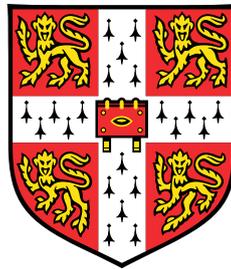


Analysing the relationships between immune infiltration and tissue architecture in High Grade Serous Ovarian Cancer



Rowan Barker-Clarke

Supervisor: Prof. James Brenton

Cancer Research UK Cambridge Institute
University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

I would like to dedicate this thesis to my loving family; my parents and my sister. Also to my grandma and grandad; key influences upon me who passed away during the years of this project. I hope I would have made them incredibly proud. . .

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This dissertation contains fewer than 65,000 words including appendices, bibliography, footnotes, tables and equations and has fewer than 150 figures.

Rowan Barker-Clarke
January 2020

Acknowledgements

And I would like to acknowledge Sarwah, Tedi, Ania, Maria and Carolin along with the rest of the Brenton lab for their academic support with this thesis and being open to discuss ideas and help this ex-physicist get lab work done. Thank you to the core facilities and in particular Jodie Miller and Fadwa Joud in the Cambridge Institute for their unparalleled support of institute staff such as myself.

I thank my supervisor, James Brenton, without whom I would never have taken on this challenge in the first place. I would also like to acknowledge Cancer Research UK for their funding of this PhD.

I thank my parents who have inspired and supported me throughout my education and I thank my best friend Charlotte whose support over the last ten years has made me a better scientist.

Lastly I would like to thank Luke, for tolerating all the long nights and supporting me throughout. Without him, writing this thesis would have been so much more difficult, Luke has been my rock on this journey and will be my side for all my future ones.

Abstract

The aim of this project was to understand the distribution and compartmental effects of immune populations and the tissue architecture that underlies them in high grade serous ovarian cancer (HGSOC).

The quantitative distribution of three key immune infiltrates was initially assessed in the SEARCH cohort (n=332); CD8⁺ cytotoxic T-cells, CD45RO⁺ memory T-cells and CD68⁺ macrophages. The stromal infiltration of CD68⁺ (HR = 0.44, $p = 3 \times 10^{-4}$) and CD45RO⁺ (HR = 0.76, $p = 7 \times 10^{-4}$) cells was positively associated with survival and the positive prognostic value of CD8⁺ cells increased when the density was averaged across both stroma and tumour compartments (HR = 0.79, $p = 6 \times 10^{-4}$).

Multiple correlated immune infiltrates were combined to view patterns across samples. This analysis was limited to the subset of patients with both tumour and stroma present (n=152). A multi-dimensional view of these immune populations (an 'immunospace') was obtained and the principal components of this space were analysed and used to build complex survival models. The largest variation between patient samples was the strength of a coordinated multi-cell immune response that was associated with improved survival (HR = 0.88, $p = 0.016$).

Tumour architecture was investigated by deriving metrics from the positions of epithelial cells in samples in the OV04 cohort (n=101). Both multiplexed pancytokeratin (CK), FOXP3 and CD8 staining and H&E serial sections were used for the localisation of epithelial cells. The nearest neighbour distance between epithelial cells was significantly correlated across serial sections ($R = 0.65$, $p = 2.7 \times 10^{-5}$), between sections from different cores from the same tissue block ($R = 0.58$, $p = 0.00064$) and sections from differing tumour sites ($R = 0.78$, $p = 7.8 \times 10^{-5}$). This measure of epithelial cell packing is therefore tumour-intrinsic. Correlations were also examined between these structural metrics and the FOXP3 and CD8 immune infiltration. Epithelial CD8⁺ density was strongly correlated with epithelial nearest neighbour distance ($R = 0.91$, $p = 1 \times 10^{-14}$) and moderately with border cell percentage and epithelial fraction of the tumour. Regulatory T-cell (FOXP3⁺) infiltration was positively correlated with the number of epithelial clusters across a core.

The conservation of structural measures across tumours and the link between structure and immune infiltrate supported the physical border hypothesis of immune exclusion. It was hypothesised that the link between these two features was collagen deposition. Collagen structure was assessed via dual IF/SHG imaging. Structural metrics were correlated with collagen Energy, Correlation and Coherency features. There was no correlation between these collagen features and immune infiltration.

Imaging Mass Cytometry was carried out on the ICON7 cohort (n=313) to produce a novel data set. Preliminary analyses of these images were carried out to investigate further investigate links between collagen, infiltrates and structure. The thesis presents novel data and numerous novel conclusions regarding the prognostic value of specific immune infiltrate variables and quantitative links between tissue structure and immune response in HGSOC.

Table of contents

List of figures	xvii
List of tables	xxi
Nomenclature	xxiii
1 Introduction	1
1.1 High Grade Serous Ovarian Cancer	1
1.1.1 Types of ovarian cancer	1
1.2 The hallmarks of cancer in HGSOC	2
1.3 The genome in HGSOC	2
1.4 The microenvironment in HGSOC	3
1.4.1 Immune micro-environment	4
1.4.2 Extracellular matrix and stromal micro-environment	6
1.5 Immune recruitment and control	7
1.6 Subtyping HGSOC	8
1.6.1 Genomic subtypes	9
1.6.2 Gene expression sub-types	9
1.6.3 Morphological sub-types	10
1.6.4 Immune subtyping	12
1.7 Methodologies for the imaging of the microenvironment	13
1.7.1 Tissue microarrays (TMAs)	13
1.7.2 Haematoxylin & Eosin (H&E)	13
1.7.3 Immunohistochemistry (IHC)	13
1.7.4 Immunofluorescence (IF)	14
1.7.5 Increasing the number of markers	14
1.7.6 Second-harmonic generation (SHG)	15
1.8 Methodologies for the image analysis of the microenvironment	15

1.8.1	Nuclei segmentation	15
1.8.2	Tissue Classification	15
1.8.3	Spatial Immune Analysis	16
1.9	Project Hypotheses, Outline and Motivations	17
2	Methods	19
2.1	External Contributions	19
2.2	Patient Cohorts	20
2.3	Chromogenic Immunohistochemistry (IHC)	21
2.3.1	Sample preparation	21
2.3.2	Manual Staining	21
2.3.3	Automated Staining	22
2.4	Second Harmonic Generation	24
2.5	Dual IF	24
2.6	Imaging Mass Cytometry	25
2.6.1	Imaging Mass Cytometry (IMC) Antibody purification	25
2.6.2	Staining and Imaging	26
2.6.3	Image analysis	26
2.7	Statistical Analysis and Mathematical Background	28
2.7.1	Analysis of immune infiltrate	28
2.7.2	Survival Analysis	29
2.7.3	Spatial Statistics	29
2.7.4	Density based analysis	30
2.7.5	Distance based analysis	31
2.7.6	Intensity analysis	32
3	Digital pathology and the quantification of immune infiltrate	35
3.1	Introduction	35
3.1.1	History of the project	36
3.2	SEARCH Cohort	36
3.3	Methods	37
3.3.1	Experimental methodology	37
3.3.2	Statistical analyses	38
3.4	Results	39
3.4.1	Patient characteristics	39
3.4.2	Data cleaning, quality assurance and spatial bias	40
3.4.3	Stroma and tumour composition of cores	43

3.4.4	Mutant allele fraction and epithelial content	44
3.4.5	Immune cell densities	46
3.4.6	Immune exclusion	47
3.4.7	Functional form of infiltrates in building a survival model	49
3.4.8	Prognostic value of individual infiltrates	50
3.4.9	Averaging immune infiltrate over a core	52
3.4.10	Epithelial core subset survival analysis	54
3.4.11	Combined model	54
3.4.12	Is the immune exclusion ratio of intra-epithelial infiltrate to stromal infiltrate prognostic?	56
3.4.13	Principal components and the combined immunospace	57
3.4.14	Comparing Survival Models	61
3.4.15	Are genetic defects associated with HGSOC driving individual infiltrates or the coordinated immune response in tumours?	62
3.4.16	Other ovarian cancer subtypes	64
3.5	Discussion	67
4	Quantifying and classifying tissue structure	71
4.1	Introduction	71
4.2	Methodology	72
4.2.1	Patient Cohorts	72
4.2.2	IHC	72
4.2.3	H&E	72
4.2.4	Statistical Analysis	73
4.3	Results	73
4.3.1	Patient Characteristics	73
4.3.2	Cell classifier performance and comparison between H&E and cytokeratin stained images	74
4.3.3	Metrics for tissue structure analysis	77
4.3.4	Stroma/Epithelium percentage	77
4.3.5	Epithelial Nearest Neighbour Distance	78
4.3.6	Quantifying border cells and the invasive front	79
4.3.7	Entropy	85
4.3.8	Cell Clustering	87
4.3.9	Correlations between structural metrics	91
4.3.10	Principal components	93
4.3.11	Morphology classification	93

4.3.12	Relationship between Structure and Immune infiltration	94
4.3.13	CD8 ⁺ Correlation	94
4.3.14	FOXP3 ⁺ Correlation	94
4.4	Discussion	98
5	Collagen as a potential structural driver and influence on immune infiltration	103
5.1	Introduction	103
5.1.1	History of the project and Collaborator roles	104
5.2	Methodology	105
5.2.1	Cohort Summary	105
5.2.2	SHG	107
5.2.3	Dual IF	107
5.2.4	Immuno Metal Conjugation Marker Panel	107
5.2.5	IMC Staining and Imaging	109
5.2.6	Image Analysis	109
5.3	Results	109
5.3.1	Patient cohorts	109
5.3.2	SHG Imaging	110
5.3.3	Correlation of structural metrics with collagen features	111
5.3.4	Dual IF and counterstain	111
5.3.5	IF and SHG combined	115
5.3.6	Correlation between FOXP3 ⁺ and CD8 ⁺ infiltration and collagen properties	115
5.3.7	IMC image analysis	115
5.3.8	Correlations between more immune populations	117
5.3.9	Collagen deposition and hypoxia	118
5.3.10	Comparing immune cell densities between epithelium, high density and low density collagen	119
5.4	Discussion	119
6	Conclusions and Outlook	125
6.1	Summary	125
6.2	Conclusions	126
6.3	Future Work	127
	References	131

Appendix A Probability Density **149**

Appendix B Published Work **151**

 B.1 Combining measures of immune infiltration shows additive effect on survival
 prediction in high-grade serous ovarian carcinoma 151

