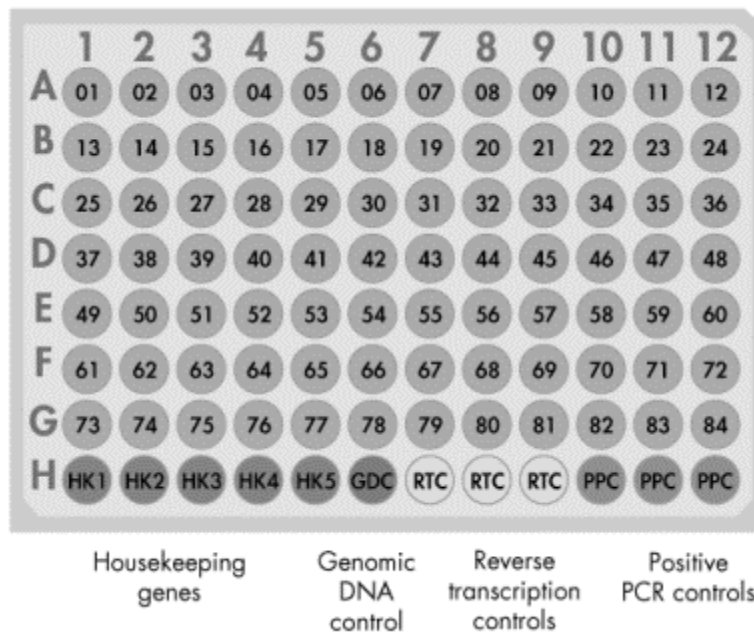
APPENDIX X: Samples that passed RNA integrity test

Sample No	ng/UI	Subtype
Sample 346	56.994	Triple negative
Sample 304	87.275	ER/PR positive
Sample 341	161.599	ER/PR positive
Sample 133	88.074	ER/PR positive
Sample 38	152.878	ER/PR positive
Sample 342	201.453	ER/PR positive
Sample 308	191.435	Triple negative
Sample 37	269.65	ER/PR positive
Sample 123	277.23	ER/PR positive
Sample 272	64.464	Triple Negative
Sample 324	51.85	ER/PR positive
Sample 234	82.087	Triple Negative

APPENDIX XI: Components of the RT2 Profiler PCR array

Housekeeping genes: there are 5 housekeeping genes in each array which enable normalization of the data.

Genomic DNA control: this specifically detects non transcribed genomic DNA contamination with a high level of sensitivity.



Reverse transcription controls: these test the efficiency of the reverse transcription reaction performed with the RT2 first strand kit by detecting template synthesized from the kits in-built external RNA control.

Positive PCR controls: This consists of a pre-dispensed artificial DNA sequence and the assay that detects it. This control tests for the PCR reaction itself. The controls in replicate allow for intra plate and inter well consistency.

Human Cancer Inflammation & Immunity Crosstalk PCR Array

(http://www.sabiosciences.com/rt_pcr_product/HTML/PAHS-181Z.html)

The Human Cancer Inflammation & Immunity Crosstalk RT² Profiler PCR Array profiles the expression of 84 key genes involved in mediating communication between tumour cells and the cellular mediators of inflammation and immunity. In addition to epithelial and stromal compartments, the tumour microenvironment contains several cell types of the innate and adaptive immune systems including B and T lymphocytes, dendritic cells, and macrophages. In response to tumor-associated antigens presented via MHC Class I molecules, or to abnormal molecular patterns recognized by Toll-like receptors, the immune system eliminates target cells using a variety of effector enzymes and the engagement of pro-apoptotic signals including TRAIL and FAS ligand. If normal homeostasis is not resolved quickly, a state of chronic inflammation can ensue, including locally increased levels of reactive oxygen and nitrogen species that promote genomic instability. Immune cells produce a variety of cytokines that coordinate the inflammatory response, which is fueled by positive feedback loops commonly

involving the STAT and NF κ B signalling pathways in tumour cells. The resulting up regulation of anti-apoptotic and immunosuppressive factors enables transformed cells to proliferate unchecked by the immune system. During cancer progression, the repertoire of chemokines, cytokines, and growth factors that orchestrates normal immune responses can be commandeered to create an immunosuppressive state that facilitates invasion and metastasis. The genes profiled with this array include mediators and effectors of the crosstalk between tumours and the immune system that influences the course of cancer progression. A set of controls present on each array enables data analysis using the $\Delta\Delta$ CT method of relative quantification as well as assessment of reverse transcription performance, genomic DNA contamination, and PCR performance. Using real-time PCR, researchers can easily and reliably analyze the expression of a focused panel of genes involved in cancer inflammation and immune crosstalk with this array.

The RT² Profiler PCR Arrays are intended for molecular biology applications. This product is not intended for the diagnosis, prevention, or treatment of a disease.

96-well Plate, 384-well (4 $\times$ 96) Plate, and 100-well Disc formats are available.

Immune & Inflammatory Responses:

Immunostimulatory Factors: IFNG, IL2, IL12A, IL12B, IL15, TNF.

Immunosuppressive Factors: CD274 (PD-L1), CSF2 (GM-CSF), CTLA4, CXCL12 (SDF1), CXCL5, IDO1 (IDO), IL10, IL13, IL4, IL8, MIF, NOS2 (iNOS), PDCD1 (PD1), PTGS2 (COX2), TGF B1, VEGFA.

Pro-Inflammatory Genes: CCL2 (MCP-1), CCL20 (MIP-3A), IFNG, IL1A, IL1B, IL2, IL6, IL12A, IL12B, IL17A, IL23A, PTGS2 (COX2), TLR4, TNF, VEGFA.

Anti-Inflammatory Genes: IL4, IL10, IL13, TGF B1.

Enzymatic Modulators of Inflammation & Immunity: AICDA (AID), GZMA, GZMB, IDO1 (IDO), NOS2 (iNOS), PTGS2 (COX2).

Antigen Presentation: HLA-A, HLA-B, HLA-C, MICA, MICB.

Chemokines: CCL2 (MCP-1), CCL4 (MIP-1B), CCL5 (RANTES), CCL18 (PARC), CCL20 (MIP-3A), CCL21, CCL22 (MDC), CCL28, CXCL1, CXCL2, CXCL5, CXCL9 (MIG), CXCL10 (IP-10), CXCL11 (I-TAC, IP-9), CXCL12 (SDF1).

Chemokine Receptors: ACKR3 (CXCR7), CCR1, CCR2, CCR4, CCR7, CCR9, CCR10, CXCR1 (IL8RA), CXCR2 (IL8RB), CXCR3, CXCR4, CXCR5.

Interleukins: IL1A, IL1B, IL2, IL4, IL6, IL8, IL10, IL12A, IL12B, IL13, IL15, IL17A, IL23A.

Other Cytokines: KITLG (SCF), MIF, SPP1, TNF, TNFSF10 (TRAIL).

Growth Factors & Receptors: CSF1 (MCSF), CSF2 (GM-CSF), CSF3 (GCSF), EGF, EGFR, IGF1, TGF B1, VEGFA.

Signal Transduction:

Interferon Signalling: GBP1, IFNG, IL6, IRF1.

Interferon-Responsive Genes: CCL2 (MCP-1), CCL5 (RANTES), CXCL9 (MIG), CXCL10 (IP-10), GBP1, IRF1, MYD88, STAT1, TLR3, TNFSF10 (TRAIL).

NF κ B Targets: BCL2L1 (BCL-XL), CCL2 (MCP-1), CCL5 (RANTES), CSF1 (MCSF), CSF2 (GM-CSF), CSF3 (GCSF), IFNG, IL8, TNF.

STAT Targets: CCL2 (MCP-1), CCL4 (MIP-1B), CCL5 (RANTES), CSF1 (MCSF), CSF2 (GM-CSF), CSF3 (GCSF), CXCL9 (MIG), CXCL10 (IP-10), CXCL11 (I-TAC, IP-9), CXCL12 (SDF1), IL1B, IL6, IL8, IL10, IL17A, IL23A, MYC.

Toll-Like Receptor Signalling: TLR2, TLR3, TLR4, MYD88.

Transcription Factors: FOXP3, HIF1A, IRF1, MYC, NF κ B1, STAT1, STAT3, TP53 (p53).

Apoptosis: Pro-Apoptotic: FASLG (TNFSF6), TNF, TNFSF10 (TRAIL), TP53 (p53). Anti-Apoptotic: BCL2, BCL2L1 (BCL-XL), MYC, STAT3

APPENDIX XII: Quantification of RNA

Sample Name	Nucleic Acid (ng/ μ L)	A260/A280	A260/A230	A260	A280	Nucleic Acid Factor	Baseline Correction (nm)	Baseline Absorbance
Sample 327	25.821	1.838	1.744	0.646	0.351	40	340	-0.006
Sample 346	56.994	1.913	2.048	1.425	0.745	40	340	-0.023
Sample 304	87.275	1.933	2.107	2.182	1.129	40	340	-0.014
Sample 341	161.599	1.936	1.943	4.04	2.087	40	340	0.02
Sample 322	25.283	1.822	1.537	0.632	0.347	40	340	0.087
Sample 133	88.074	1.941	2.159	2.202	1.134	40	340	-0.165
Sample 38	152.878	1.95	2.075	3.822	1.96	40	340	-0.011
Sample 311	23.639	1.717	1.149	0.591	0.344	40	340	0.264
Sample 331	35.132	1.928	2.038	0.878	0.456	40	340	-0.004
Sample 342	201.453	1.955	2.164	5.036	2.576	40	340	-0.012
Sample 308	91.435	1.964	2.177	2.286	1.164	40	340	-0.009
Sample 129	1.87	1.135	0.497	0.047	0.041	40	340	0.005
Sample 129	1.887	1.444	0.528	0.047	0.033	40	340	0.032
Sample 37	269.65	1.963	2.126	6.741	3.435	40	340	0.026
Sample 123	277.23	1.959	2.112	6.931	3.537	40	340	0.046
Sample 272	64.464	1.91	2.006	1.612	0.844	40	340	0.031
Sample 42	-0.426	1.084	0.423	-0.011	-0.01	40	340	0.119
Sample 324	51.85	1.861	1.853	1.296	0.697	40	340	0.086
Sample 35	36.749	1.95	2.126	0.919	0.471	40	340	0.011
Sample 234	82.087	1.936	2.085	2.052	1.06	40	340	0.014

N=20	Nucleic Acid (ng/ μ L)	260/A280	A260/A230	A260/A280	Baseline	Absorbance
Mean	86.74720	1.80695	1.74485	2.16870	1.12005	.024850
Median	60.72900	1.93050	2.04300	1.51850	.79450	.012500

Notes:

1. Concentration of RNA can be determined by measuring the absorbance at 260nm (A260nm) in a spectrophotometer.
2. Buffer was diluted in neutral pH and buffer solution was used to calibrate the spectrophotometer.

Quantification of RNA: This was performed on the Nanodrop OneC (Thermo Scientific). 1ul of RNA elute was pipetted from each collection tube (placed on a box of crushed ice to avoid denaturation of RNA). The Nanodrop sampling area was cleaned with tissue made wet buffer.