List of figures

1.1	Schematic illustrating the hallmarks of cancer[1]	3
1.2	Cells in the immune system and the progenitors from which they are derived[2]	4
1.3	Morphology types in HGSOE	11
2.1	Individual CD8 and FOXP3 antibody stains.	21
2.2	Multiplex CD8/FOXP3/CK staining	22
2.3	Nearest neighbour distances diagram	31
3.1	Visual Abstract for Chapter 3	36
3.2	Segmentation in Definiens	38
3.3	REMARK diagram for the SEARCH cohort.	40
3.4	Heatmap of total core area across TMA	42
3.5	Histogram of the distribution of total tissue area for each core in the SEARCH cohort.	42
3.6	Heatmap of density of infiltrates in each core across TMA	43
3.7	Scatter plot and linear fit for the epithelial and stromal areas of each core	44
3.8	Distribution of Mutant Allele Fractions	45
3.9	<i>TP53</i> mutant allele fraction against epithelial area of a core	45
3.10	<i>TP53</i> allele fraction and sequencing depth for technical replicates.	46
3.11	Distribution of immune densities.	47
3.12	Correlation between infiltrates	48
3.13	Immune exclusion ratio of CD8 and CD45RO T-cells	49
3.14	Martingale residuals of each infiltrate	51
3.15	Residuals of model fit	52
3.16	KM Survival curves for individual infiltrates	55
3.17	Patients plotted on axes giving their values of Principal Component 1 and 2.	58
3.18	Kaplan Meier Curves for Principal Component 1	59
3.19	Examples of tissue sections with largest values of principal components.	60

3.20	AIC for most significant models.	61
3.21	Boxplots of infiltrate distribution against genomic phenotype.	63
3.22	Histogram of morphological subtypes	65
3.23	Area of epithelium and stroma across morphological subtypes	65
3.24	Distribution of CD8 ⁺ , CD45RO ⁺ and CD68 ⁺ infiltrate in stroma and epithelium for different morphological subtypes of ovarian cancer	66
4.1	Visual abstract for Chapter 4.	72
4.2	Analysis in project outline	73
4.3	UpSetR plot of OV04 patients	75
4.4	An example of the stroma and tumour classifier built using QuPath software	76
4.5	Boxplots of the distribution of epithelial percentages in the OV04 cohort in each of the sample types.	78
4.6	Epithelial percentage as compared between samples.	79
4.7	L-function plotted against distance for a single core epithelial cell point pattern.	80
4.8	Correlation between median Epithelial NNdist	80
4.9	Images with very high and very low border cell percentages	81
4.10	Histogram of border cell distribution.	82
4.11	Correlation between percentage of epithelial cells classified as border cells for matched Ovarian samples	82
4.12	Serial sections in H&E and CK staining with significantly different border cell percentages.	83
4.13	Outlying cores	84
4.14	Correlation between Shannon entropy for matched patient samples	85
4.15	Examples of point pattern of cells and the division of the region for calculating Batty entropy	86
4.16	Correlation between Batty entropy for matched patient samples	86
4.17	k-means neighbour clustering of cells	87
4.18	k nearest neighbour elbow plot to find optimal <i>eps</i>	88
4.19	DBSCAN example	89
4.20	Boxplots for distribution of number of stromal, epithelial and all cell types	89
4.21	Boxplots for distribution of number of clusters across different sample and staining types.	90
4.22	Correlation between number of epithelial clusters for matched patient samples	90
4.23	Correlations between structural metrics	92
4.24	Cores with high and low structural PC values	95

5.1	Visual abstract for Chapter 4	105
5.2	REMARK diagram for the BriTROC cohort.	106
5.3	Emission spectrum for Cy5 with excitation at 633nm	108
5.4	Emission spectrum for Cy3 with excitation at 542nm	108
5.5	Emission spectrum for DAPI with excitation at 405nm	108
5.6	Examples of SHG collagen signal	110
5.7	Correlation between nearest neighbour metrics and collagen.	112
5.8	Correlation between entropy metrics and collagen.	113
5.9	Correlation between border cell metrics and collagen.	114
5.10	Images from IF protocols.	115
5.11	Images measuring IF and SHG signal.	116
5.12	Example of BRITROC and ICON7 classifiers	117
5.13	Correlation between markers of T-cell subsets CD3, CD8, PD-1	118
5.14	Correlation between markers of macrophage subsets CD68, CD163 and PD-1	119
5.15	Correlations between cells positive for markers in IMC data.	120
5.16	Heatmap of correlations between number of cells positive for each marker	121
5.17	Distribution of densities of CD68+ and CD163+ cells in collagen	121

List of tables

2.1	Bulk Reagents (Roche)	23
2.2	Chromogenic IHC antibodies used	23
2.3	Negative control IHC antibodies used	23
2.4	Antibodies used for Dual IF CD8/CD68 staining with DAPI counterstain . .	25
2.5	IMC antibodies	27
3.1	Patient characteristics for the HGSOC subset of the SEARCH cohort. . . .	41
3.2	Patient characteristics for the subset of HGSOC patients in the SEARCH cohort	41
3.3	P-values for the association of each functional form of the variable with survival	50
3.4	Individual immune infiltrates Cox regression	53
3.5	Epithelial core immune infiltrate and survival	54
3.6	Cox regression HR and p-values for combined and reduced survival models.	56
3.7	Cox regression for ratio of Tumour to Stroma infiltrate	56
3.8	Composition of the principal components across all infiltrates	58
3.9	Cox proportional hazard regression for principal components	59
3.10	Significance of Kruskal Wallis tests of infiltrate against genomic subtype . .	64
4.1	OV04 study patient characteristics.	74
4.2	Confusion matrices for QuPath H&E and CK based classifiers.	76
4.3	Linear regression model coefficients	84
4.4	Loadings for the principal components of structural metrics selected from analysis.	93
4.5	Correlations between spatial metrics and CD8 ⁺ Epithelial Density.	96
4.6	Pearson Correlation between FOXP3 ⁺ epithelial infiltrate density and spatial metrics.	97
5.1	Definitions of collagen features	111

5.2 Correlation between CD8+ epithelial density and collagen properties. . . . 116

Nomenclature

Acronyms / Abbreviations

AIC Akaike Information Criterion

CK Cytokeratin

CN Copy Number

ECM Extra-cellular Matrix

FFPE Formalin Fixed Paraffin Embedded

GLCM Grey Level Co-occurrence Matrix

H&E Haematoxylin and Eosin

HGSOC High Grade Serous Ovarian Cancer

HRD Homologous Recombination Deficiency

IF Immunofluorescence

IHC Immuno histo chemistry

IMC Imaging Mass Cytometry

KM Kaplan Meier

LGSC Low Grade Serous Ovarian Cancer

MAF Mutant Allele Fraction

NN Nearest Neighbour

OC Ovarian Cancer

OS Overall Survival

PFS Progression Free Survival

SHG Second Harmonic Generation

TMA Tissue Micro Array

Chapter 1

Introduction

1.1 High Grade Serous Ovarian Cancer

Cancer is defined by the uncontrolled proliferation of cells and formation of a mass of malignant cells known as a neoplasm or tumour. The malignant cells may also acquire the ability to invade surrounding tissues and also to metastasise to other sites within the body. It is this metastasis rather than the primary tumour which usually results in death[3–5]. High grade serous ovarian cancer (HGSOC) specifically is a cancer of the ovaries that originates in the fallopian tube[6] and is characterized by ubiquitous *TP53* mutations and genomic instability[7].

1.1.1 Types of ovarian cancer

HGSOC is a cancer of epithelial origin, but just one of multiple histologies within epithelial ovarian cancer (EOC). HGSOC is the most common (>70% of cases)[8–10] with clear cell, mucinous and low grade serous cancers present in the population at reported frequencies of approximately 3-5%, 10-15% and <10% respectively[11]. HGSOC is not only the most common subtype but has the worst survival[12, 13]. Standard treatment for HGSOC varies by stage at diagnosis and regional practices but typically consists of either debulking surgery followed by paclitaxel/carboplatin combination chemotherapy or neoadjuvant chemotherapy with subsequent interval debulking surgery (IDS). Despite these treatments, overall survival has not improved in over 20 years. The reasoning for this is twofold; firstly the majority of patients are diagnosed with Stage 3 or later neoplasms and then secondly, despite initial successful treatment, the majority of patients develop platinum resistance and subsequent metastases which prove fatal[14].

1.2 The hallmarks of cancer in HGSOC

The hallmarks of cancer are a concept that was summarised by Hanahan and Weinberg in 2000[15] and revisited in 2011[1]. The idea is that cancer is distinguished by a set of processes which define and aid this malignant and uncontrolled proliferation and survival of cells. A review of the hallmarks in 2011 highlighted that the interaction of the cancer cell with its environment is also an intrinsic part of cancer progression and as such framed cancer as a disease of both the malignant epithelium and its microenvironment[1]. Within the hallmarks of cancer are 4 hallmarks that are of interest in ovarian cancer and specifically this project;

- **Genome instability & mutation**
- **Avoiding immune destruction**
- **Invasion & metastasis**
- **Tumour-promoting inflammation**

These are summarized within the hallmarks as a whole in Figure 1.1[1]. The most characteristic hallmark of HGSOC is genomic instability, HGSOC tumours are accepted to have extremely unstable and heavily rearranged genomes[16]. The other three of the hallmarks mentioned are ones which are fundamentally linked to the microenvironment and the factors that are of interest in this thesis. The avoidance of immune destruction is of relevance as both the presence or absence of immune cells and potential factors affecting this are to be investigated. Spatially dependent immune infiltration or the lack of it can both be considered immune evasion by the cancer cells. Invasion and metastasis are strongly linked to immune cell subsets and structural features in the tumour, both of which are to be investigated. Tumour promoting inflammation will be investigated through evaluating the potential negative prognostic value of several immune infiltrates. Understanding the properties of these hallmarks in HGSOC will allow for a better understanding of HGSOC progression and open avenues for potential new treatments.

1.3 The genome in HGSOC

As mentioned, *TP53* mutation is ubiquitous in HGSOC. *TP53* mutations alter the expression of the TP53 protein, reducing a cells ability to control proliferation[17]. It should be noted that although *TP53* mutations are considered a driver of HGSOC, *TP53* plays a key role in all cell proliferation. This means that mutations of this gene are extremely common in

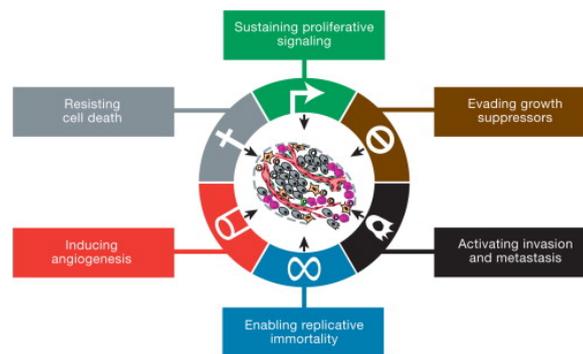


Fig. 1.1 Schematic illustrating the hallmarks of cancer[1]

multiple cancer types [17, 18] and also frequently found in normal fallopian tube cells[19, 20]. This suggests that *TP53* mutations are necessary but not sufficient for the development of HGSOC. Unlike some other cancer types there are not many common mutations that occur alongside *TP53* mutations in HGSOC, making the direct cause of many HGSOC cases hard to determine and therapy targets more difficult to locate. *BRCA1/BRCA2* mutations are present in approximately 10-20% of patients. The *BRCA1* and *BRCA2* genes encode for breast cancer type 1 and type 2 susceptibility proteins, proteins which take part in DNA repair. Mutations in these genes alter the ability of a cell to produce these BRCA1 and BRCA2 proteins and to repair its genome or to undergo cell death if the genome is damaged. A common genomic event in HGSOC is Phosphatase and tensin homolog deleted on chromosome 10q23.3 (*PTEN*) in the tumour which is prevalent in over 30% of patients and has been identified as a prognostic driver event[21]. The predominant genomic landscape in HGSOC however is not a mutation but large numbers of copy number alterations across the genome as a whole[22, 16].