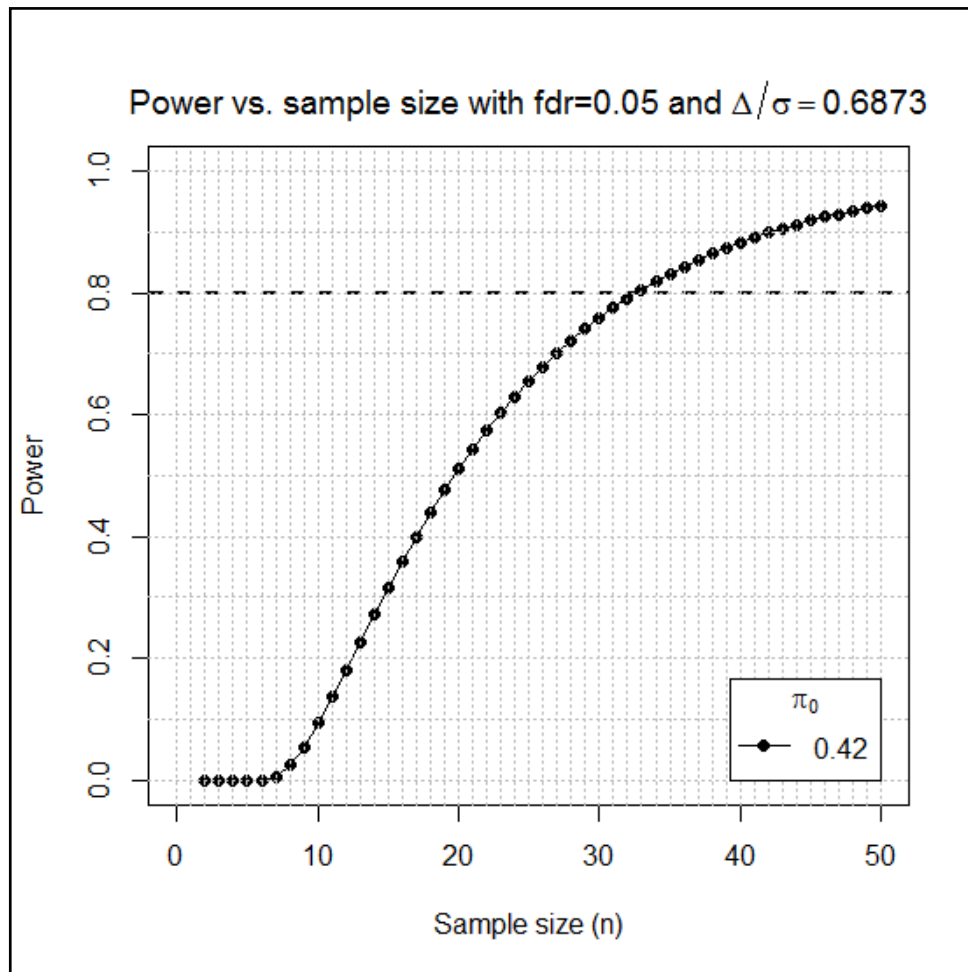
Purity of RNA: The ratio of the absorbance between A260nm and 280nm provides an estimate of RNA purity with respect to protein and other contaminants that absorb in the UV spectrum. For accurate values, absorbance is measured in 10mM Tris-Cl, pH 7.5. Pure RNA has an absorbance ratio at 260/280 of 1.9-2.1 in 10mM TrisCl, pH 7.5. Calibration of spectrophometer must always be done with same solution used for dilution.

DNA contamination: An absorbance reading of 1.0 at 260 nm in a 1 cm detection path corresponds to an RNA concentration of 40ug/ml. A260/A230 ratio should be greater than 1.7. A260/A280 ratio should be 1.8 to 2.0

Concentration determined by A260 should be greater than 40ug/ml (Ref: Qiagen RT2 PreAMP cDNA Synthesis Handbook- March 2011- page 14)

PR positive Sample 272	64.464	Triple Negative
Sample 324	51.85	ER/PR positive
Sample 234	82.087	Triple Negative

APPENDIX XIII: Sample size calculations for Validation experiment



**FOLD
CHANGE
Sample Size
= 33 Per
Group**

APPENDIX XIV: Multi-Analyte ELISArray Kits

These are designed to simultaneously profile the level of multiple cytokines and/or chemokines using the conventional and simple sandwich-based enzyme-linked immunosorbent assay (ELISA) technique. The 96-well ELISA microplate has been coated with a panel of twelve target-specific capture antibodies, one in each eight-well strip allowing you to obtain qualitative relative profiling results from up to six samples. Figure 1 displays the typical layout of a catalogued ELISArray Kit. Each kit also includes the corresponding detection antibodies, Antigen Standards, and a complete set of detection reagents for a colorimetric ELISA.

Figure 1: Layout of Catalogued Multi-Analyte ELISArray Kits

		Cytokines											
		1	2	3	4	5	6	7	8	9	10	11	12
Negative Control	A												
Sample 1	B												
Sample 2	C												
Sample 3	D												
Sample 4	E												
Sample 5	F												
Sample 6	G												
Positive Control	H												

The Multi-Analyte ELISArray Kit uses a standard ELISA technique. Incubation allows the capture antibodies to bind their specific protein of interest. After washing away unbound protein, biotinylated detection antibodies added to the wells also bind the captured analyte. After washing again to remove unbound material, an avidin-horseradish peroxidase conjugate is added. The wells are again washed, and a colorimetric substrate solution is added, which produces a blue color in direct proportion to the amount of protein analyte present in the initial sample.

The color development is stopped by adding stop solution, and the absorbance at 450 nm is read and compared across your samples.

All wells of an eight-well strip are coated with the same capture antibody for the same cytokine or chemokine. A set of twelve strips in one ELISArray microplate therefore represents twelve cytokines or chemokines. The corresponding lettered wells of each row will characterize the same biological sample.

Each ELISArray microplate therefore characterizes up to six biological samples plus positive and negative controls.

Protocol:**Sample Preparation and Handling:**

High quality input material is **ESSENTIAL** for obtaining good results.

The most important prerequisite for any ELISA analysis experiment is consistent, high- quality experimental samples. Therefore, the sample preparation and handling procedures are critical to the success of the experiment. Specimens should be clear, because residual traces of particulate matter, heme, lipids, or other contaminants will interfere with the performance of the ELISA.

1. Recommended Preparation Methods:

Roughly 650 μ l of each sample is required to complete the experiment.

High quality protein samples for the ELISA experiment must be prepared using one of the following methods, each specific for your biological sample:

Cell Culture Supernatants: Remove any particulate material by centrifugation for 10 minutes at 1000 x g and assay immediately, or aliquot and store samples at $\leq -20^{\circ}\text{C}$. Avoid repeated freeze / thaw cycles.

Reagent Preparation: Use only polypropylene tubes for the antigen and detection antibody steps.

Briefly centrifuge any 1.5 ml tube in the kit before opening it in order to collect its contents at the bottom of the tube.

2. ELISArray Kit Reagents: Thaw the 10% BSA and Donkey Serum at room temperature. Store on ice.

Bring the Wash Buffer Concentrate, Assay Buffer Stock, Sample Dilution Buffer Stock and ELISArray plate to room temperature. Do not remove the ELISArray plate from its pouch at this time.

Once finished with each reagent, return it to its proper storage conditions.

- a. Keep all other kit components at their recommended storage temperature until recommended by the protocol.

3. Wash Buffer

Visually inspect the Wash Buffer Concentrate to ensure that all components are in solution. If any precipitates are visible, briefly shake the bottle to suspend any precipitate. Dilute 50 ml of Wash Buffer Concentrate into de-ionized or distilled water (dH₂O) to a final volume of 500 ml. Transfer to wash bottle. Keep at room temperature.

4. Assay Buffer

Dilute 0.6 ml of 10% BSA into a final volume of 30 ml with Assay Buffer Stock. Keep at room temperature.

5. Sample Dilution Buffer:

For cell culture supernatant samples, prepare Sample Dilution Buffer 1: Dilute 2 ml of 10% BSA into a final volume of 20 ml with Sample Dilution Buffer Stock. Store

on ice.

Assay Procedure:

NOTE: *Very accurate and precise pipetting is critical to the reliability and consistency of any ELISA experiment. Make sure that all your single- and multi-channel pipettors are calibrated before starting this protocol. Maintain incubation times and temperatures as consistent as possible across arrays and experiments for best results.*

6. Generate an Antigen Standard Cocktail:

- a. Just before use, thaw the Antigen Standards on ice for 20 minutes. Vortex gently to mix well.
- b. In a single tube, prepare the Concentrated Antigen Standard Cocktail containing all 12 of the Antigen Standards by pipetting 10 μ l of each Antigen Standard into the same 880 μ l volume of the appropriate Sample Dilution Buffer to yield 1 ml of a Concentrated Antigen Standard Cocktail.

NOTE: *For Mix-N-Match ELISArrays see your Product Specification Sheet for Generating Antigen Standard Cocktail*

- c. Dilute 200 μ l of the concentrated Standard Antigen Cocktail into 800 μ l of the appropriate Sample Dilution Buffer for the final Antigen Standard Cocktail.

7. Remove the ELISArray plate from its pouch.

Using a multi-channel pipette, add 50 μ l of Assay Buffer into each well of the ELISArray plate.

NOTE: *For best results, add all samples and reagents carefully by touching the pipette tips to the side of the wells just above the bottom and allowing the dispensed volume to run down the well wall. Do not blow out the last drop of volume dispensed to avoid introducing bubbles into the wells.*

8. Add 50 μ l of the appropriate Sample Dilution Buffer into each well of Row A in the ELISArray plate to setup the negative control.

NOTE: *Change your pipette tips with every addition to avoid cross-contamination of the wells.*

9. Add 50 μ l of each experimental sample or their dilutions to each well of their respective rows (B through G) in the ELISArray plate.

10. Add 50 μ l of the final Antigen Standard Cocktail into each well of Row H in the ELISArray plate to set up the positive control.

11. Cover the plate and gently shake or tap the plate for 10 seconds to mix. Incubate for 2 hours at room temperature on your bench top.

12. Allow the Development Solution and Stop Solution to warm to room temperature in lab bench drawer protected from light.

NOTE: *Avoid prolonged exposure of the Development Solution to light or contact with water, air, or extreme temperature.*

13. Dilute the Detection Antibodies just before the next step:
 - a. Thaw the Detection Antibodies on ice for 30 minutes.
 - b. Add 855 μ l of Assay Buffer to each tube of Detection Antibody.
 - c. Mix well but gently.

Transfer each detection antibody to its own empty tube of the Detection

Multi-Analyte ELISArray Kits

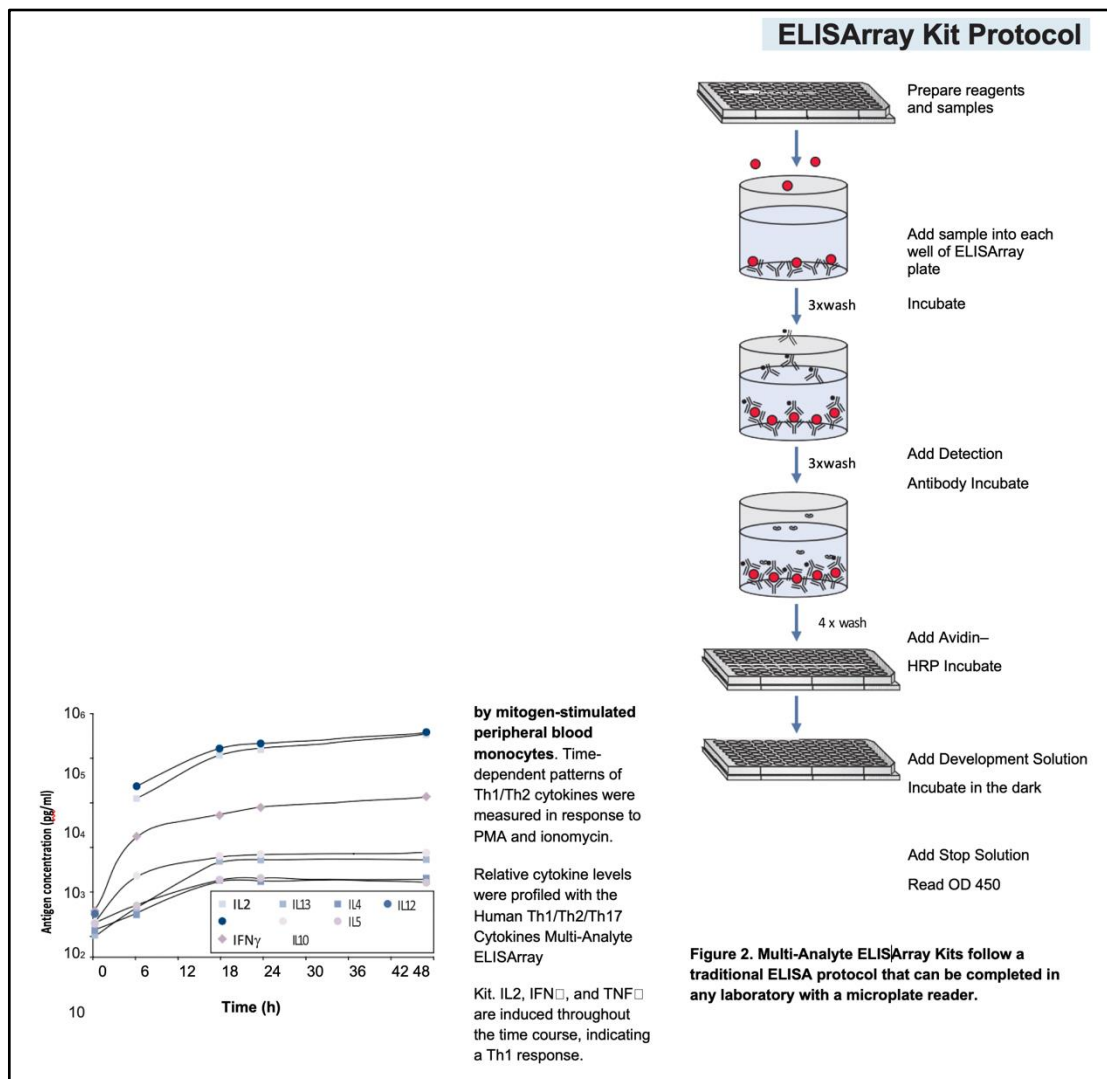
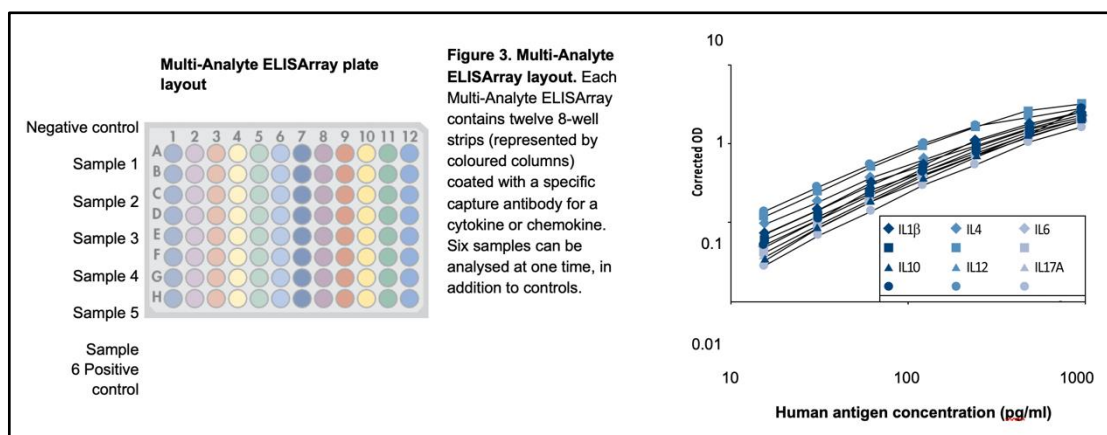