1.4 The microenvironment in HGSOC

The microenvironment of a tumour comprises the interaction of the malignant cells with the cells surrounding and infiltrating the tumour and the physical and chemical properties of the region. The tumour microenvironment (TME) has been a topic of great interest in recent work in HGSOC, with rapidly increasing numbers of studies of its importance and the potential of the environment of a tumour to facilitate metastasis or drive tumour progression[23–28]. Tumour infiltrating immune cells and the extracellular matrix structure are both examples of micro-environmental features which have been shown to be prognostic in ovarian cancer.

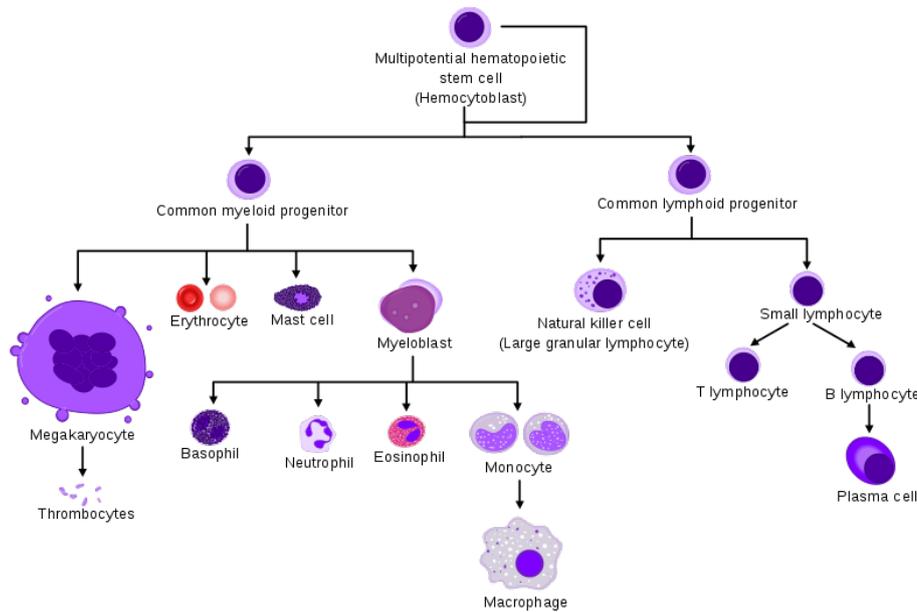


Fig. 1.2 Cells in the immune system and the progenitors from which they are derived[2]

1.4.1 Immune micro-environment

HGSOC tumours, although heterogeneous, are on average more inflamed than other epithelial ovarian tumours[29, 30] and within the HGSOC subtype many links have been found between the quantity of this immune infiltration and survival[31, 32]. With both a wealth of evidence to support the prognostic impact of the immune micro-environment and the advent of new technologies making it easier to image multiple cell types at once, the key component and the most studied part of the tumour microenvironment, immune cells, is still a key feature of research. The success of targeted immunotherapies and the interest in personalised medicine and patient stratification make this a popular research area. Of course the umbrella term of the immune system covers a large variety of cells with a large number of functions, comprising the innate and adaptive immune systems. Figure 1.2 shows the lineage of immune cells and their progenitors. The most frequently studied cells of the groups in this diagram are T-cells (T-lymphocytes) and macrophages.

T-cells

The reason for the focus on T-cells in cancer is predominantly because of evidence of their direct impact on tumour progression. Key historical events such as the discovery of T-cells in 1967[33] and the development of the first single surface antigen based vaccine[34] increased scientific interest in utilising the immune system to cure disease. In the following years,

the ability to create engineered mouse models and evidence showing that CD8+ T-cells have a positive impact on survival in many cancers made T-cells a focal point of immune oncology research [27, 35]. The first positive association between T-cells and Ovarian cancer prognosis was published in 1991[36]; since then many studies have attempted to categorise lymphocytes within Ovarian cancer. A large meta-analysis of the survival impact of various T-cell subsets in Ovarian cancer demonstrated that CD8+ cytotoxic T-cells were a significant prognostic infiltrate over 10 studies[37].

FOXP3+ regulatory T cells (Tregs) are a subclass of CD4+ T cells with immunosuppressive characteristics but studies have demonstrated an association of Tregs with both improved [31] and worse survival[32]. Such variety in reporting could be why Hwang *et al.* saw no significant impact of FOXP3+ infiltrate upon survival in the meta-analysis[37]

Most previous attempts to study tumour infiltrating lymphocytes (TILs) have used simple and manual methods to either count lymphocytes or evaluate the quantity as "low" or "high". Such methods don't account for tissue area variation between samples and unlike computational segmentation of cells, results from manual counting methods cannot be carried forward to analyses of spatial localisation. In the most modern research, quantitative measures have been related to outcome, an example of which is the observation of loglinear dose-response with survival of CD8+ cells. This study was manually scored but demonstrates a quantitative survival benefit of CD8+ cells on a very large number of patients [27]. Automated measures have started to generate quantitative data sets using modern cell-segmentation software.

Results are not limited to CD8+ cells within HGSOC. The presence of other T-cell subsets such as CD45RO+ and CD3+ cells have been shown to improve survival in HGSOC. Despite the evidence for multiple prognostic immune populations in many cancers, immune population analysis has not often taken into account the presence of multiple immune cell types. Lack of this more nuanced analysis could explain conflicting findings regarding the prognostic significance of lymphocytes in some cancers. Studies by Barua *et al.* amongst others, have demonstrated improved modelling of patient survival when the average distance between two immune cell types in a sample is taken into account[38–41]

Macrophages and their functionality

Macrophages are another very commonly studied subgroup of immune cells. Markers such as CD68+ and CD163+ are used to identify different macrophage populations in tissue. The prognostic impact and role of macrophages in the pro- and anti-tumour response has been controversial. Attempts to define their affect resulted in research focusing on presenting macrophages as one of two options, M1-anti tumour and M2-pro tumour macrophages.

More recent research illustrates macrophages to be a highly diverse population with a much wider range of phenotypes[42] and with context specific function. For example in HGSOC, Yafei *et al.* demonstrated a negative association between stromal macrophages and survival [43] whereas Tanaka *et al.* demonstrated that upregulation of Bikulin by macrophages was associated with improved survival[44]. The prognostic impact is variable amongst other cancers too; in gastric cancer, Huang *et al.* demonstrate that CD68+CD163+ macrophages are associated with improved survival in univariable analysis[45].

Immune checkpoints and Programmed Death Ligands (PD-1/PD-L1)

As mentioned in the discussion of the hallmarks of cancer, immune escape and evasion are key factors in the survival and propagation of cancer cells. This is as even in the presence of immune cells, tumour cells can still evade immune destruction. One method of evasion is the deactivation of T-cells, utilised by cells in the normal immune response to avoid overreaction and damage by the immune system[46]. PD-1 and PD-L1 are part of one such immune checkpoint system. PD-1 is the receptor on the immune cell and the corresponding ligand is PD-L1. When PD-1 and PD-L1 bind, they modulate cytotoxic T-cell activity and produce signals that promote T-regulatory cell activity. This mechanism to reduce the immune response has been shown to be co-opted by tumour cells and PD-1/PD-L1 are upregulated by tumour cells to evade immune destruction[47]. As such, we can identify tumour cells that are successfully evading the immune system by imaging or quantifying a cell's PD-1 or PD-L1 expression. Multi-marker imaging methods can therefore allow not only T-cells, but also the markers they express, to be identified. Such studies can demonstrate the functionality and activation of T-cells. Measuring PD-1 and PD-L1 expression can therefore be informative as to the effectiveness of the immune response to the cancer, for example expression of PD-1 correlated with exhaustion markers in carcinoma of the liver[48, 49].

1.4.2 Extracellular matrix and stromal micro-environment

Another factor effecting both immune and cancer cells is the surrounding extracellular matrix (ECM). The ECM structure and its mechanical stress are structural properties of the tumour micro-environment. The ECM is deregulated and disorganised in cancer; this deregulates stroma and facilitates angiogenesis[50–52]. Changes in biomechanical properties of the ECM have been shown to affect cell migration[53–55]. A key protein family that both regulates and is produced by the ECM is the lysyl oxidase (LOX) family of proteins, this family is upregulated in hypoxic environments showing that both structural and biochemical gradients can be linked. These LOX proteins cross-link collagen fibres and change the biomechanical

properties of the ECM. LOX expression has been directly linked to metastasis [56–58] but over-expression of LOX alone is not enough to induce tumorigenesis[56].

The ECM has been shown to become dysregulated during HGSOC tumorigenesis. This dysregulation is why the remodelling of collagen, as imaged by second harmonic generation imaging, can distinguish normal from ovarian cancer samples[59]. On a gene expression level, changes in nuclear LOX expression have been linked to poor prognosis and over expression of collagen remodelling genes in HGSOC has been associated with reduced survival[60, 61]. The over expression of other ECM proteins, such as fibronectin, has also been associated with poor survival and metastasis in HGSOC[62]. In other work on the importance of the stroma in cancer progression, carcinoma associated fibroblasts were shown to be capable of transforming non-tumorigenic epithelial cells into tumorigenic ones and inducing tumour growth[63]. Fibroblasts have also been linked to metastasis, a key factor in patient survival, through reorganisation of ECM structure[64].

Collagen

Collagen is the main component of the extracellular matrix and collagen deposition is a key part of desmoplasia[65]. In experiments that assessed the expression of genes in both cisplatin resistant and sensitive ovarian cancer cells, many ECM genes were elevated in cisplatin-resistant cells. COL6A3 was one of the most highly upregulated genes, and cultivation of cisplatin-sensitive cells in the presence of collagen VI protein promoted resistance in vitro[52]. Work by Januchowski *et al.* in 2016 found similar results[66]. There is much evidence to suggest that collagen remodelling is utilized by cells to aid survival in the presence of chemotoxic drugs[67, 68]. Collagen remodelling has also been linked to increased motility and promotion of metastasis[69, 70, 61, 71] and the normalization of the tumor stroma has been linked to improved sensitivity to chemotherapy in ovarian cancer[72].

Collagen also provides structure for cell migration and has been linked to immune cell motility as well as metastasis. CD8+ cells were shown to predominantly move along collagen fibres and to move faster in the epithelial regions than in the collagen filled stroma[73]. They were also shown to move faster along aligned fibres than unaligned ones[74].

1.5 Immune recruitment and control

When immune cell populations are imaged in tumours, what is measured is the infiltration of immune cells into specific regions of tissue. It is this recruitment of cells which is assigned prognostic value. It is therefore a fundamental question of interest as to what drives or changes this infiltration which can vary dramatically between patients. If immune cell

recruitment and modulation is understood, it can perhaps be controlled in patients to improve survival. Immune checkpoint inhibitor therapies (eg. anti-PDL1) aim to do this by stopping the existing immune inhibition by the tumour.

Activation of specific DNA repair pathways may be responsible for some of the variation in immune responses in HGSOC patients. Heavily rearranged genomes are prevalent in HGSOC and *BRCA1* mutations are associated with homologous recombination deficiency[75]. DNA damage has been shown to trigger the cytosolic DNA, active cyclic GMP-AMP synthase-stimulator of interferon genes (cGAS-STING) pathway through the formation of micronuclei[76] and upregulation of the STING pathway has been shown to be associated with immune signalling in breast cancer. The rearrangement and reduced ability to repair DNA damage in *BRCA1* mutated HGSOC patients may increase STING pathway activation. This increase in STING activated signalling may be the cause of higher immune infiltration and hence improved survival in *BRCA1* mutated patients. In work supporting these ideas a mouse model of *BRCA1*-deficient HGSOC tumours, PARP inhibition (immune checkpoint inhibition) prompted a STING-dependent immune response[77, 78].

Another potential factor in the control of immune infiltration is the physical structure of the tissue. As mentioned, remodelling of the extra cellular matrix (ECM) by fibroblasts has been shown to increase thickness and to alter the orientation of collagen fibres which can in turn affect immune cell motility[79]. Importantly, observed differences in tissue architecture in *BRCA1* mutated tumours could be the driver of altered immune infiltration[80]. *BRCA1* mutated tumours are associated with higher immune infiltration as well as different structure of the tissue[81, 80, 82].

Therefore we arrive at two hypotheses about how infiltration into the epithelial compartment may be regulated;

- **The physical barrier hypothesis:** A modified epithelial structure and collagen environment could act as a physical barrier, reducing immune infiltration into the epithelium [83, 84].
- **The chemo-control hypothesis:** Cell signalling, of various origins, at the margins between tumour and stroma may attract or repel immune cells by control through chemotaxis.[85]

1.6 Subtyping HGSOC

The set of all HGSOC tumours is diverse but still poorly categorised. Developing a robust classification of patients is still a key focus of HGSOC research; the aim of which is to be

able to group patients for clinical trials more effectively and to predict differences in and mechanisms of response when trialling new treatments.

1.6.1 Genomic subtypes

As mentioned *TP53* mutations are believed to be ubiquitous in high grade serous ovarian cancer[86]. Other specific driver mutations are uncommon but over 30% of HGSOC present with loss of PTEN when the stromal DNA contribution is excluded, this loss was found to be prognostic[87]. *BRCA1/2* mutations are present in 10-20% of HGSOC patients and are associated with altered morphology, survival and immune infiltration[80, 81]. Otherwise, the HGSOC genome is characterized by being extremely complex and highly rearranged[16]. Without specific driver mutations. The direct impact of such heterogeneous genomes is hard to elucidate but signatures addressing copy number rearrangement suggest some patterns of rearrangement are associated with increased infiltration[88].

Copy number (CN) signatures have proven to be the best classification of genomic instability in HGSOC so far, identifying different forms of chromosomal rearrangement that have contributed to a patient genome[16]. Wang *et al.* also discovered CN aberrations and the presence of foldback inversions stratified patients in HGSOC[89]. These recent works show that genomic rearrangements on a larger level may stratify HGSOC patients and may affect immune infiltration. The exact mechanism by which immune modulation occurs is not yet known but it has been suggested that the activation of pathways such as the STING pathway or the generation of micronuclei during genomic rearrangement may be driving factors.

1.6.2 Gene expression sub-types

Many attempts have been made to categorise HGSOC patients into subtypes by gene expression. The molecular subtypes as defined by Tothill *et al.* began a series of validations and improvements across the literature. Despite the initial analysis being carried out with k-means clustering and as such requiring pre-defined numbers of clusters, the consensus is now that of the 6 initially identified subtypes there exist four gene expression molecular subtypes in HGSOC and these have been validated across multiple cohorts[90, 91]. These subtypes are labelled as mesenchymal, immunoreactive, differentiated and proliferative. These subtypes are associated with differences in survival, with the immunoreactive gene expression signature conferring a survival advantage. These molecular subtypes have also been shown to delineate the affects of treatment response, in the ICON7 trial, the proliferative

subtype was seen to have a significant improvement in overall survival when treated with Bevacizumab[92].

1.6.3 Morphological sub-types

Another microenvironmental aspect of the tumour is the architecture and growth patterns of the epithelium. This tissue structure is known as architectural grade and a robust classification has been elusive for decades[93, 94]. There are many different versions of these tissue architectures even within HGSOE. Classifications all usually contain versions of glandular, papillary and solid architectures, with histopathological grading scoring the whether the tumour is predominantly any of these classifications[13]. Architectural grading has repeatedly been associated with survival in patients[94, 93, 95] but since the important development of robust low grade and high grade classifications, nuances in architectural features within these subtypes have not received as much attention, possibly due to difficulties with reproducing classifications[96].

Some recent work focusing on the architecture of tissues has reinforced the relationship between prognosis and structure. Metastatic tissues from HGSOE patients were classified as pushing pattern (either predominantly or exclusively), an exclusively micropapillary infiltrative pattern, or an infiltrative pattern composed of papillae, micropapillae, glands and nests (mixed infiltrative pattern) and these classifications were found to correlate with prognosis[95].

In work by Soslow *et al.* it was observed that Solid, pseudoEndometrioid, and Transitional cell carcinoma-like morphology (SET) features overlapped significantly in samples and these were grouped into a single classification with tumours labelled as displaying SET features or not[80]. *BRCA* mutations were associated with SET features and papillary patterns associated with non-*BRCA* mutants.

Examples of solid, papillary and glandular morphologies are shown in Figure 1.3. Many tumours can display a mixture of these growth patterns[80].

Alternate morphologies across HGSOE tumours have also been described by Murakami [97]. Murakami *et al.* defined a mesenchymal type with desmoplasia, an immunoreactive type, a papillary glandular type and solid proliferative type. These subtypes are loosely related to the Tothill signatures and are not widely accepted architectural grades.

There is no existing automated classification of growth structures but Murakami *et al.* found interobserver agreement when 6 pathologists classified slides and found that the structure in metastases from the same patients were of the same class as the primary tumour. These morphological subtypes were also associated with the Tothill molecular gene expression subtypes of mesenchymal, immunoreactive, differentiated and proliferative but the

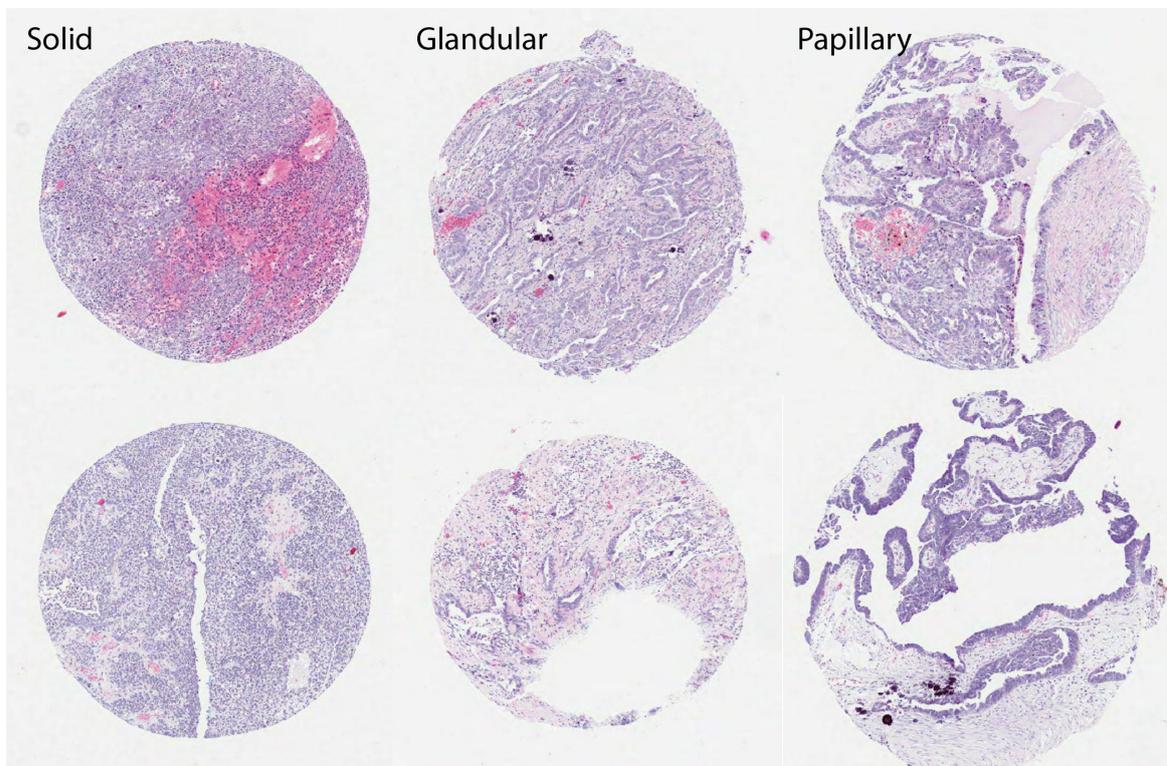


Fig. 1.3 Heterogeneous morphology of HGSOC tissues. Examples show solid, glandular and papillary architectures

subtypes were not automatically derived or scored. It is worth noting that several observers misclassified the immuno-reactive subtype. This highlights that an immuno-reactive subtype defined this way is not a reliable architectural subtype. This is evidence that some examples of all of the architectural patterns can be seen to have high numbers of tumour infiltrating lymphocytes (TILs).