Figure. Multi-Analyte ELISArray Kits profile cytokines produced



- d. Ensure to add the Detection Antibodies to the tubes in the same left-to-right order as the capture antibodies in columns 1 through 12. See the Product Specification Sheet for the correct orientation.

14. Washing the ELISArray Plate:

- a. Decant or aspirate the contents of the wells.
- b. Wash wells by filling each well with 1x Wash Buffer (350 μ L per well).
- c. Gently shake or tap the plate for 10 seconds to mix and follow by decanting or aspirating the solution.
- d. Blot the plate upside down on absorbent paper to remove any residual buffer.
- e. Repeat twice for a total of three washes.

15. Using a 12-channel pipettor, transfer 100 μ l of the dilute Detection Antibodies from the Dilution Tube Strip to the appropriate rows of the ELISArray plate.

Cover the plate and gently shake or tap the ELISArray plate for 10 seconds to mix. Incubate for 1 hour at room temperature on your bench top.

NOTE: Be sure to correctly orient your Detection Antibody Dilution Tube Strip above row A of the ELISArray plate to transfer the correct Detection Antibody to its correct wells.

16. Prepare Avidin-HRP:

Just before use, thaw the Avidin-HRP Conjugate on ice for 20 minutes. Vortex gently to mix well.

Add 11 μ l of the Avidin-HRP Conjugate to 11 ml of Assay Buffer.

17. NOTE: Avoid prolonged exposure of the Avidin-HRP Solution to light.

18. Wash ELISA wells as in step 9.

19. Add 100 μ l of dilute Avidin-HRP into all wells.

Cover the plate and gently shake or tap the ELISArray for 10 seconds to mix. Incubate for 30 minutes at room temperature in the dark.

20. Wash ELISA wells as in step 9, except for a total of 4 washes.

Just before use, transfer 12 ml of the Development Solution from its original bottle into a clean multi-channel pipette reservoir.

NOTE: Use a different clean multi-channel reservoir for each solution to be dispensed.

21. Add 100 μ l of the Development Solution to each well.

Incubate the plate for 15 minutes at room temperature in the dark.

22. Add 100 μ l of Stop Solution to each well in the same order as the Development Solution was added. The color in the wells should change from blue to yellow.
23. Read absorbance at 450 nm within 30 minutes of stopping the reaction. If wavelength correction is available, subtract readings at 570 nm from the reading at 450 nm.

NOTE: Subtracting the A450 reading by the A570 reading corrects for any minor optical imperfections

Data Analysis:

24. Processing the Raw Data

- a. Typical absorbance values should range from 0.00 to 2.50.

Absorbance values greater than 2.50 are not within the linear range of the assay.

Also, do not interpret absorbance values less than two times the negative control absorbance values for each antigen.

- b. For each antigen, subtract the observed absorbance by the absorbance of the negative control to obtain the corrected absorbance values.

25. Profiling Experiment

- a. For each antigen separately, compare the corrected absorbance values between samples to determine which cytokine or chemokine changes its protein level most or least dramatically. Be sure to account for any dilution of the samples.
- b. Follow up the profiling experiment with a more detailed analysis using the corresponding Single Analyte ELISArray Kit.

NOTE: In the ELISA plate. Uncorrected A450 readings may yield artificially high signals.

APPENDIX XV: Cytokine Elisa Raw Data

Table 24: Plate1 311019

	1	2	3	4	5	6	7	8	9	10	11	12
NEG -A	0.088	0.063	0.082	0.099	0.105	0.08	0.092	0.164	0.091	0.123	0.132	0.085
	0.03	0.039	0.041	0.041	0.041	0.037	0.042	0.042	0.041	0.041	0.4	0.03
B-(ER)	0.054	0.044	0.048	0.044	0.047	0.053	0.051	0.071	0.068	0.097	0.126	0.091
	0.04	0.039	0.039	0.038	0.04	0.039	0.039	0.04	0.041	0.04	0.03	0.029
C-(ER)	0.055	0.045	0.052	0.048	0.035	0.06	0.054	0.093	0.067	0.101	0.132	0.111
	0.039	0.039	0.039	0.039	0.029	0.04	0.039	0.04	0.037	0.039	0.04	0.042
D-(ER)	0.059	0.035	0.051	0.045	0.052	0.068	0.053	0.083	0.073	0.102	0.15	0.106
	0.039	0.031	0.038	0.039	0.039	0.041	0.039	0.039	0.039	0.039	0.04	0.039
E-(TN)	0.05	0.044	0.049	0.046	0.044	0.047	0.049	0.067	0.063	0.082	0.109	0.087
	0.04	0.04	0.04	0.041	0.04	0.04	0.039	0.039	0.041	0.041	0.041	0.04
F-(TN)	0.034	0.043	0.047	0.045	1.397	0.05	0.132	0.084	0.06	0.075	0.112	0.191
	0.033	0.036	0.04	0.039	1.295	0.041	0.035	0.042	0.041	0.04	0.041	0.04
G-(TN)	0.055	0.2	0.043	0.035	0.045	0.034	0.064	0.069	0.065	0.083	0.112	0.12
	0.039	0.039	0.037	0.03	0.04	0.03	0.04	0.04	0.041	0.042	0.041	0.042
POS-H	0.968	3.09	0.923	2.241	2.512	2.55	3.145	2.168	0.758	0.685	1.814	0.735
	0.04	0.053	0.042	0.047	0.049	0.057	0.053	0.048	0.042	0.043	0.073	0.043

Background OD removal

	1	2	3	4	5	6	7	8	9	10	11	12
NEG -A	0.058	0.024	0.041	0.058	0.064	0.043	0.05	0.122	0.05	0.082	-0.268	0.055
B-(ER)	0.014	0.005	0.009	0.006	0.007	0.014	0.012	0.031	0.027	0.057	0.096	0.062
C-(ER)	0.016	0.006	0.013	0.009	0.006	0.02	0.015	0.053	0.03	0.062	0.092	0.069
D-(TN)	0.02	0.004	0.013	0.006	0.013	0.027	0.014	0.044	0.034	0.063	0.11	0.067
E-(TN)	0.01	0.004	0.009	0.005	0.004	0.007	0.01	0.028	0.022	0.041	0.068	0.047
F-(ER)	0.001	0.007	0.007	0.006	0.102	0.009	0.097	0.042	0.019	0.035	0.071	0.151
G-(ER)*	0.016	0.161	0.006	0.005	0.005	0.004	0.024	0.029	0.024	0.041	0.071	0.078
POS-H	0.928	3.037	0.881	2.194	2.463	2.493	3.092	2.12	0.716	0.642	1.741	0.692

Neg control removal

	1	2	3	4	5	6	7	8	9	10	11	12
NEG -A	0	0	0	0	0	0	0	0	0	0	0	0
B-(ER)	-0.044	-0.019	-0.032	-0.052	-0.057	-0.029	-0.038	-0.091	-0.023	-0.025	0.364	0.007
C-(ER)	-0.042	-0.018	-0.028	-0.049	-0.058	-0.023	-0.035	-0.069	-0.02	-0.02	0.36	0.014
D-(TN)	-0.038	-0.02	-0.028	-0.052	-0.051	-0.016	-0.036	-0.078	-0.016	-0.019	0.378	0.012
E-(TN)	-0.048	-0.02	-0.032	-0.053	-0.06	-0.036	-0.04	-0.094	-0.028	-0.041	0.336	-0.008
F-(ER)	-0.057	-0.017	-0.034	-0.052	0.038	-0.034	0.047	-0.08	-0.031	-0.047	0.339	0.096
G-(ER)*	-0.042	0.137	-0.035	-0.053	-0.059	-0.039	-0.026	-0.093	-0.026	-0.041	0.339	0.023
POS-H	0.87	3.013	0.84	2.136	2.399	2.45	3.042	1.998	0.666	0.56	2.009	0.637

Table 25: Plate 2 311019

	1	2	3	4	5	6	7	8	9	10	11	12
NEG - A	0.126	0.132	0.161	0.121	0.125	0.178	0.142	0.312	0.146	0.221	0.216	0.245
	0.038	0.043	0.042	0.043	0.042	0.056	0.042	0.041	0.04	0.044	0.042	0.059
B-(ER)	0.104	0.098	0.102	0.103	0.118	0.159	0.116	0.16	0.123	0.218	0.206	0.182
	0.04	0.048	0.04	0.04	0.04	0.04	0.041	0.042	0.041	0.04	0.041	0.048
C-(TN)	0.102	0.092	0.152	0.16	0.1	0.163	0.163	0.22	0.156	0.212	0.264	0.227
	0.04	0.04	0.04	0.051	0.042	0.04	0.043	0.045	0.041	0.047	0.041	0.043
D-(TN)	0.14	0.093	0.098	0.105	0.077	0.226	0.154	0.415	0.348	0.475	0.385	0.434
	0.045	0.041	0.041	0.04	0.039	0.041	0.049	0.042	0.04	0.042	0.043	0.045
E-(TN)	0.089	0.089	0.123	0.198	0.163	0.18	0.115	0.268	0.186	0.273	0.331	0.205
	0.042	0.042	0.043	0.042	0.041	0.041	0.04	0.045	0.042	0.051	0.052	0.059
F-(TN)	0.081	0.063	0.077	0.299	0.078	0.173	0.231	0.171	0.683	0.215	0.258	0.248
	0.028	0.03	0.04	0.041	0.041	0.042	0.04	0.048	0.053	0.046	0.051	0.047
G-(ER)	0.085	0.084	0.171	0.139	0.201	0.139	0.451	0.253	0.171	0.26	0.298	0.462
	0.029	0.036	0.042	0.059	0.0042	0.041	0.042	0.042	0.041	0.042	0.045	0.043
POS-H	0.278	0.134	0.248	0.082	0.109	0.786	0.576	0.408	0.285	0.354	0.405	0.342
	0.03	0.03	0.041	0.042	0.39	0.041	0.042	0.042	0.042	0.043	0.06	0.044