Methods of clustering patients by gene expression profiles have similar issues when structure and immune infiltration are conflated. The prognostic impact of some gene expression subtypes can potentially be explained entirely from being driven by the T-cell infiltrate contribution to the signature. Schwede *et al.* suggest that once adjusted for stromal content or T-cell infiltration, the gene expression signatures are no longer prognostic[98]. The importance and interplay of structure, stromal content and immune infiltrate is further reason to investigate the link between them and the possibilities of structural barriers to T-cell infiltration.

1.6.4 Immune subtyping

Categorising patients into different immune subgroups is of particular interest to those designing immunotherapy trials. As such, analyses of tumour infiltrations, and in particular TILs, have been used to classify patients into subcategories such as immune desert, immune inflamed and immune excluded[99]. These classes have been allocated using tools such as a basic density analysis or visual comparison between regions. Chen *et al.* estimated that approximately 50% of their cohort exhibited immune exclusion, approximately 25% inflamed and approximately 25% immune desert phenotypes but do not detail the classification method[99]. A classification in bladder cancer defined immune exclusion as $3\times$ as much infiltrate in the stromal as compared with the epithelial compartment[100]. There is, however, a wealth of evidence that suggests that overall immune phenotyping for even a single patient may be more complex. Zhang *et al.* demonstrated clonal dynamics and spatial immune variation within a patient at diagnosis that associated with clonal subsets and that reflected the clonal direction of the patient[101]. Kather *et al.* also do not observe clear subtypes when using unsupervised clustering[102]. In HGSOC, Jimenez-Sanchez *et al.* showed wildly diverse microenvironments[103] across metastases in a single HGSOC patient and varying immune infiltrations were associated with region specific regulation of the Wnt pathway. Robust immune subtyping remains elusive in the presence of such complex heterogeneity.

1.7 Methodologies for the imaging of the microenvironment

In order to assess the microenvironment of a tumour we must first be able to see it. HGSOC tissue must be obtained and then the cells and structures of interest must be marked and then imaged. In HGSOC, samples of tissue are collected from primary or interval debulking surgery(IDS) and stored in Fixed Formalin Paraffin Embedded(FFPE) blocks which preserve the tissue for many years. The majority of samples are stored this way as they are easier and cheaper to store than fresh frozen samples but the DNA is partially degraded[104].

1.7.1 Tissue microarrays (TMAs)

When analysing immune populations across whole cohorts, instead of staining, imaging and analysing whole slide tissue sections for every patient, tissue microarrays are often constructed from the FFPE blocks. Tissue microarrays consist of cores, punched from the tumour block, of approximately 1mm diameter which are then arranged in a grid containing between tens and hundreds of these samples. Sections of these new TMA blocks are then cut, dewaxed, mounted on slides and stained. It has been shown in multiple studies that samples in TMA sections are representative of the wider tumour [105–107]. This result is reliant upon careful pathological sampling and there are clear limitations to the representation of the whole tumour with one region. This is of particular note in tumours with fluid sacs or observed distinct tissue types. Recommendations for ensuring representative sampling include sampling from tumour peripheries where fixatives are most effective[108].

1.7.2 Haematoxylin & Eosin (H&E)

When staining the TMA or slides of patient tissue, the most common staining type used is Haematoxylin and Eosin (H&E). Haematoxylin stains the nuclei blue and Eosin stains the cytoplasm of the cell pink. H&E staining therefore is good at visualising the layout of the cell and thereby the tissue. For this reason H&E stains are routinely used for diagnosis in cancer[109] and to guide TMA construction[110]. Due to the popularity of this staining technique, there are many repositories of H&E slides and digitally scanned images of H&E sections and TMAs across research departments worldwide, as evidenced by many studies utilising these slides for further analysis[111].

1.7.3 Immunohistochemistry (IHC)

Immunohistochemistry is a method of antigen visualisation that utilises the fact that antibodies bind to specific antigens. The method uses conjugated antibodies to identify specific

antigens which may be expressed on cells and as such can be used to identify specific cell subsets. In the most common and cheapest method, chromogenic IHC, antigen specific antibodies are attached to chromogenic dyes for visualisation, these dyes are visible under natural light and the resultant stained sample is imaged for digital pathology using digital light microscopy. Chromogenic IHC is frequently used for the evaluation of p53 expression and the identification of wildtype, gain of function (GOF) or loss of function (LOF) mutations [112].

1.7.4 Immunofluorescence (IF)

Fluorescent IHC is the process by which fluorophores are attached to the antibodies in IHC instead of chromogenic dyes. The number of markers on a section/single image is then limited by the absorption and emission spectra of the fluorophores and detector limits rather than the more limited visual spectral range. For such reason, measuring multiple markers is easier than that of chromogenic IHC. IF staining is visualised using laser illumination under a confocal microscope. One downside of IF is that the fluorescent nuclear counterstains, such as DAPI, which are used to visualise underlying tissue structure do not highlight the tissue morphology in nearly as much detail as H&E or the haematoxylin counterstain in chromogenic IHC methods. Furthermore, even with careful refrigerated storage, fluorophores have a limited lifetime and IF signal will eventually fade. In contrast to IHC and H&E stains which retain high levels of detail over years and even decades, the quality of imaging of IF stained slides is greatly reduced within weeks of staining.[113]

1.7.5 Increasing the number of markers

The imaging of more than 3 or so markers using IHC or IF becomes very complex. Very rapidly overlapping spectra in the range detectable by standard confocal microscopy become hard to distinguish and leakage occurs between channels. Imaging mass cytometry (IMC) is a novel solution to this problem, using immuno-metal conjugation of antibodies to metals. The tissue is stained with these antibodies but then the tissue is ablated, pixel by pixel, and these ablated antibodies are detected and identified by mass spectrometry. This method allows for over 30 distinct markers and subsequent detection channels on a single section[114]. Other methods for imaging multiple markers include consecutive fluorescent staining which requires labour intensive washing off of markers[115]. IMC is, despite its advantages, still a more expensive and time consuming method than IHC based methods. Metal-conjugated antibodies cost thousands of pounds and the ablation steps take over an hour for each 1mm² of tissue core/tissue area[116].

1.7.6 Second-harmonic generation (SHG)

Second harmonic generation (SHG) is a nonlinear optical process in which two photons with the same frequency interact with a nonlinear material. The linearity of a material is defined by the relationship between the polarization and electric field in the material. In a linear material these quantities have a linear relationship and the polarization is directly proportional to the electric field. A non-linear material is one in which this relationship is non-linear. For second order effects to occur such as second harmonic generation, the material must have non-centrosymmetric structure. In non-linear materials, such as collagen, two photons of the same wavelength can combine to make a single new photon with twice the energy of the initial photons and therefore half the wavelength. Collagen fibres can be imaged in this way as the bundles of fibres constitute a non centrosymmetric, non-linear material. SHG and Third harmonic generation imaging of non-linear materials was first carried out in 1974[117] but was first demonstrated in tissue in 2003[79]. An advantage of this method of imaging collagen over the use of staining such as picosirius red is that SHG imaging is marker free and so can be carried out without antigen retrieval[118]. SHG imaging can also often be carried out on tissues which have already had other staining applied[119]. This method of collagen imaging in tissue has seen a large surge in publications over the last 10 years including in HGSOC[120–124].

1.8 Methodologies for the image analysis of the microenvironment

1.8.1 Nuclei segmentation

The segmentation of cells in an image is a difficult computational problem; the identification of cell membranes which do not stain strongly in H&E images is particularly difficult and different software tools have shown wildly different results[125]. Several papers have shown that it is possible to accurately and automatically segment nuclei from H&E images[126, 127] and most basic segmentation methods therefore focus on the segmentation of strongly stained nuclei in H&E imaging and produce robust results.

1.8.2 Tissue Classification

In order to assess the spatial nature of immune infiltration in epithelium and stroma when analysing immune infiltrate in digital pathology, the epithelium and stroma must be segmented.

An experimental method that has improved the ability to do such tissue subtyping is the use of markers such as pan-cytokeratin; a cocktail of two antibodies that marks multiple families of cytokeratins. The use of pan-cytokeratin or alternatively specific cytokeratin markers in multi-marker imaging methods has allowed both epithelial cells and immune cells to be simultaneously identified.

In downstream analyses, approaches such as stain deconvolution have improved accuracy in the field[128]. As the abundance of publically available H&E images and their use in digital pathology is so high, tissue classification is still the focus of research today[129, 130]. Classification and segmentation methods based on deep learning continue to improve classification[131]. A potential drawback with these methods is that commercial software which may perform such classification is expensive. Free software such as QuPath[132] is improving access to effective tools for tissue classification.

1.8.3 Spatial Immune Analysis

Digital pathology, more accessible image analysis methods, more accessible software and the availability of multiplexing has allowed for the easier extraction of the spatial location of many cell types and consequently has increased the number of studies including spatial analysis of immune cells in tissue.

Quantitative metrics and spatial statistics in the TME are most easily derived from images and frequently prognostic. Carstens *et al.* carried out quantitative image analysis of densities and distributions of T cells in pancreatic ductal adenocarcinoma. Using Ripley's L function they found heterogeneous infiltrating T cell populations whose quantity and proximity to cancer cells was prognostic[133]. Ali *et al.* found that median density of lymphocytes was prognostic in breast cancer. This metric was calculated using k-nearest neighbours to find the average distance of a lymphocyte from the nearest 50 lymphocytes in the tumour section[134]. Barua *et al.* measured the distance of T-regs from epithelium in non small cell lung cancer and found, using the area under the curve (AUC) of the G-cross function, that the smaller this distance, the worse the survival. They also found that smaller distances between FOXP3+ and CD8+ cells were associated with improved survival[41]. A multitude of work, including measures such as spatial grouping of B and T cells and densities of T follicular helper cells, has shown that spatial measures are prognostic and a large quantity of literature emphasises the relevance of inter-cell TME interactions that can be measured only through image analysis[135].

The composition and development of the tumour microenvironment has also been examined by many authors from the perspective of evolving and competing populations of cells. Ecology assesses population growth and competition and methods from this field have been

applied to the cells in tumour samples. Yuan *et al.* have quantified the TME in this way in multiple cancers, often looking at the tumour tissue from an ecological perspective. They also found that the number of co-localized immune and cancer hot spots, regions where high values of a variable are clustered, was prognostic in oestrogen receptor negative breast cancer patients. They also found that this spatial metric was not correlated with gene expression in the samples[136]. Their work may suggest that tumour micro-environmental composition is an independent driving factor in addition to cell-intrinsic effects[137].

In HGSOE specifically, Yuan *et al.* showed that stroma and tumour composition and stratification of patients into those with high or low lymphocyte density were prognostic[138]. Yuan *et al.* also investigated the similarity and diversity of TME in HGSOE metastases using ‘ecologically inspired’ ideas such as niches, resources, diversity and Shannon entropy[136]. This measure of diversity was prognostic, with more diverse micro-environments leading to metastasis[139].

Massi *et al.* combined multiple aspects of cell based image analysis, analysing the spatial distribution of T-cells in triple negative breast cancer in order to test the hypothesis that immune infiltration into tumour clusters is mediated by a physical barrier [140]. They found no evidence linking the depth and quantity of infiltration of T-cells to the collagen structure. They also used a model of the alternate hypothesis of chemoattractants and chemorepellants at the border of epithelial clusters and found that this model matched their data.

1.9 Project Hypotheses, Outline and Motivations

- Hypothesis 1: Immune infiltration in both the epithelial and stromal regions of tissue sections is prognostic.
- Hypothesis 2: The combination of multiple immune infiltrations into a survival model will improve the accuracy of survival modelling.
- Hypothesis 3: Increased stroma-epithelium contact in a sample will increase epithelial immune infiltration.
- Hypothesis 4: Physical barrier hypothesis - Increased density of collagen deposition and increased disorder of fibres will decrease epithelial immune infiltration.

This thesis aimed to investigate the structural nature of immune infiltration and immune exclusion. I aimed to assess these hypotheses by developing and comparing quantitative measures of immune infiltrate, tissue morphology and collagen. This project investigated these hypotheses in order, starting with an exploration into immune quantification, comparing

metrics and regions of interest and evaluating survival, then developing automated measures of underlying tissue architecture. Thereby allowing relationships between structural metrics and immune infiltration to be evaluated. Finally the physical barrier hypothesis was investigated by analysing measures of collagen and associations of this collagen structure with both epithelial architecture and quantity of infiltrating immune cells.

Chapter 2

Methods

2.1 External Contributions

Anne Montfort (AM) - AM carried out the staining of tissue from the SEARCH cohort and used Definiens to extract cells and annotate stromal and epithelial regions of SEARCH cores.

Jaqueline McDermott (JMcD) - JMcD, a pathologist, was involved in the manual verification of Definiens tissue region and cell classifications.

Anna Piskorz (AP) - AP carried out the sWGS and analysis of sWGS data for SEARCH.

Sarwah Al Khalidi (SAK) - SAK carried out the optimisation and staining of the OV04 cohort with pancytokeratin, FOXP3 and CD8 IHC antibodies. SAK carried out the follow up sWGS and analysis for TP53 mutation on a small subset of SEARCH samples. SAK was part of the creation of the BRITROC TMA and the staining of BRITROC with IHC and IMC antibodies. SAK optimised the antibody panel for IMC.

Jodi Miller (JM) - JM of the histopathology core of the Cancer Research UK Cambridge Institute (CI), carried out the staining and assisted with optimisation of the immunofluorescence imaging of nuclei, CD8⁺ and CD68⁺ cells.

Fatime Qosaj (FQ) - FQ carried out the staining of ICON7 with IMC antibodies.

Richard Grenfell (RG) - RG carried out the running of the Hyperion system for imaging the ICON7 IMC slides.

Jennifer Alsop (JA) - JA carried out the construction and sectioning of ICON7 TMA blocks.

Sectioning and any automated staining of OV04 samples was carried out by the histopathology core services at the CI.

2.2 Patient Cohorts

Multiple patient cohorts were used throughout the thesis and are patient characteristics are detailed below.

- **SEARCH** (Studies of Epidemiology and Risk Factors in Cancer Heredity)[141]
SEARCH is an ongoing population based study covering the regions of East Anglia and the West Midlands in the UK. The study was approved by all the Local Research Ethics Committees in East Anglia. (Reference 99/5/007)
(n=332 HGSOC patients)
- **BriTROC-1 Study** (British Translational Research Ovarian Cancer Collaborative)[142]
BriTROC-1 is a sample collection study to investigate the role of Homologous Recombination Deficiency in platinum sensitivity in recurrent high grade serous ovarian cancer. Relapse samples were collected but only data from initial biopsy tissues were assessed in this PhD.
Ethics approval from Cambridge Central Research Ethics Committee (Reference 12/EE/0349).
(n= 202 HGSOC patients)
- **CTCR-OV04**[143]
Molecular Analysis of Response to Treatment in Ovarian Cancer. To investigate the mechanisms of treatment response in ovarian cancer through the study of sequential tissue, ascitic and circulating tumour cells and the establishment of an ex-vivo model. Prospective clinical study designed to identify biomarkers of heterogeneity in ovarian cancer.
Ethics approval from Cambridgeshire Research Ethics Committee (Reference 08/H0306/61)
(n=156 HGSOC patients)
- **ICON7** (International Collaboration on Ovarian Neoplasms)[144, 145]
ICON7 was randomised, two-arm, multicentre Gynaecologic Cancer InterGroup trial of adding bevacizumab to standard chemotherapy (carboplatin and paclitaxel) in patients with epithelial ovarian cancer. The samples were made available for further research.
Trial no. ISRCTN91273375
(n=313 HGSOC patients)

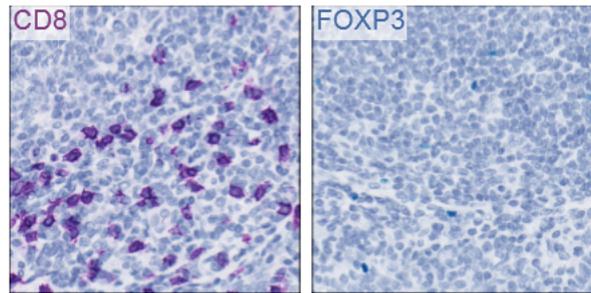


Fig. 2.1 Individual CD8 and FOXP3 antibody stains.

2.3 Chromogenic Immunohistochemistry (IHC)

2.3.1 Sample preparation

3 μ m sections of FFPE blocks were cut using a Leica microtome, transferred to a water bath pre-heated to 40°C and collected on a glass slides. The sections were dried over-night at 37°C then baked at 60°C for 1h to remove some of the paraffin.

2.3.2 Manual Staining

CD8, CD3 and FOXP3 dual staining for OV04 and BRITROC

FFPE sections were further dewaxed and rehydrated using the Leica Autostainer ST020 by incubating them twice for 10 min in xylene, followed by two 5 min incubations in 100% ethanol, one 5 min incubation in 70% Ethanol and finally washing them in water. Antigens were retrieved in a Tris (10 mmol)-EDTA (1mM) buffer (pH 9.2) preheated to 96°C in a water bath. After 1h incubation, slides were cooled in a water bath for 5 min, washed in Milli-Q water for 5 min and incubated in 3%BSA-PBS for 1h in a hydration chamber at room temperature. ImmEdge pen was used to draw a hydrophobic border around the tissue, and 200 μ l of primary antibody pre-diluted to the desired concentration in 1%BSA-PBS were added to the sections. Following an overnight incubation in a hydration chamber at 4°C, the antibody solution was drained and the slides were washed twice for 5 min in 1% Tween-TBS, succeeded by two washes for 5 min in TBS. For Horseradish Peroxidase (HRP)-conjugated primary antibodies, DAB was directly added to the slides. For secondary antibody staining, the pre-diluted secondary antibodies were added and incubated at room temperature for 30 min, followed by 2 washes in TBS. The secondary antibodies are listed in Table 2.2. The slides were incubated in DAB chromogen for 5 min, washed in Milli-Q water, and dehydrated and cover-slipped using the Leica Autostainer ST020. Examples of CD8, FOXP3 and the triple Cytokeratin, CD8 and FOXP3 stain are shown in Figures 2.1 and 2.2[88].

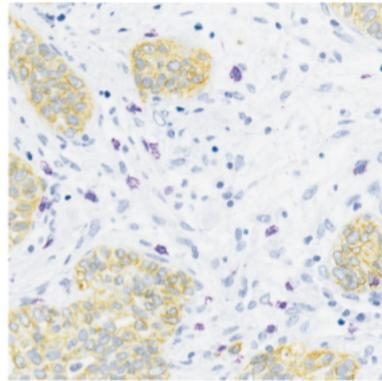


Fig. 2.2 An example of tissue with pancytokeratin(yellow), CD8(purple) and FOXP3(teal) staining.

SEARCH

Sectioning and staining was carried out as follows on the SEARCH cohort. Microarray slides composed of FFPE embedded ovarian tumour cores were dewaxed and rehydrated prior to heat induced epitope retrieval (HIER) using a pressure cooker and a citrate-based antigen unmasking solution (Vector Laboratory). Detection of CD8⁺ T cells, CD45RO⁺ memory lymphocytes and CD68⁺ macrophages was performed using the mouse anti-human CD8 (Clone C8/144B, Dako), mouse anti-human CD45RO (Clone UCLH, Dako) and mouse anti-human CD68 (Clone M0876, Dako) antibodies, using ultrasensitive Polymer-HRP IHC Detection system (Biogenex). Immunohistochemical protocols and slide hybridizations were carried out manually. Sections were counterstained with haematoxylin and mounted with DPX mounting medium (Sigma). Stained slides were scanned using the Panoramic Slash Scanner (3D Histech).

2.3.3 Automated Staining

Automated staining was carried out using the Ventana Discovery Ultra platform. All bulk reagents used were purchased from Roche and are listed in Table 2.1. The slides were rehydrated by incubating them in EZ prep solution for 32 min at 69°C, then incubated for 1h in VentanaCell Conditioning buffer 1 (CC1) at 96°C for antigen retrieval (pH 8.5). Hydrogen peroxide was then applied and incubated for 4 min to quench endogenous peroxidase activity, followed by primary antibody incubation for 1h at 37°C. Details of antibodies used for chromogenic IHC are listed in Table 2.2.

After a 16 min incubation with an HRP-conjugated secondary antibody, the chromogenic signal was developed in either 3, 3'-Diaminobenzidine (DAB) chromogen for 8 min, in

Reagent	Code
EZ prep	950-102
Ultra Cell Conditioning Solution (CC1)	950-224
Ultra Liquid Cover Slip (LCS)	650-210
Reaction Buffer	950-300
ChromoMap DAB Kit	760-159
Purple Kit	760-229
Yellow Kit	760-239
Antibody Diluent	760-108
Haematoxylin II	760-2208
Bluing reagent	760-2037

Table 2.1 Bulk Reagents (Roche)

Target Secondary	Source	Catalog No.	Conc($\mu\text{g/ml}$)
CD8	Spring Bioscience	M5394	0.5
OmniMap Anti-Rb-HRP	Roche	760-4311	prediluted
FOXP3	Spring Bioscience	M3974	0.25
OmniMap Anti-Rb-HRP	Roche	760-4311	prediluted
Pan-Keratin	Roche	760-2135	prediluted
OmniMap Anti-Ms-HRP	Roche	760-4310	prediluted

Table 2.2 Chromogenic IHC antibodies used and their secondary antibodies.