Background OD removal

	1	2	3	4	5	6	7	8	9	10	11	12
NEG -A	0.088	0.089	0.119	0.078	0.083	0.122	0.1	0.271	0.106	0.177	0.174	0.186
B-(ER)	0.064	0.05	0.062	0.063	0.078	0.119	0.075	0.118	0.082	0.178	0.165	0.134
C-(ER)	0.062	0.052	0.112	0.109	0.058	0.123	0.12	0.175	0.115	0.165	0.223	0.184
D-(TN)	0.095	0.052	0.057	0.065	0.038	0.185	0.105	0.373	0.308	0.433	0.342	0.389
E-(TN)	0.047	0.047	0.08	0.156	0.122	0.139	0.075	0.223	0.144	0.222	0.279	0.146
F-(ER)	0.053	0.033	0.037	0.258	0.037	0.131	0.191	0.123	0.63	0.169	0.207	0.201
G-(ER)*	0.056	0.048	0.129	0.08	0.1968	0.098	0.409	0.211	0.13	0.218	0.253	0.419
POS-H	0.248	0.104	0.207	0.04	-0.281	0.745	0.534	0.366	0.243	0.311	0.345	0.298

Neg control removal

	1	2	3	4	5	6	7	8	9	10	11	12
NEG -A	0	0	0	0	0	0	0	0	0	0	0	0
B-142 (ER)	-0.024	-0.039	-0.057	-0.015	-0.005	-0.003	-0.025	-0.153	-0.024	0.001	-0.009	-0.052
C- 147 (ER)	-0.026	-0.037	-0.007	0.031	-0.025	0.001	0.02	-0.096	0.009	-0.012	0.049	-0.002
D-169(TN)	0.007	-0.037	-0.062	-0.013	-0.045	0.063	0.005	0.102	0.202	0.256	0.168	0.203
E-170 (TN)	-0.041	-0.042	-0.039	0.078	0.039	0.017	-0.025	-0.048	0.038	0.045	0.105	-0.04
F-183(ER)	-0.035	-0.056	-0.082	0.18	-0.046	0.009	0.091	-0.148	0.524	-0.008	0.033	0.015
G- 190(ER)*	-0.032	-0.041	0.0098	0.002	0.1138	-0.024	0.309	-0.06	0.024	0.041	0.079	0.233
POS-H	0.16	0.015	0.088	-0.038	-0.364	0.623	0.434	0.095	0.137	0.134	0.171	0.112

Table 26: Plate 2: 291019

	1	2	3	4	5	6	7	8	9	10	11	12
NEG -A	0.076	0.047	0.078	0.066	0.085	0.078	0.075	0.182	0.131	0.112	0.135	0.083
	0.039	0.041	0.042	0.041	0.041	0.041	0.042	0.043	0.042	0.043	0.043	0.042
B-(ER)	0.06	0.054	0.052	0.053	0.052	0.056	0.065	0.07	0.086	0.086	0.103	0.063
	0.044	0.047	0.042	0.045	0.044	0.045	0.042	0.043	0.045	0.042	0.042	0.042
C-(ER)	0.056	0.05	0.051	0.059	0.049	0.05	0.052	0.079	0.079	0.081	0.085	0.06
	0.041	0.04	0.041	0.043	0.041	0.04	0.04	0.044	0.04	0.04	0.043	0.041
D-(TN)	0.057	0.049	0.053	0.059	0.06	0.054	0.052	0.068	0.094	0.073	0.085	0.062
	0.046	0.042	0.044	0.052	0.049	0.043	0.042	0.04	0.04	0.041	0.045	0.044
E-(TN)	0.049	0.048	0.052	0.048	0.049	0.049	0.05	0.068	0.105	0.095	0.086	0.063
	0.04	0.042	0.042	0.041	0.042	0.041	0.042	0.041	0.041	0.049	0.042	0.041
F-(ER)	0.052	0.048	0.052	0.049	0.05	0.049	0.055	0.074	0.086	0.093	0.105	0.068
	0.043	0.042	0.041	0.04	0.042	0.04	0.044	0.043	0.04	0.046	0.042	0.041
G-(ER)*	0.053	0.05	0.049	0.05	0.047	0.048	0.05	0.073	0.051	0.067	0.08	0.075
	0.047	0.041	0.04	0.041	0.041	0.04	0.04	0.041	0.042	0.047	0.043	0.042
POS-H	1.248	2.794	1.33	2.544	0.949	2.628	3.213	2.646	1.005	1.314	2.216	1.288
	0.044	0.051	0.044	0.05	0.042	0.05	0.055	0.052	0.044	0.048	0.047	0.044

Background OD removal

	1	2	3	4	5	6	7	8	9	10	11	12
NEG -A	0.037	0.006	0.036	0.025	0.044	0.037	0.033	0.139	0.089	0.069	0.092	0.041
B-(ER)	0.016	0.007	0.01	0.008	0.008	0.011	0.023	0.027	0.041	0.044	0.061	0.021
C-(ER)	0.015	0.01	0.01	0.016	0.008	0.01	0.012	0.035	0.039	0.041	0.042	0.019
D-(TN)	0.011	0.007	0.009	0.007	0.011	0.011	0.01	0.028	0.054	0.032	0.04	0.018
E-(TN)	0.009	0.006	0.01	0.007	0.007	0.008	0.008	0.027	0.064	0.046	0.044	0.022
F-(ER)	0.009	0.006	0.011	0.009	0.008	0.009	0.011	0.031	0.046	0.047	0.063	0.027
G-(ER)*	0.006	0.009	0.009	0.009	0.006	0.008	0.01	0.032	0.009	0.02	0.037	0.033
POS-H	1.204	2.743	1.286	2.494	0.907	2.578	3.158	2.594	0.961	1.266	2.169	1.244

Neg control removal

	1	2	3	4	5	6	7	8	9	10	11	12
NEG -A	0	0	0	0	0	0	0	0	0	0	0	0
B-142 (ER)	-0.021	0.001	-0.026	-0.017	-0.036	-0.026	-0.01	-0.112	-0.048	-0.025	-0.031	-0.02
C-147 (ER)	-0.022	0.004	-0.026	-0.009	-0.036	-0.027	-0.021	-0.104	-0.05	-0.028	-0.05	-0.022
D-169(TN)	-0.026	0.001	-0.027	-0.018	-0.033	-0.026	-0.023	-0.111	-0.035	-0.037	-0.052	-0.023
E-170 (TN)	-0.028	0	-0.026	-0.018	-0.037	-0.029	-0.025	-0.112	-0.025	-0.023	-0.048	-0.019
F-183(ER)	-0.028	0	-0.025	-0.016	-0.036	-0.028	-0.022	-0.108	-0.043	-0.022	-0.029	-0.014
G-190(ER)*	-0.031	0.003	-0.027	-0.016	-0.038	-0.029	-0.023	-0.107	-0.08	-0.049	-0.055	-0.008
POS-H	1.167	2.737	1.25	2.469	0.863	2.541	3.125	2.455	0.872	1.197	2.077	1.203

Table 27: PLATE 3:291019

	1	2	3	4	5	6	7	8	9	10	11	12
NEG -A	0.062	0.049	0.07	0.066	0.054	0.071	0.073	0.15	0.097	0.105	0.142	0.102

	0.04	0.041	0.041	0.042	0.041	0.044	0.041	0.042	0.041	0.041	0.042	0.047
B-(ER)	0.055	0.049	0.05	0.048	0.049	0.066	0.053	0.074	0.08	0.085	0.112	0.07
	0.041	0.042	0.045	0.042	0.041	0.048	0.043	0.043	0.042	0.042	0.044	0.046
C-(ER)	0.05	0.061	0.056	0.048	0.054	0.054	0.054	0.062	0.064	0.065	0.085	0.056
	0.042	0.047	0.049	0.041	0.044	0.045	0.044	0.041	0.04	0.041	0.048	0.04
D-(ER)	0.054	0.05	0.054	0.206	0.062	0.055	0.056	0.061	0.07	0.077	0.09	0.07
	0.043	0.044	0.046	0.143	0.052	0.047	0.04	0.04	0.039	0.04	0.05	0.045
E-(ER)	0.053	0.061	0.05	0.062	0.051	0.051	0.058	0.07	0.088	0.078	0.109	0.068
	0.042	0.066	0.042	0.057	0.046	0.043	0.045	0.048	0.045	0.044	0.053	0.045
F-(ER)	0.057	0.055	0.061	0.051	0.067	0.067	0.07	0.118	0.086	0.109	0.126	0.082
	0.045	0.048	0.05	0.043	0.054	0.048	0.052	0.08	0.044	0.054	0.043	0.049
G-(TN)	0.047	0.052	0.051	0.06	0.352	0.059	0.36	0.889	0.077	0.072	0.088	0.07
	0.04	0.044	0.043	0.05	0.122	0.049	0.123	0.694	0.044	0.043	0.049	0.048
POS-H	0.801	1.938	0.77	2.268	2.552	2.246	3.175	2.652	0.787	0.192	1.668	1.047
	0.042	0.047	0.042	0.05	0.048	0.048	0.054	0.051	0.043	0.041	0.052	0.043

Background OD removal

	1	2	3	4	5	6	7	8	9	10	11	12
NEG -A	0.022	0.008	0.029	0.024	0.013	0.027	0.032	0.108	0.056	0.064	0.1	0.055
B-(ER)	0.014	0.007	0.005	0.006	0.008	0.018	0.01	0.031	0.038	0.043	0.068	0.024
C-(ER)	0.008	0.014	0.007	0.007	0.01	0.009	0.01	0.021	0.024	0.024	0.037	0.016
D-(TN)	0.011	0.006	0.008	0.063	0.01	0.008	0.016	0.021	0.031	0.037	0.04	0.025
E-(TN)	0.011	-0.01	0.008	0.005	0.005	0.008	0.013	0.022	0.043	0.034	0.056	0.023
F-(ER)	0.012	0.007	0.011	0.008	0.013	0.019	0.018	0.038	0.042	0.055	0.083	0.033
G-(ER)*	0.007	0.008	0.008	0.01	0.23	0.01	0.237	0.195	0.033	0.029	0.039	0.022
POS-H	0.759	1.891	0.728	2.218	2.504	2.198	3.121	2.601	0.744	0.151	1.616	1.004

Neg control removal

	1	2	3	4	5	6	7	8	9	10	11	12
NEG -A	0	0	0	0	0	0	0	0	0	0	0	0
B-(ER)	-0.008	-0	-0.024	-0.02	-0.005	-0.01	-0.02	-0.08	-0.02	-0.02	-0.032	-0.031
C-(ER)	-0.014	0.006	-0.022	-0.02	-0.003	-0.02	-0.02	-0.09	-0.03	-0.04	-0.063	-0.039
D-(TN)	-0.011	-0	-0.021	0.039	-0.003	-0.02	-0.02	-0.09	-0.03	-0.03	-0.06	-0.03
E-(TN)	-0.011	-0.01	-0.021	-0.02	-0.008	-0.02	-0.02	-0.09	-0.01	-0.03	-0.044	-0.032
F-(ER)	-0.01	-0	-0.018	-0.02	0	-0.01	-0.01	-0.07	-0.01	-0.01	-0.017	-0.022
G-(ER)*	-0.015	0	-0.021	-0.01	0.217	-0.02	0.205	0.087	-0.02	-0.04	-0.061	-0.033
POS-H	0.737	1.883	0.699	2.194	2.491	2.171	3.089	2.493	0.688	0.087	1.516	0.949