Purple chromogen for 40 min, in Yellow chromogen for 28 min or in Teal chromogen for 8 mins. For chromogenic counter-staining, the slides were incubated for 4 min in Copper, 8 min in haematoxylin and 4 min in Bluing Reagent. Slides were washed with Reaction Buffer after each incubation. Tonsil sections were used a positive control in optimisation due to high immune infiltration. As a negative control, the monoclonal anti-mouse and anti-rabbit antibodies in Table 2.3 replaced the mouse and rabbit primary antibodies, respectively. Liver, kidney and spleen samples were also included in every TMA as control tissues. Stained slides were dehydrated using the Leica Autostainer ST020, manually cover-slipped and digitally scanned by Aperio Scanscope XT.

Target	Source	Catalog No.	Conc($\mu\text{g/ml}$)
Rabbit IgG	Roche	760-1029	prediluted
Mouse IgG	Roche	760-2014	prediluted

Table 2.3 Antibodies used for negative controls.

2.4 Second Harmonic Generation

Second Harmonic Generation (SHG) images were acquired from unstained and IHC stained tumour sections to assess the collagen content. SHG microscopy was performed using a Leica SP5 laser scanning confocal microscope (Leica Microsystems Inc, Wetzlar, Germany) with a mode-locked fs-pulsed IR-laser (Coherent Chameleon Ultra II, Ti-Sapphire, 680-1080nm excitation range; Coherent Scotland Ltd., Glasgow, UK). Illumination and emission wavelengths for the acquisition of collagen SHG signal were determined from illumination and emission spectra as 920nm and 460nm respectively which are consistent with the principles of SHG. Images were acquired with an HC PL APO 40x oil immersion and 20x Dry objective. Once optimised, imaging acquisition parameters were kept constant across all imaging sessions. SHG microscopy images were automatically acquired across TMAs using tilescan and allocating a 1mm×1mm field for each core. Imaging was optimised on a single slide.

The SHG signal over the whole image was quantified using ImageJ and the OrientationJ and GLCM plugins. The energy and entropy of each image were calculated using GLCM. Once optimised, threshold parameters were kept constant for all images. For each tumour core the mean SHG signal, entropy, energy were calculated. The image was then split into 100 squares (each 100 μ m × 100 μ m) and a minimum signal threshold was set, the non background squares were analysed for orientation and coherency of the fibres.

2.5 Dual IF

All slides were run using Leica's Research Detection System 2 (DS9777) on their automated Bond RX platform. Prior to loading, the slides were dewaxed and re-hydrated (as standard) on Leica's automated ST5020 and then transferred to the Bond RX for subsequent retrieval and staining. The antibodies were applied sequentially (CD8, anti-rabbit 546, CD68, anti-mouse 647), with washes in Leica's Bond Wash (AR9590) applied between each immunoreagent. All antibody incubations were for 30 min except the CD8, which incubated for 60 min. Antigen retrieval was performed with Tris EDTA (Leica, AR9640) at 100°C for 20 minutes. The specific antibody details are listed in Table 2.4.

Slides were counterstained in DAPI (Cell Signaling, 4083 – diluted to 10 μ g/ml in 1x PBS) for 5 min and mounted in Vector's VECTASHIELD HardSet Antifade Mounting Medium, reference H-1400.

Slides were imaged using a Leica SP5 laser scanning confocal microscope (Leica Microsystems Inc, Wetzlar, Germany). UV(405nm), Argon(543nm) and HeNe(633nm) lasers

Target	Catalogue No.	Dilution/Conc.	Secondary	Dilution
CD8	Thermo, RM-9116	1:25	Anti-Rabbit 546 (Invitrogen, A11010)	1:250
CD68	Novocastra, NCL-L-CD68	1:25	Anti-Mouse 647 (Jackson, 715-607-003)	1:250

Table 2.4 Antibodies used for Dual IF CD8/CD68 staining with DAPI counterstain

were used to excite the fluorophores. Excitation and emission spectra are shown in Figures 5.4.

2.6 Imaging Mass Cytometry

2.6.1 Imaging Mass Cytometry (IMC) Antibody purification

Carrier proteins like bovine serum albumin (BSA) are commonly added to the storage buffer of antibodies as preservatives. These proteins can replace antibodies in the conjugation process, reducing the efficiency of antibody conjugation to metals. Antibodies with carrier proteins were therefore purified using the Pierce™ Antibody Clean-up Kit (ThermoFisher, 44600) following the manufacturer's instruction. This involves a desalting step, where the antibody is passed through a Zeba spin desalting column containing a resin bed and centrifuged for 2 min at 1000×g. The antibody was then added to a spin column containing a Purification Support slurry, incubated at room temperature for 5 min with end-to-end mixing, then centrifuged for 1 min at 4000×g to collect the purified antibody. The purity of the antibody was tested using a NuPAGE™4-12% Bis-Tris Gel (Invitrogen, NP0321BOX). The antibody was mixed with LDS Sample buffer (Invitrogen, NP0007) and Sample Reducing agent (Invitrogen, NP0009) and incubated at 70°C for 10 min. 1 μg of the original antibody and the purified antibody were loaded onto the gel, with a BSA-only sample used as a control. The gel was run in Bolt MPOS SDS Running Buffer (Invitrogen, B0001) at 120V for 30 min followed by 2h at 60V. The gel was stained with InstantBlue (Expedeon, ISB1L) to visualise the protein bands, and imaged using Antibody-Metal conjugation. Carrier-free antibodies were metal-conjugated using the Maxpar®X8 Antibody Labelling Kits (Fluidigm) according to the manufacturer's instructions. A proprietary polymer is loaded with a lanthanide metal by mixing them in the provided L-buffer and incubating them in a 37°C water bath for 35 min. The metal-loaded polymer is filtered to remove excess metal by centrifuging it twice in a 30kDa Amicon Ultra500 μV bottom (Millipore, UFC505096) at 14,000×g for a total of 55 min. The antibody is centrifuged at 14,000×g in a 50kDa Amicon Ultra500

μ lV bottom (Millipore, UFC505096) for 10 min, and then partially reduced by incubating it in 4mM Pierce™ Bond-Breaker®TCEP Solution (ThermoFisher Scientific,77720) at 37°C for 30 min. This is immediately followed by three washes in the provided C-buffer, with a 10 min centrifugation at 14,000×g after each wash. The metal-loaded polymer is added to the partially reduced antibody and incubated in a 37°C water bath for 90 min. The metal conjugated antibody is washed four times in the provided W-buffer, with a 10 min centrifugation at 14,000×g after each wash. The antibody is finally suspended in 100 μ l PBS containing 0.05% sodium azide as a preservative, and collected by centrifuging the inverted filter in a fresh microcentrifuge tube for 2 min at 1,000×g. The concentration of the antibody was measured using NanoDrop Spectrophotometer ND-1000 (LabTech International).

2.6.2 Staining and Imaging

The manual staining protocol outlined in section 2.3.2 was also followed for IMC. Metal-conjugated primary antibodies were used instead of unconjugated or HRP-conjugated antibodies. Details of all antibodies used are displayed in Table 2.5. Following the two washes in 1% Tween-TBS and two washes in TBS, the slides were counter-stained with Cell-ID™ Intercalator-Ir for 30 min at room temperature. The slides were then washed in Milli-Q water for 5 min, and air dried for at least 20 min before imaging on the Hyperion™ Imaging System (Fluidigm). Images were acquired using a UV laser at a frequency of 200 Hz. Metals in each pixel ablated were plasma-ionised and their count was detected using Cytometry by Time of Flight (CyTOF) technology. This was used to generate a stack of false-colour images, each layer corresponding to a single metal and the intensity of each pixel reflects the amount of each antigen present.

2.6.3 Image analysis

SEARCH

Definiens Tissue Studio was used for the tissue classification of the images from the SEARCH cohort. Definiens image analysis algorithms for detection of epithelial and stromal areas were trained and the segmentation for each core was manually refined by AM and a consultant gynaecological-histopathologist (J. McD.).

OV04 - Cytokeratin, H&E

QuPath 0.1.2 was used for TMA segmentation and the segmentation and classification of individual cells from images of Haemotoxylin and Eosin and Cytokeratin7 stained slides.

Antigen	Clone	Tag	Source	Code	Conc./Dilution ($\mu\text{g/ml}$)
CA9	polyclonal	141Pr	Novus	BioAF2188	2
p53	DO-7	143Nd	Fluidigm	3143026D	5
CD16	EPR16784	146Nd	Fluidigm	3146020D	10
CD163	EDHu-1	147Sm	Fluidigm	3147021D	5
PanK	C11	148Nd	Fluidigm	3148020D	1
CD56	MRQ-42	151Eu	Cell Marque	156R-96	1
CD45	D9M8I	152Sm	Fluidigm	3152018D	1.5
LAG3	D2G40	153Eu	Fluidigm	3153028D	2.5
TIM3	D5D5R	154Sm	Fluidigm	3154024D	2.5
FoxP3	236A/E7	155Gd	Fluidigm	3155016D	5
CD4	EPR6855	156Gd	Fluidigm	3156033D	5
CD68	KP-1	159Tb	Fluidigm	3159035D	2
PD1	D4W2J	160Gd	Cell Signalling	86163S	5
CD20	H1	161Dy	Fluidigm	3161029D	5
CD8	aC8/144B	162Dy	Fluidigm	3162034D	5
CD103	EP206	163Dy	Cell Marque	437R-16	3
CK7	RCK105	164Dy	Fluidigm	3164028D	10
pH2AX	N1431	165Ho	Fluidigm	3165036D	0.5
B7h4	D1M8I	166Er	CST14572	5	
GranzymeB	EPR20129-217	167Er	Fluidigm	3167021D	1
Ki67	B56	168Er	Fluidigm	3168022D	1.5
Collagen 1	Polyclonal	169Tm	Fluidigm	3169023D	1
CD3	Polyclonal	170Er	Fluidigm	3170019D	7.5
CTLA4	BSB88	171Yb	BioSB	BSB 2885	7.5
OX40	ACT-35	172Yb	CST98785	5	
CD45RO	UCHL1	173Yb	Fluidigm	3173016D	3
HLADR	TAL 1B5	174Yb	ThermoFisher	MA1-46109	0.5
ICOS	D1K2T	175Lu	CST	3148021D	5
Histone H3	D1H2	176Yb	Fluidigm	3176023D	0.25
Intercalator		191Ir, 193Ir	Fluidigm	201192A	1:1000

Table 2.5 IMC antibodies

TMA Dearthayer, Watershed cell detection algorithms and Random Tree classifiers were used. Smoothed features at $25\mu\text{m}$, $50\mu\text{m}$ and $100\mu\text{m}$ were generated, all 133 available features were used for the classifier and the labelled cells were split 50:50 into training and validation sets.