Table 28: Plate 3:311019

	1	2	3	4	5	6	7	8	9	10	11	12
NEG -A	0.369	0.271	0.243	0.086	0.208	0.212	0.201	0.348	0.22	0.219	0.327	0.1
	0.07	0.241	0.16	0.052	0.13	0.127	0.135	0.075	0.108	0.097	0.214	0.043
B-(TN)	0.342	0.504	0.307	0.071	0.114	0.187	0.123	0.132	0.116	0.156	0.203	0.095
	0.284	0.454	0.233	0.049	0.085	0.116	0.089	0.045	0.045	0.043	0.048	0.04
C-(TN)	0.083	0.056	0.062	0.051	0.067	0.067	0.07	0.103	0.102	0.141	0.185	0.09
	0.047	0.042	0.044	0.04	0.042	0.04	0.041	0.041	0.041	0.041	0.042	0.041
D-(TN)	0.217	0.226	0.059	0.075	0.091	0.07	0.254	0.148	0.111	0.11	0.141	0.269
	0.048	0.045	0.042	0.043	0.044	0.042	0.042	0.042	0.041	0.041	0.041	0.04
E-(TN)	0.056	0.052	0.055	0.055	0.064	0.055	0.107	0.118	0.079	0.109	0.112	0.085
	0.041	0.041	0.042	0.042	0.042	0.041	0.062	0.042	0.041	0.044	0.042	0.041
F-(TN)	0.082	0.097	0.063	0.054	0.07	0.061	0.197	0.147	0.075	0.125	0.181	0.788
	0.05	0.043	0.042	0.041	0.042	0.04	0.042	0.042	0.041	0.041	0.042	0.043
G-(TN)	0.071	0.058	0.059	0.053	0.06	0.054	0.102	0.111	0.085	0.123	0.162	0.106
	0.042	0.044	0.041	0.04	0.044	0.04	0.04	0.04	0.041	0.041	0.042	0.041
POS-H	0.939	3.142	1.179	2.838	2.333	3.013	3.177	2.324	0.939	1.04	2.268	1.013
	0.046	0.061	0.046	0.052	0.051	0.052	0.054	0.049	0.046	0.044	0.05	0.045

Background OD removal

	1	2	3	4	5	6	7	8	9	10	11	12
NEG -A	0.299	0.030	0.083	0.034	0.078	0.085	0.066	0.273	0.112	0.122	0.113	0.057
B-(ER)	0.058	0.05	0.074	0.022	0.029	0.071	0.034	0.087	0.071	0.113	0.155	0.055
C-(ER)	0.036	0.014	0.018	0.011	0.025	0.027	0.029	0.062	0.061	0.1	0.143	0.049
D-(TN)	0.169	0.181	0.017	0.032	0.047	0.028	0.212	0.106	0.07	0.069	0.1	0.229
E-(TN)	0.015	0.011	0.013	0.013	0.022	0.014	0.045	0.076	0.038	0.065	0.07	0.044
F-(ER)	0.032	0.054	0.021	0.013	0.028	0.021	0.155	0.105	0.034	0.084	0.139	0.745
G-(ER)*	0.029	0.014	0.018	0.013	0.016	0.014	0.062	0.071	0.044	0.082	0.12	0.065
POS-H	0.893	3.081	1.133	2.786	2.282	2.961	3.123	2.275	0.893	0.996	2.218	0.968

Neg control removal

	1	2	3	4	5	6	7	8	9	10	11	12
NEG -A	0	0	0	0	0	0	0	0	0	0	0	0
B-(ER)	-0.24	0.02	-0.01	-0.012	-0.049	-0.014	-0.03	-0.186	-0.041	-0.01	0.042	-0.002
C-(ER)	-0.26	-0.016	-0.07	-0.023	-0.053	-0.058	-0.04	-0.211	-0.051	-0.02	0.03	-0.008
D-(TN)	-0.13	0.151	-0.07	-0.002	-0.031	-0.057	0.146	-0.167	-0.042	-0.05	-0.01	0.172
E-(TN)	-0.28	-0.019	-0.07	-0.021	-0.056	-0.071	-0.02	-0.197	-0.074	-0.06	-0.04	-0.013
F-(ER)	-0.27	0.024	-0.06	-0.021	-0.05	-0.064	0.089	-0.168	-0.078	-0.04	0.026	0.688
G-(ER)*	-0.27	-0.016	-0.07	-0.021	-0.062	-0.071	0	-0.202	-0.068	-0.04	0.007	0.008
POS-H	0.594	3.051	1.05	2.752	2.204	2.876	3.057	2.002	0.781	0.874	2.105	0.911

Table 29: Sheet 1 (2):29101

9

	1	2	3	4	5	6	7	8	9	10	11	12
A	0.06	0.05	0.07	0.06	0.07	0.07	0.07	0.15	0.10	0.10	0.13	0.09
	0.04	0.04	0.05	0.05	0.04	0.04	0.04	0.04	0.05	0.04	0.05	0.04
B (TN)	0.06	0.05	0.07	0.05	0.05	0.05	0.06	0.08	0.07	0.08	0.10	0.06
	0.05	0.04	0.06	0.04	0.04	0.04	0.05	0.05	0.04	0.05	0.06	0.06
C (ER)	0.05	0.05	0.05	0.05	0.05	0.05	0.06	0.09	0.10	0.08	0.09	0.07
	0.05	0.05	0.04	0.04	0.04	0.04	0.05	0.06	0.06	0.05	0.05	0.06
D (TP)	0.05	0.08	0.05	0.05	0.06	0.05	0.07	0.08	0.06	0.09	0.09	0.06
	0.04	0.07	0.04	0.04	0.05	0.04	0.05	0.05	0.05	0.05	0.05	0.04
E (ER)	0.06	0.06	0.06	0.05	0.05	1.91	0.06	0.08	0.06	0.08	0.08	0.06
	0.05	0.06	0.05	0.05	0.05	1.97	0.05	0.05	0.04	0.05	0.04	0.04
F (ER)	0.05	0.06	0.06	0.06	0.07	0.07	0.06	0.08	0.05	0.07	0.08	0.06
	0.05	0.06	0.05	0.05	0.06	0.06	0.05	0.05	0.04	0.05	0.05	0.04
G (ER)	0.07	0.06	0.06	0.07	0.05	0.06	0.07	0.08	0.04	0.09	0.10	0.06
	0.06	0.05	0.05	0.05	0.05	0.05	0.05	0.04	0.04	0.04	0.04	0.04
H	1.00	2.92	1.15	2.82	2.72	3.22	0.08	2.11	0.04	0.99	1.75	0.87
	0.05	0.06	0.06	0.07	0.05	0.06	0.04	0.05	0.04	0.04	0.05	0.05

Background OD removal

	1	2	3	4	5	6	7	8	9	10	11	12
A	0.018	0.006	0.024	0.014	0.032	0.032	0.03	0.106	0.044	0.059	0.074	0.045
B (TN)	0.012	0.009	0.011	0.006	0.006	0.007	0.013	0.029	0.029	0.037	0.046	0.007
C (ER)	0.006	0.006	0.007	0.006	0.008	0.009	0.01	0.035	0.046	0.029	0.041	0.015
D (TP)	0.01	0.014	0.007	0.008	0.01	0.008	0.011	0.028	0.014	0.038	0.033	0.019
E (ER)	0.008	0.007	0.007	0.006	0.006	-0.057	0.012	0.026	0.021	0.029	0.036	0.02
F (ER)	0.005	0.007	0.008	0.008	0.01	0.01	0.013	0.029	0.014	0.027	0.031	0.016
G (ER)	0.011	0.008	0.01	0.017	0.009	0.009	0.014	0.039	0.002	0.047	0.056	0.026
H	0.954	2.862	1.097	2.757	2.672	3.166	0.038	2.068	0.003	0.941	1.697	0.815

Neg control removal

	1	2	3	4	5	6	7	8	9	10	11	12
NEG -A	0	0	0	0	0	0	0	0	0	0	0	0
B-142 (ER)	-0.006	0.003	-0.013	-0.008	-0.026	-0.025	-0.017	-0.077	-0.015	-0.022	-0.028	-0.038
C-147 (ER)	-0.012	0	-0.017	-0.008	-0.024	-0.023	-0.02	-0.071	0.002	-0.03	-0.033	-0.03
D-169(TN)	-0.008	0.008	-0.017	-0.006	-0.022	-0.024	-0.019	-0.078	-0.03	-0.021	-0.041	-0.026
E-170 (TN)	-0.01	0.001	-0.017	-0.008	-0.026	-0.089	-0.018	-0.08	-0.023	-0.03	-0.038	-0.025
F-183(ER)	-0.013	0.001	-0.016	-0.006	-0.022	-0.022	-0.017	-0.077	-0.03	-0.032	-0.043	-0.029
G-190(ER)*	-0.007	0.002	-0.014	-0.003	-0.023	-0.023	-0.016	-0.067	-0.042	-0.012	-0.018	-0.019
POS-H	0.936	2.856	1.073	2.743	2.64	3.134	0.008	1.962	-0.041	0.882	1.623	0.77

Table 30: Sheet 2: 291019

	1	2	3	4	5	6	7	8	9	10	11	12
A	0.063	0.048	0.058	0.06	0.131	0.076	0.077	0.126	0.074	0.095	0.13	0.089
	0.04	0.041	0.041	0.041	0.087	0.044	0.041	0.043	0.04	0.041	0.043	0.04
B (TN)	0.053	0.062	0.05	0.048	0.046	0.051	0.059	0.067	0.078	0.067	0.102	0.058
	0.04	0.045	0.044	0.042	0.04	0.045	0.046	0.043	0.052	0.045	0.046	0.044
C(ER)	0.055	0.052	0.049	0.048	0.048	0.05	0.11	0.065	0.074	0.188	0.114	0.058
	0.041	0.045	0.042	0.042	0.042	0.043	0.043	0.041	0.043	0.107	0.061	0.045
D(TP)	0.049	0.051	0.051	0.048	0.05	0.062	0.054	0.062	0.063	0.077	0.081	0.055
	0.044	0.041	0.044	0.042	0.043	0.054	0.041	0.041	0.04	0.049	0.041	0.04
E(ER)	0.053	0.05	0.055	0.048	0.049	0.048	0.051	0.107	0.067	0.071	0.084	0.067
	0.044	0.043	0.046	0.042	0.043	0.042	0.041	0.118	0.047	0.047	0.047	0.049
F (ER)	0.052	0.053	0.053	0.05	0.05	0.063	0.335	0.068	0.058	0.07	0.78	0.076
	0.043	0.045	0.046	0.045	0.044	0.056	0.458	0.046	0.044	0.045	0.636	0.056
G (ER)	0.052	0.132	0.055	0.618	0.247	0.298	0.056	1.086	0.062	0.224	0.127	0.076
	0.042	0.119	0.043	0.524	0.208	0.23	0.042	0.963	0.042	0.129	0.044	0.051
H	0.737	1.651	0.34	1.43	1.233	1.614	2.438	1.935	0.597	0.992	1.668	0.824
	0.043	0.046	0.045	0.045	0.045	0.044	0.049	0.048	0.042	0.044	0.047	0.044

Background OD removal

	1	2	3	4	5	6	7	8	9	10	11	12
A	0.023	0.007	0.017	0.019	0.044	0.032	0.036	0.083	0.034	0.054	0.087	0.049
B (TN)	0.013	0.017	0.006	0.006	0.006	0.006	0.013	0.024	0.026	0.022	0.056	0.014
C(ER)	0.014	0.007	0.007	0.006	0.006	0.007	0.067	0.024	0.031	0.081	0.053	0.013
D(TP)	0.005	0.01	0.007	0.006	0.007	0.008	0.013	0.021	0.023	0.028	0.04	0.015
E(ER)	0.009	0.007	0.009	0.006	0.006	0.006	0.01	-0.011	0.02	0.024	0.037	0.018
F (ER)	0.009	0.008	0.007	0.005	0.006	0.007	-0.12	0.022	0.014	0.025	0.144	0.02
G (ER)	0.01	0.013	0.012	0.094	0.039	0.068	0.014	0.123	0.02	0.095	0.083	0.025
H	0.694	1.605	0.295	1.385	1.188	1.57	2.389	1.887	0.555	0.948	1.621	0.78

Neg control removal

	1	2	3	4	5	6	7	8	9	10	11	12
NEG -A	0	0	0	0	0	0	0	0	0	0	0	0
B-142 (ER)	-0.01	0.01	-0.011	-0.013	-0.038	-0.026	-0.023	-0.059	-0.008	-0.032	-0.031	-0.035
C- 147 (ER)	-0.009	0	-0.01	-0.013	-0.038	-0.025	0.031	-0.059	-0.003	0.027	-0.034	-0.036
D-169(TN)	-0.018	0.003	-0.01	-0.013	-0.037	-0.024	-0.023	-0.062	-0.011	-0.026	-0.047	-0.034
E-170 (TN)	-0.014	0	-0.008	-0.013	-0.038	-0.026	-0.026	-0.094	-0.014	-0.03	-0.05	-0.031
F-183(ER)	-0.014	0.001	-0.01	-0.014	-0.038	-0.025	-0.159	-0.061	-0.02	-0.029	0.057	-0.029
G-190(ER)*	-0.013	0.006	-0.005	0.075	-0.005	0.036	-0.022	0.04	-0.014	0.041	-0.004	-0.024
POS-H	0.671	1.598	0.278	1.366	1.144	1.538	2.353	1.804	0.521	0.894	1.534	0.731

Table 31: Sheet 3: 291019

	1	2	3	4	5	6	7	8	9	10	11	12
--	---	---	---	---	---	---	---	---	---	----	----	----

A	0.069	0.049	0.06	0.061	0.055	0.064	0.075	0.175	0.087	0.107	0.105	0.08
	0.04	0.041	0.04	0.044	0.042	0.042	0.04	0.042	0.041	0.042	0.041	0.04
B (ER)	0.057	0.047	0.048	0.049	0.051	0.053	0.057	0.079	0.082	0.09	0.108	0.058
	0.04	0.04	0.041	0.042	0.043	0.044	0.042	0.042	0.042	0.041	0.047	0.041
C(TN)	0.048	0.051	0.057	0.049	0.06	0.049	0.145	0.061	0.05	0.075	0.101	0.739
	0.039	0.04	0.048	0.041	0.043	0.042	0.042	0.041	0.04	0.044	0.044	0.043
D(TN)	0.049	0.047	0.051	0.049	0.047	0.049	0.051	0.079	0.057	0.073	0.072	0.056
	0.04	0.04	0.044	0.042	0.04	0.042	0.041	0.049	0.041	0.041	0.04	0.041
E(ER)	0.053	0.053	0.05	0.047	0.055	0.049	0.054	0.077	0.066	0.077	0.136	0.059
	0.041	0.046	0.042	0.041	0.047	0.041	0.042	0.045	0.043	0.044	0.074	0.042
F (TN)	0.05	0.048	0.053	0.049	0.062	0.055	0.056	0.081	0.06	0.073	0.092	0.058
	0.04	0.041	0.045	0.043	0.05	0.046	0.042	0.045	0.042	0.044	0.044	0.041
G(ER)	0.052	0.048	0.05	0.047	0.052	0.053	0.062	0.085	0.059	0.075	0.141	0.067
	0.04	0.041	0.042	0.041	0.042	0.043	0.046	0.05	0.043	0.042	0.042	0.044
H	1.113	2.084	0.747	2.133	1.77	2.285	2.976	2.411	0.539	1.237	1.625	0.862
	0.043	0.048	0.041	0.048	0.046	0.049	0.055	0.069	0.041	0.045	0.046	0.043

Background OD removal

	1	2	3	4	5	6	7	8	9	10	11	12
NEG -A	0.029	0.008	0.02	0.017	0.013	0.022	0.035	0.133	0.046	0.065	0.064	0.04
B-(ER)	0.017	0.007	0.007	0.007	0.008	0.009	0.015	0.037	0.04	0.049	0.061	0.017
C-(ER)	0.009	0.011	0.009	0.008	0.017	0.007	0.103	0.02	0.01	0.031	0.057	0.696
D-(TN)	0.009	0.007	0.007	0.007	0.007	0.007	0.01	0.03	0.016	0.032	0.032	0.015
E-(TN)	0.012	0.007	0.008	0.006	0.008	0.008	0.012	0.032	0.023	0.033	0.062	0.017
F-(ER)	0.01	0.007	0.008	0.006	0.012	0.009	0.014	0.036	0.018	0.029	0.048	0.017
G-(ER)*	0.012	0.007	0.008	0.006	0.01	0.01	0.016	0.035	0.016	0.033	0.099	0.023
POS-H	1.07	2.036	0.706	2.085	1.724	2.236	2.921	2.342	0.498	1.192	1.579	0.819

Neg control removal

	1	2	3	4	5	6	7	8	9	10	11	12
NEG -A	0	0	0	0	0	0	0	0	0	0	0	0
B-142 (ER)	-0.012	-0.001	-0.013	-0.01	-0.005	-0.013	-0.02	-0.096	-0.006	-0.016	-0.003	-0.023
C- 147 (ER)	-0.02	0.003	-0.011	0.009	0.004	-0.015	0.068	-0.113	-0.036	-0.034	-0.007	0.656
D-169(TN)	-0.02	-0.001	-0.013	-0.01	-0.006	-0.015	-0.025	-0.103	-0.03	-0.033	-0.032	-0.025
E-170 (TN)	-0.017	-0.001	-0.012	0.011	-0.005	-0.014	-0.023	-0.101	-0.023	-0.032	-0.002	-0.023
F-183(ER)	-0.019	-0.001	-0.012	0.011	-0.001	-0.013	-0.021	-0.097	-0.028	-0.036	-0.016	-0.023
G-190(ER)*	-0.017	-0.001	-0.012	0.011	-0.003	-0.012	-0.019	-0.098	-0.03	-0.032	-0.035	-0.017
POS-H	1.041	2.028	0.686	2.068	1.711	2.214	2.886	2.209	0.452	1.127	1.515	0.779

Table 32: SHEET 4: 311019

	1	2	3	4	5	6	7	8	9	10	11	12
A	0.104	0.089	0.103	0.1	0.107	0.125	0.118	0.334	0.192	0.209	0.2	0.162
	0.04	0.042	0.04	0.041	0.042	0.041	0.042	0.044	0.042	0.042	0.043	0.041
B (ER)	0.071	0.071	0.073	0.065	0.072	0.069	0.081	0.117	0.112	0.11	0.106	0.119
	0.04	0.04	0.039	0.04	0.04	0.04	0.04	0.041	0.04	0.04	0.041	0.039
C(ER)	0.139	0.087	0.078	0.094	0.104	0.107	0.114	0.149	0.121	0.216	0.125	0.148
	0.041	0.04	0.04	0.041	0.04	0.04	0.04	0.041	0.041	0.041	0.041	0.041
D(ER)	0.071	0.073	0.068	0.073	0.169	0.082	0.141	0.105	0.064	0.151	0.123	0.108
	0.039	0.039	0.039	0.038	0.039	0.038	0.038	0.029	0.031	0.029	0.039	0.039
E(ER)	0.065	0.115	0.071	0.071	0.072	0.096	0.102	0.387	0.093	0.115	0.125	0.141
	0.038	0.041	0.041	0.041	0.04	0.04	0.04	0.079	0.04	0.041	0.041	0.041
F (TN)	0.065	0.081	0.069	0.067	0.082	0.178	0.181	0.154	0.122	0.14	0.118	0.235
	0.038	0.04	0.039	0.04	0.035	0.042	0.04	0.041	0.04	0.041	0.04	0.04
G (ER)	0.07	0.073	0.069	0.07	1.445	0.082	0.074	0.13	0.116	0.14	0.128	0.139
	0.039	0.04	0.041	0.04	0.839	0.04	0.04	0.04	0.04	0.041	0.041	0.042
H	1.16	3.108	1.000 0	2.998	2.879	2.824	3.262	2.973	1.024	1.126	2.918	1.505
	0.04	0.054	0.042	0.049	0.05	0.048	0.051	0.048	0.042	0.042	0.049	0.042

Background OD removal

	1	2	3	4	5	6	7	8	9	10	11	12
NEG -A	0.064	0.047	0.063	0.059	0.065	0.084	0.076	0.29	0.15	0.167	0.157	0.121
B-(ER)	0.031	0.031	0.034	0.025	0.032	0.029	0.041	0.076	0.072	0.07	0.065	0.08
C-(ER)	0.098	0.047	0.038	0.053	0.064	0.067	0.074	0.108	0.08	0.175	0.084	0.107
D-(TN)	0.032	0.034	0.029	0.035	0.13	0.044	0.103	0.076	0.033	0.122	0.084	0.069
E-(TN)	0.027	0.074	0.03	0.03	0.032	0.056	0.062	0.308	0.053	0.074	0.084	0.1
F-(ER)	0.027	0.041	0.03	0.027	0.047	0.136	0.141	0.113	0.082	0.099	0.078	0.195
G-(ER)*	0.031	0.033	0.028	0.03	0.606	0.042	0.034	0.09	0.076	0.099	0.087	0.097
POS-H	1.12	3.054	0.958	2.949	2.829	2.776	3.211	2.925	0.982	1.084	2.869	1.463

Neg control removal

	1	2	3	4	5	6	7	8	9	10	11	12
NEG -A	0	0	0	0	0	0	0	0	0	0	0	0
B-142 (ER)	0.033	0.016	0.029	0.034	0.033	0.055	0.035	0.214	0.078	0.097	0.092	0.041
C- 147 (ER)	0.034	0	0.025	0.006	0.001	0.017	0.002	0.182	-0.07	0.008	0.073	0.014
D-169(TN)	0.032	0.013	0.034	0.024	0.065	-0.04	0.027	0.214	0.117	0.045	0.073	0.052
E-170 (TN)	0.037	0.027	0.033	0.029	0.033	0.028	0.014	0.018	0.097	0.093	0.073	0.021
F-183(ER)	0.037	0.006	0.033	0.032	0.018	0.052	0.065	0.177	0.068	0.068	0.079	0.074
G-190(ER)*	0.033	0.014	0.035	0.029	0.541	0.042	0.042	-0.2	0.074	0.068	-0.07	0.024
POS-H	1.056	3.007	0.895	2.89	2.764	2.692	3.135	2.635	0.832	0.917	2.712	1.342

Table 33: Sheet 5: 311019

	1	2	3	4	5	6	7	8	9	10	11	12
A	0.222	0.279	0.129	0.211	0.185	0.501	0.167	0.238	0.239	0.297	0.415	0.247
	0.04	0.041	0.041	0.041	0.042	0.042	0.044	0.041	0.041	0.043	0.046	0.032
B(TN)	0.102	0.174	0.102	0.154	0.148	0.169	0.144	0.206	0.178	1.222	0.383	0.213
	0.034	0.041	0.04	0.041	0.04	0.04	0.04	0.039	0.04	0.044	0.045	0.038
C (TN)	0.187	0.217	0.129	0.154	0.148	0.104	0.236	0.175	0.182	0.21	0.221	0.221
	0.041	0.042	0.039	0.04	0.04	0.04	0.041	0.04	0.04	0.04	0.04	0.039
D(ER)	0.096	0.094	0.137	0.169	0.266	0.079	0.511	0.196	0.202	0.361	0.238	0.937
	0.038	0.04	0.041	0.04	0.04	0.038	0.04	0.04	0.039	0.039	0.037	0.036
E(ER)	0.109	0.193	0.176	0.141	0.1	0.181	0.233	0.254	0.404	0.204	0.163	0.189
	0.039	0.041	0.04	0.041	0.041	0.041	0.041	0.041	0.04	0.044	0.04	0.041
F(ER)	0.107	0.392	0.236	0.187	0.657	0.131	0.442	0.255	0.18	0.15	0.187	0.882
	0.039	0.041	0.041	0.041	0.049	0.042	0.041	0.042	0.042	0.041	0.041	0.041
G (ER)	0.067	0.08	0.105	0.241	0.162	0.11	0.127	0.376	0.293	0.305	0.302	0.202
	0.038	0.039	0.039	0.039	0.04	0.04	0.042	0.04	0.04	0.041	0.04	0.04
H	1.045	2.45	0.962	2.75	2.618	2.706	3.166	2.746	0.998	1.688	1.583	1.302
	0.042	0.047	0.04	0.048	0.049	0.049	0.051	0.049	0.041	0.042	0.043	0.041

Background OD removal

	1	2	3	4	5	6	7	8	9	10	11	12
NEG -A	0.182	0.238	0.088	0.17	0.143	0.459	0.123	0.197	0.198	0.254	0.369	0.215
B-(ER)	0.068	0.133	0.062	0.113	0.108	0.129	0.104	0.167	0.138	1.178	0.338	0.175
C-(ER)	0.146	0.175	0.09	0.114	0.108	0.064	0.195	0.135	0.142	0.17	0.181	0.182
D- (TN)	0.058	0.054	0.096	0.129	0.226	0.041	0.471	0.156	0.163	0.322	0.201	0.901
E-(TN)	0.07	0.152	0.136	0.1	0.059	0.14	0.192	0.213	0.364	0.16	0.123	0.148
F-(ER)	0.068	0.351	0.195	0.146	0.608	0.089	0.401	0.213	0.138	0.109	0.146	0.841
G-(ER)*	0.029	0.041	0.066	0.202	0.122	0.07	0.085	0.336	0.253	0.264	0.262	0.162
POS-H	1.003	2.403	0.922	2.702	2.569	2.657	3.115	2.697	0.957	1.646	1.54	1.261

Neg control removal

	1	2	3	4	5	6	7	8	9	10	11	12
NEG -A	0	0	0	0	0	0	0	0	0	0	0	0
B-142 (ER)	-0.114	-0.105	-0.026	-0.057	-0.035	-0.33	-0.019	-0.03	-0.06	0.924	-0.031	-0.04
C- 147 (ER)	-0.036	-0.063	0.002	-0.056	-0.035	-0.395	0.072	-0.062	-0.056	-0.084	-0.188	-0.033
D-169(TN)	-0.124	-0.184	0.008	-0.041	0.083	-0.418	0.348	-0.041	-0.035	0.068	-0.168	0.686
E-170 (TN)	-0.112	-0.086	0.048	-0.07	-0.084	-0.319	0.069	0.016	0.166	-0.094	-0.246	-0.067
F-183(ER)	-0.114	0.113	0.107	-0.024	0.465	-0.37	0.278	0.016	-0.06	-0.145	-0.223	0.626
G-190(ER)*	-0.153	-0.197	-0.022	0.032	-0.021	-0.389	-0.038	0.139	0.055	0.01	-0.107	-0.053
POS-H	0.821	2.165	0.834	2.532	2.426	2.198	2.992	2.5	0.759	1.392	1.171	1.046

Table 34: SHEET 6: 311019

	1	2	3	4	5	6	7	8	9	10	11	12
A	0.09	0.052	0.075	0.049	0.059	0.07	0.084	0.186	0.083	0.119	0.111	0.083
	0.039	0.041	0.041	0.026	0.028	0.03	0.033	0.029	0.038	0.043	0.041	0.04
B(TN)	0.071	0.049	0.05	0.048	0.048	0.051	0.056	0.082	0.092	0.094	0.106	0.072
	0.037	0.04	0.04	0.039	0.039	0.039	0.039	0.04	0.039	0.046	0.039	0.038
C (TN)	0.051	0.06	0.047	0.047	0.051	0.05	0.068	0.076	0.057	0.081	0.074	0.076
	0.039	0.04	0.039	0.039	0.041	0.039	0.045	0.04	0.04	0.04	0.039	0.039
D(TN)	0.053	0.047	0.051	0.046	0.063	0.048	0.065	0.084	0.064	0.077	0.084	0.074
	0.04	0.039	0.041	0.035	0.04	0.038	0.039	0.038	0.039	0.041	0.039	0.039
E(TN)	0.057	0.055	0.051	0.052	0.061	0.057	0.066	0.081	0.063	0.077	0.088	0.07
	0.042	0.042	0.041	0.041	0.042	0.042	0.04	0.041	0.041	0.041	0.042	0.04
F (187-TN)	0.056	0.048	0.05	0.049	0.052	0.057	0.054	0.097	0.059	0.076	0.086	0.071
	0.039	0.04	0.041	0.041	0.04	0.041	0.04	0.042	0.04	0.041	0.042	0.042
G(TN)	0.052	0.05	0.05	0.049	0.053	0.054	0.056	0.081	0.06	0.073	0.083	0.084
	0.039	0.04	0.04	0.04	0.041	0.043	0.04	0.04	0.041	0.043	0.041	0.041
H	1.891	2.754	1.675	3.085	2.061	3.05	3.467	3.093	0.957	1.549	2.457	0.648
	0.052	0.049	0.045	0.058	0.068	0.053	0.085	0.054	0.044	0.035	0.05	0.041

Background OD removal

	1	2	3	4	5	6	7	8	9	10	11	12
A	0.051	0.011	0.034	0.023	0.031	0.04	0.051	0.157	0.045	0.076	0.07	0.043
B(TN)	0.034	0.009	0.01	0.009	0.009	0.012	0.017	0.042	0.053	0.048	0.067	0.034
C (TN)	0.012	0.02	0.008	0.008	0.01	0.011	0.023	0.036	0.017	0.041	0.035	0.037
D(TN)	0.013	0.008	0.01	0.011	0.023	0.01	0.026	0.046	0.025	0.036	0.045	0.035
E(TN)	0.015	0.013	0.01	0.011	0.019	0.015	0.026	0.04	0.022	0.036	0.046	0.03
F (TN)	0.017	0.008	0.009	0.008	0.012	0.016	0.014	0.055	0.019	0.035	0.044	0.029
G(TN)	0.013	0.01	0.01	0.009	0.012	0.011	0.016	0.041	0.019	0.03	0.042	0.043
H	1.839	2.705	1.63	3.027	1.993	2.997	3.382	3.039	0.913	1.514	2.407	0.607

Neg control removal

	1	2	3	4	5	6	7	8	9	10	11	12
A	0	0	0	0	0	0	0	0	0	0	0	0
B(TN)	-0.017	-0.002	-0.02	-0.01	-0.02	-0.03	-0.034	-0.12	0.008	-0.03	-0.003	-0.009
C (TN)	-0.039	0.009	-0.03	-0.02	-0.02	-0.03	-0.028	-0.12	-0.028	-0.04	-0.035	-0.006
D(TN)	-0.038	-0.003	-0.02	-0.01	-0.01	-0.03	-0.025	-0.11	-0.02	-0.04	-0.025	-0.008
E(TN)	-0.036	0.002	-0.02	-0.01	-0.01	-0.03	-0.025	-0.12	-0.023	-0.04	-0.024	-0.013
F (TN)	-0.034	-0.003	-0.03	-0.02	-0.02	-0.02	-0.037	-0.1	-0.026	-0.04	-0.026	-0.014
G(TN)	-0.038	-1E-03	-0.02	-0.01	-0.02	-0.03	-0.035	-0.12	-0.026	-0.05	-0.028	0
H	1.788	2.694	1.596	3.004	1.962	2.957	3.331	2.882	0.868	1.438	2.337	0.564

Table 34: Issue Plate: 311019

	1	2	3	4	5	6	7	8	9	10	11	12
A	0.072	0.066	0.069	0.052	0.053	0.083	0.07	0.262	0.115	0.173	0.251	0.088
	0.04	0.041	0.047	0.042	0.041	0.041	0.047	0.042	0.043	0.044	0.044	0.04
B(TN)	0.056	0.053	0.054	0.054	0.058	0.064	0.061	0.083	0.095	0.105	0.195	0.079
	0.041	0.041	0.041	0.041	0.04	0.04	0.04	0.041	0.04	0.041	0.041	0.041
C(TN)	0.06	0.059	0.052	0.053	0.056	0.064	0.055	0.085	0.078	0.099	0.103	0.076
	0.04	0.042	0.041	0.042	0.042	0.04	0.04	0.04	0.04	0.041	0.04	0.04
D(TN)	0.049	0.052	0.05	0.052	0.049	0.05	0.051	0.086	0.069	0.09	0.09	0.083
	0.038	0.041	0.04	0.04	0.039	0.039	0.039	0.04	0.041	0.042	0.039	0.038
E (TN)	0.05	0.054	0.054	0.057	0.051	0.05	0.071	0.079	0.07	0.113	0.144	0.076
	0.045	0.039	0.04	0.046	0.04	0.04	0.046	0.041	0.04	0.042	0.041	0.04
F (TN)	0.056	0.057	0.052	0.059	0.061	0.054	0.095	0.086	0.055	0.094	0.107	0.107
	0.04	0.04	0.04	0.047	0.039	0.04	0.04	0.041	0.04	0.045	0.041	0.047
G (TN)	0.049	0.05	0.051	0.049	0.053	0.094	0.067	0.102	0.09	0.124	0.169	0.08
	0.04	0.039	0.039	0.039	0.04	0.047	0.039	0.041	0.04	0.04	0.04	0.041
H	0.068	0.066	0.069	0.068	0.063	0.09	0.071	0.252	0.111	0.172	0.2	0.113
	0.039	0.038	0.042	0.039	0.039	0.038	0.038	0.039	0.04	0.04	0.04	0.039

Background OD removal

	1	2	3	4	5	6	7	8	9	10	11	12
NEG -A	0.032	0.025	0.022	0.01	0.012	0.042	0.023	0.22	0.072	0.129	0.207	0.048
B-(ER)	0.015	0.012	0.013	0.013	0.018	0.024	0.021	0.042	0.055	0.064	0.154	0.038
C-(ER)	0.02	0.017	0.011	0.011	0.014	0.024	0.015	0.045	0.038	0.058	0.063	0.036
D-(TN)	0.011	0.011	0.01	0.012	0.01	0.011	0.012	0.046	0.028	0.048	0.051	0.045
E-(TN)	0.005	0.015	0.014	0.011	0.011	0.01	0.025	0.038	0.03	0.071	0.103	0.036
F-(ER)	0.016	0.017	0.012	0.012	0.022	0.014	0.055	0.045	0.015	0.049	0.066	0.06
G-(ER)*	0.009	0.011	0.012	0.01	0.013	0.047	0.028	0.061	0.05	0.084	0.129	0.039
POS-H	0.029	0.028	0.027	0.029	0.024	0.052	0.033	0.213	0.071	0.132	0.16	0.074

Neg control removal

	1	2	3	4	5	6	7	8	9	10	11	12
NEG -A	0	0	0	0	0	0	0	0	0	0	0	0
B-142 (ER)	-0.017	-0.013	-0.009	0.003	0.006	-0.018	-0.002	-0.178	-0.017	-0.065	-0.053	-0.01
C- 147 (ER)	-0.012	-0.008	-0.011	0.001	0.002	-0.018	-0.008	-0.175	-0.034	-0.071	-0.144	-0.012
D-169(TN)	-0.021	-0.014	-0.012	0.002	-0.002	-0.031	-0.011	-0.174	-0.044	-0.081	-0.156	-0.003
E-170 (TN)	-0.027	-0.01	-0.008	0.001	-0.001	-0.032	0.002	-0.182	-0.042	-0.058	-0.104	-0.012
F-183(ER)	-0.016	-0.008	-0.01	0.002	0.01	-0.028	0.032	-0.175	-0.057	-0.08	-0.141	0.012
G-190(ER)*	-0.023	-0.014	-0.01	0	0.001	0.005	0.005	-0.159	-0.022	-0.045	-0.078	-0.009
POS-H	-0.003	0.003	0.005	0.019	0.012	0.01	0.01	-0.007	-0.001	0.003	-0.047	0.026

Table 35: Sheet 8: 311019

	1	2	3	4	5	6	7	8	9	10	11	12
A	0.051	0.045	0.053	0.048	0.065	0.078	0.072	0.066	0.061	0.07	0.077	0.051
	0.029	0.031	0.03	0.032	0.031	0.037	0.041	0.038	0.031	0.041	0.04	0.031
B (ER)	0.05	0.033	0.035	0.043	0.043	0.048	0.046	0.051	0.054	0.056	0.112	0.071
	0.029	0.029	0.029	0.039	0.031	0.039	0.039	0.04	0.03	0.03	0.043	0.039
C (TN)	0.043	0.042	0.033	0.045	0.046	0.036	0.046	0.055	0.058	0.071	0.094	0.063
	0.03	0.036	0.029	0.04	0.039	0.028	0.039	0.041	0.04	0.036	0.041	0.039
D(TN)	0.035	0.035	0.033	0.032	0.055	0.048	0.235	0.051	0.048	0.064	0.069	0.453
	0.031	0.031	0.029	0.029	0.029	0.041	0.042	0.029	0.041	0.039	0.029	0.04
E (TN)	0.043	0.035	0.034	0.034	0.039	0.046	0.076	0.053	0.068	0.057	0.059	0.109
	0.038	0.032	0.03	0.031	0.031	0.04	0.042	0.041	0.042	0.043	0.042	0.038
F(TN)	0.047	0.091	0.048	0.046	0.045	0.051	0.056	0.059	0.05	0.052	0.081	0.068
	0.037	0.03	0.035	0.04	0.04	0.033	0.04	0.042	0.039	0.042	0.041	0.04
G (ER)	0.036	0.032	0.034	0.123	0.033	0.046	0.05	0.051	0.049	0.042	0.049	0.049
	0.031	0.029	0.029	0.029	0.029	0.041	0.04	0.04	0.039	0.034	0.031	0.038
H	0.321	1.131	0.301	1.157	1.293	1.235	1.172	0.033	0.039	0.414	0.469	0.324
	0.03	0.033	0.032	0.034	0.045	0.034	0.034	0.03	0.036	0.033	0.04	0.037

Background OD removal

	1	2	3	4	5	6	7	8	9	10	11	12
NEG -A	0.022	0.014	0.023	0.016	0.034	0.041	0.031	0.028	0.03	0.029	0.037	0.02
B-(ER)	0.021	0.004	0.006	0.004	0.012	0.009	0.007	0.011	0.024	0.026	0.069	0.032
C-(ER)	0.013	0.006	0.004	0.005	0.007	0.008	0.007	0.014	0.018	0.035	0.053	0.024
D-(TN)	0.004	0.004	0.004	0.003	0.026	0.007	0.193	0.022	0.007	0.025	0.04	0.413
E-(TN)	0.005	0.003	0.004	0.003	0.008	0.006	0.034	0.012	0.026	0.014	0.017	0.071
F-(ER)	0.01	0.061	0.013	0.006	0.005	0.018	0.016	0.017	0.011	0.01	0.04	0.028
G-(ER)*	0.005	0.003	0.005	0.094	0.004	0.005	0.01	0.011	0.01	0.008	0.018	0.011
POS-H	0.291	1.098	0.269	1.123	1.248	1.201	1.138	0.003	0.003	0.381	0.429	0.287

Neg control removal

	1	2	3	4	5	6	7	8	9	10	11	12
NEG -A	0	0	0	0	0	0	0	0	0	0	0	0
B-142 (ER)	-0.001	-0.01	-0.017	-0.012	-0.022	-0.032	-0.024	-0.017	-0.006	-0.003	0.032	0.012
C- 147 (ER)	-0.009	-0.008	-0.019	-0.011	-0.027	-0.033	-0.024	-0.014	-0.012	0.006	0.016	0.004
D-169(TN)	-0.018	-0.01	-0.019	-0.013	-0.008	-0.034	0.162	-0.006	-0.023	-0.004	0.003	0.393
E-170 (TN)	-0.017	-0.011	-0.019	-0.013	-0.026	-0.035	0.003	-0.016	-0.004	-0.015	-0.02	0.051
F-183(ER)	-0.012	0.047	-0.01	-0.01	-0.029	-0.023	-0.015	-0.011	-0.019	-0.019	0.003	0.008
G- 190(ER)*	-0.017	-0.011	-0.018	0.078	-0.03	-0.036	-0.021	-0.017	-0.02	-0.021	-0.019	-0.009
POS-H	0.269	1.084	0.246	1.107	1.214	1.16	1.107	-0.025	-0.027	0.352	0.392	0.267

Table 36: Sheet 9: 311019

	1	2	3	4	5	6	7	8	9	10	11	12
A	0.07	0.05	0.093	0.086	0.064	0.104	0.103	0.184	0.102	0.105	0.155	0.092
	0.04	0.04	0.041	0.04	0.041	0.041	0.041	0.042	0.042	0.041	0.042	0.047
B (ER)	0.049	0.048	0.048	0.049	0.054	0.05	0.173	0.086	0.051	0.069	0.093	0.214
	0.038	0.04	0.04	0.039	0.04	0.039	0.04	0.04	0.04	0.04	0.04	0.031
C (TN)	0.042	0.054	0.046	0.047	0.048	0.051	0.053	0.073	0.041	0.044	0.041	0.041
	0.039	0.039	0.039	0.039	0.04	0.039	0.04	0.04	0.039	0.043	0.039	0.039
D (TN)	0.058	0.064	0.056	0.062	0.048	0.088	0.06	0.081	0.09	0.098	0.108	0.073
	0.038	0.04	0.039	0.04	0.039	0.039	0.04	0.039	0.039	0.039	0.039	0.039
E (ER)	0.052	0.049	0.048	0.05	0.052	0.054	0.055	0.067	0.061	0.079	0.066	0.061
	0.04	0.04	0.04	0.041	0.042	0.041	0.042	0.041	0.04	0.031	0.031	0.033
F(TN)	0.053	0.051	0.049	0.049	0.049	0.05	0.081	0.072	0.066	0.071	0.077	0.065
	0.031	0.04	0.04	0.04	0.04	0.041	0.042	0.041	0.041	0.041	0.031	0.037
G (6221/19-TN)	0.055	0.054	0.053	0.056	0.058	0.054	0.178	0.081	0.055	0.077	0.108	0.308
	0.04	0.04	0.045	0.04	0.04	0.04	0.042	0.04	0.041	0.041	0.041	0.042
H	0.891	2.117	0.712	2.238	1.788	2.084	2.78	1.681	0.561	0.669	1.244	0.663
	0.042	0.047	0.04	0.048	0.046	0.047	0.052	0.046	0.042	0.042	0.041	0.043

Background OD removal

	1	2	3	4	5	6	7	8	9	10	11	12
NEG -A	0.03	0.01	0.052	0.046	0.023	0.063	0.062	0.142	0.06	0.064	0.113	0.045
B-(ER)	0.011	0.008	0.008	0.01	0.014	0.011	0.133	0.046	0.011	0.029	0.053	0.183
C-(ER)	0.003	0.015	0.007	0.008	0.008	0.012	0.013	0.033	0.002	0.001	0.002	0.002
D-(TN)	0.02	0.024	0.017	0.022	0.009	0.049	0.02	0.042	0.051	0.059	0.069	0.034
E-(TN)	0.012	0.009	0.008	0.009	0.01	0.013	0.013	0.026	0.021	0.048	0.035	0.028
F-(ER)	0.022	0.011	0.009	0.009	0.009	0.009	0.039	0.031	0.025	0.03	0.046	0.028
G-(ER)*	0.015	0.014	0.008	0.016	0.018	0.014	0.136	0.041	0.014	0.036	0.067	0.266
POS-H	0.849	2.07	0.672	2.19	1.742	2.037	2.728	1.635	0.519	0.627	1.203	0.62

Neg control removal

	1	2	3	4	5	6	7	8	9	10	11	12
NEG -A	0	0	0	0	0	0	0	0	0	0	0	0
B-142 (ER)	-0.019	-0.002	-0.044	-0.036	-0.009	-0.052	0.071	-0.096	-0.049	-0.035	-0.06	0.138
C- 147 (ER)	-0.027	-0.005	-0.045	-0.038	-0.015	-0.051	-0.049	-0.109	-0.058	-0.063	-0.111	-0.043
D-169(TN)	-0.01	-0.014	-0.035	-0.024	-0.014	-0.014	-0.042	-0.1	-0.009	-0.005	-0.044	-0.011
E-170 (TN)	-0.018	-0.001	-0.044	-0.037	-0.013	-0.05	-0.049	-0.116	-0.039	-0.016	-0.078	-0.017
F-183(ER)	-0.008	-0.001	-0.043	-0.037	-0.014	-0.054	-0.023	-0.111	-0.035	-0.034	-0.067	-0.017
G-190(ER)*	-0.015	-0.004	-0.044	-0.03	-0.005	-0.049	-0.074	-0.101	-0.046	-0.028	-0.046	-0.221
POS-H	0.819	2.06	0.62	2.144	1.719	1.974	2.666	1.493	0.459	0.563	1.09	0.575

Table 37: PLATE 2: SEPT. 2019

	1	2	3	4	5	6	7	8	9	10	11	12
NEG -A	0.054	0.047	0.046	0.052	0.05	0.066	0.062	0.178	0.078	0.096	0.108	0.1
	0.04	0.039	0.39	0.04	0.04	0.036	0.041	0.042	0.041	0.042	0.046	0.04
B	0.05	0.054	0.048	0.052	0.057	0.05	0.24	0.081	0.053	0.074	0.153	0.146
	0.039	0.043	0.041	0.042	0.04	0.041	0.178	0.041	0.04	0.039	0.055	0.039
C	0.051	0.051	0.047	0.047	0.046	0.047	0.064	0.076	0.057	0.068	0.069	0.059
	0.042	0.042	0.04	0.039	0.039	0.04	0.051	0.042	0.042	0.043	0.041	0.042
D	0.056	0.053	0.045	0.057	0.051	0.05	0.093	0.077	0.052	0.073	0.099	0.113
	0.042	0.042	0.038	0.049	0.038	0.041	0.039	0.04	0.04	0.043	0.04	0.042
E	0.046	0.049	0.043	0.044	0.054	0.057	0.048	0.071	0.053	0.067	0.072	0.06
	0.039	0.041	0.037	0.039	0.047	0.053	0.039	0.04	0.042	0.043	0.042	0.044
F	0.056	0.06	0.058	0.046	0.056	0.046	0.069	0.075	0.059	0.072	0.059	0.049
	0.044	0.042	0.048	0.039	0.046	0.039	0.04	0.042	0.041	0.041	0.042	0.045
G	0.061	0.052	0.051	0.046	0.05	0.046	0.056	0.085	0.052	0.07	0.078	0.068
	0.042	0.042	0.041	0.039	0.041	0.039	0.039	0.059	0.039	0.041	0.045	0.046
POS-H	0.041	2.761	1.457	2.737	3.119	1.036	3.371	3.38	0.772	1.288	1.416	0.05
	0.039	0.06	0.054	0.049	0.054	0.255	0.051	0.065	0.053	0.044	0.044	0.039

Background OD removal

	1	2	3	4	5	6	7	8	9	10	11	12
NEG -A	0.014	0.008	-0.344	0.012	0.01	0.03	0.021	0.136	0.037	0.054	0.062	0.06
B-(ER)	0.011	0.011	0.007	0.01	0.017	0.009	0.062	0.04	0.013	0.035	0.098	0.107
C-(ER)	0.009	0.009	0.007	0.008	0.007	0.007	0.013	0.034	0.015	0.025	0.028	0.017
D-(TN)	0.014	0.011	0.007	0.008	0.013	0.009	0.054	0.037	0.012	0.03	0.059	0.071
E-(TN)	0.007	0.008	0.006	0.005	0.007	0.004	0.009	0.031	0.011	0.024	0.03	0.016
F-183(ER)	0.012	0.018	0.01	0.007	0.01	0.007	0.029	0.033	0.018	0.031	0.017	0.004
G-190(ER)*	0.019	0.01	0.01	0.007	0.009	0.007	0.017	0.026	0.013	0.029	0.033	0.022
POS-H	0.002	2.701	1.403	2.688	3.065	0.781	3.32	3.315	0.719	1.244	1.372	0.011

Neg control removal

	1	2	3	4	5	6	7	8	9	10	11	12
NEG -A	0	0	0	0	0	0	0	0	0	0	0	0
B-142 (ER)	-0.003	0.003	0.351	-0.002	0.007	-0.021	0.041	-0.096	-0.024	-0.019	0.036	0.047
C- 147 (ER)	-0.005	0.001	0.351	-0.004	0.003	-0.023	0.008	-0.102	-0.022	-0.029	0.034	0.043
D-169(TN)	0	0.003	0.351	-0.004	0.003	-0.021	0.033	-0.099	-0.025	-0.024	-0.003	0.011
E-170 (TN)	-0.007	0	0.35	-0.007	0.003	-0.026	0.012	-0.105	-0.026	-0.03	-0.032	0.044
F-183(ER)	-0.002	0.01	0.354	-0.005	0	-0.023	0.008	-0.103	-0.019	-0.023	0.045	0.056
G-190(ER)*	0.005	0.002	0.354	-0.005	0.001	-0.023	0.004	-0.11	-0.024	-0.025	0.029	0.038
POS-H	-0.012	2.693	1.747	2.676	3.055	0.751	3.299	3.179	0.682	1.19	1.31	-0.049

To get the final data for analysis, we assessed the fold difference in each of the cytokines against the tumour subtype.