Collagen Structure Analysis

ImageJ and the OrientationJ and GLCM plugins were used for analysis of collagen structure images from SHG microscopy. Images were split into a 10×10 grid and the features calculated on each subimage before the median was calculated. Macros were written to automate this process.

BriTROC, OV04, ICON7 - IMC and IHC

Analysis of images taken from both chromogenic IHC and IMC was done using the HALO Image Analysis Platform (Indica labs). Halo was used for tissue classification and immune cell identification and extraction in OV04 and BriTROC cohorts.

2.7 Statistical Analysis and Mathematical Background

R (version 3.5.1) was used for statistical analysis and R markdown documents were generated for all analyses. These files are stored in a Github repository <https://bitbucket.org/jamesdbrenton/search-montfort/src/master/>.

2.7.1 Analysis of immune infiltrate

All count data were transformed to log base 10 after adding a small offset ($\frac{x_{min}}{2}$) to zero values. Continuous data were presented as median and interquartile range (IQR) and groups were compared by the Kruskal–Wallis, pairwise Kruskal-Wallis tests and Wilcoxon signed rank test was used to compare the mean infiltrate between groups. Discrete data were presented as count and percentage.

Quality assurance tests for spatial bias in staining across TMAs was carried out visually using heatmaps and statistically using Shapiro-Wilk tests. The same methods were used to investigate whether there was spatial bias in the areas of the cores.

2.7.2 Survival Analysis

Cox regression was used to examine the relationship between predictors (X_n) such as immune infiltrate and stage and patient survival. The functional form of immune variables was assessed using comparison with cubic splines. The best approximations to the functional forms were carried forward for the Cox models.

The hazard function is the measure of relative risk over time for a patient compared to the baseline. It is a function of other variables X_n which can be things like immune infiltrate, age or stage. This relative hazard $H(t)$ is given by the following formula:

$$H(t) = H_0(t)e^{b_1X_1+\dots+b_nX_n} \quad (2.1)$$

Where H is the hazard, the predictors are denoted X_n . It is the cox regression method that estimates the coefficients of the predictors b_n , ie the contribution of the variables to the survival of the patient. The proportional hazard model is semi-parametric as there is an assumption that the hazard is a linear function of the predictors and that the hazards are proportional (the coefficients are not a function of time). There are no assumptions about the form of the baseline hazard function $H_0(t)$.

The Akaike Information Criterion (AIC), equation 2.2, was used to compare the performance of survival models, which includes the loglikelihood of the model \hat{L} and a penalty on the number of terms, k to reduce overfitting.

$$\text{AIC} = 2k - 2\ln(\hat{L}) \quad (2.2)$$

2.7.3 Spatial Statistics

Spatial statistics can most easily be carried upon 2D point patterns. In such analyses cells are assumed to be points with coordinates (x, y) and zero radius $r = 0$. The value of the center of the point for each cell is given in equation 2.3.

$$x = \frac{x_{max} + x_{min}}{2}, y = \frac{y_{max} + y_{min}}{2} \quad (2.3)$$

We will see that these assumptions are violated but information can still be gleaned from these measurements as long as the consequences and limitations of such assumptions are understood.

The spatial patterns of a distribution can be split into first order properties (density variation) and second order properties (how points influence the location of others).

2.7.4 Density based analysis

Global density

The simplest measure of the density of a point pattern is the following:

$$\lambda = \frac{n}{|A|} \quad (2.4)$$

Where λ is the number of points in a region and A is the area measure of the region.

Ripley's K and L function

Ripley's K function measures the distribution of points within a certain radius and can be used to determine whether points are random, dispersed or clustered. It is defined as:

$$K(r) = \lambda^{-1}E \quad (2.5)$$

Where λ is the density and E is the expected number of points in the region. The expected number of points is

$$E = \sum_{i=j} \frac{I(d_{ij} < r)}{n} \quad (2.6)$$

where I is the operand function and d_{ij} is the distance between i^{th} and j^{th} points.

The L function is often used in data analysis because it is variance stabilized and is defined as

$$L(r) = \left(\frac{K(r)}{\pi} \right)^{\frac{1}{2}} \quad (2.7)$$

DBSCAN

DBSCAN stands for density based clustering of applications with noise. There are many clustering methods in existence but DBSCAN is one which is ideal for a situation where the number of clusters within the data is unknown. DBSCAN works on the idea of reachability of points rather than allocating points based solely upon the distance from a central point. DBSCAN also allows for some points to remain unallocated, a method which works well for cells within images, some of which will not lie within clusters. DBSCAN is better for clustering data points which have similar densities. This works well for identifying clusters of single cell types, as all stroma or all tumour cells will have similar spatial densities within the same tissue section. Due to its ability to exclude cells from clusters it is also shown to be less sensitive to noise than methods such as K means clustering. This is important in

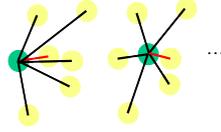


Fig. 2.3 Example of the nearest neighbour distance for two cells in this set are highlighted in red. This is calculated for all cells/points in an image.

applications to real data with noise and to cell classification algorithms with any accuracy below 100% [146].

2.7.5 Distance based analysis

Average Nearest Neighbour

This analysis measures the distance from each point to its nearest neighbouring point and takes the average over all cells/points. An example of cells within a small population and their nearest neighbour distance is shown in Figure 2.3

k-nearest neighbour

k-nearest neighbour method derives for each point in a point pattern, the k^{th} nearest neighbour to it. The theoretical expected distance to the nearest neighbour in a random 2D point pattern is given by

$$\langle r_n \rangle = \frac{1}{2\sqrt{\lambda}} \quad (2.8)$$

Where λ is the global density of points.

Batty's entropy

Batty's entropy was calculated using **R** package **SpatEntropy** [147, 148]. It is defined as follows; if a phenomenon of interest F occurs over an observation window of size T partitioned into G areas of size T_g . This defines G dummy variables identifying the occurrence of F over a generic area $g, g = 1, \dots, G$. Given that F occurs over the window, its occurrence in area g takes place with probability p_g , where $1 - p_g = \sum_{g' \neq g} p_{g'}$ and $\sum_g p_g = 1$. The phenomenon intensity is obtained as $\lambda_g = p_g/T_g$, where T_g is the area size, and is assumed constant within each area. Batty's spatial entropy is

$$H_B(F) = \sum_{g=1}^G p_g \log\left(\frac{T_g}{p_g}\right) \quad (2.9)$$

Batty's entropy is high when the phenomenon of interest F is equally intense over the G areas partitioning the observation window (i.e., when $\lambda_g = \lambda$ for all g). Batty's entropy $H_B(F)$ reaches a minimum value equal to $\log(T_{g^*})$ when $p_{g^*} = 1$ and $p_g = 0$ for all $g \neq g^*$, with g^* denoting the area with the smallest size.

2.7.6 Intensity analysis

The Grey Level Co-occurrence Matrix (GLCM) measures texture features based on second order statistics based upon greyscale pixel intensities[149]. Initially the pixel values in the image are scaled so they have integer values 1 to 8. The values are then normalised to the range $[0, 1]$, the GLCM can then be calculated.

P_{ij} is the i, j element of the normalized GLCM. Each entry P_{ij} is, by default, the number of pairs of horizontally neighbouring pixels in the scaled images with values i and j . Given that the scaled image now only has 8 possible pixel values between 0 and 1, there are 8×8 possible values for the pixels in the neighbouring pairs. The GLCM is therefore an 8×8 matrix which is filled with the counts of each possible neighbouring pixel pair. For example, P_{11} is the number of pairs of horizontally neighbouring pixels in the scaled image both with a value of 1. Offsets other than horizontal neighbours can be used. μ is the mean of the GLCM. The statistics that are commonly calculated on the GLCM[150] include the following:

Entropy

Entropy is a measure of information mixing. Increasing disorder increases entropy. Generally Shannon entropy is defined in Equation 2.10 where p_i is the probability of obtaining class i summed over all possible classes.

$$H = - \sum_{i=1} p_i \log p_i \quad (2.10)$$

Entropy can also be defined on an image or image region based on pixel grayscale values. In the GLCM notation, entropy is:

$$\text{Entropy} = - \sum_{i,j=0}^{N-1} P_{ij} \ln P_{i,j} \quad (2.11)$$

Energy

Sum of squared elements in the GLCM that measures homogeneity. A constant image has Energy=1.

$$\sum_{i,j=0}^{N-1} (P_{ij}^2) \quad (2.12)$$

Contrast

Local variation in intensity, measuring the difference in neighbouring pixel values. This measure is the sum of all neighbouring pixel differences squared.

$$\sum_{i,j=0}^{N-1} P_{i,j}(i-j)^2 \quad (2.13)$$

Homogeneity

This measures the tightness of the distribution of pixel values.

$$\sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1+(i-j)^2} \quad (2.14)$$

Correlation

This measures the correlation between neighbouring pixels; a perfectly correlated image where all pixels have the mean value would have correlation = 1.

$$\sum_{i,j=0}^{N-1} P_{ij} \frac{(i-\mu)(j-\mu)}{\sigma^2} \quad (2.15)$$

Chapter 3

Digital pathology and the quantification of immune infiltrate

This work was accepted for publication in April 2020. See Appendix A.

3.1 Introduction

Modern approaches to pathology have begun to digitise and automate measures of immune cell responses in a wide variety of tumours. The availability to normalise densities, apply reproducible automated classification and obtain higher resolution data offers a large number of possibilities for analysis[151]. The real strengths of such digital pathology approaches lie in the more complex questions around quantitation and spatial resolution of such immune responses. Furthermore, as the number of simultaneous populations we can measure in a single sample increases with technology and the resolution and accuracy of those measures increases, the reliance on digital pathology methods will increase and importantly must find new ways to interpret the multi-dimensional data that results.

In this chapter I aimed to develop a robust statistical workflow for the analysis of quantitative digital pathology data. The TMAs were investigated for bias; the cores that were outlying in size or infiltration were extracted and examined. The infiltration densities were calculated and normalised such that they were appropriate for a variety of statistical tests.

I worked to complete density based survival analyses on a dataset from the SEARCH cohort that contained the density of CD8⁺, CD45RO⁺ and CD68⁺ immune cells in epithelium and stroma for each patient. I then expanded upon this density analysis to ask questions about the properties of the multi-dimensional immune response and worked on optimising complex survival model building and interpreting biologically patterns of immune variation. I analysed

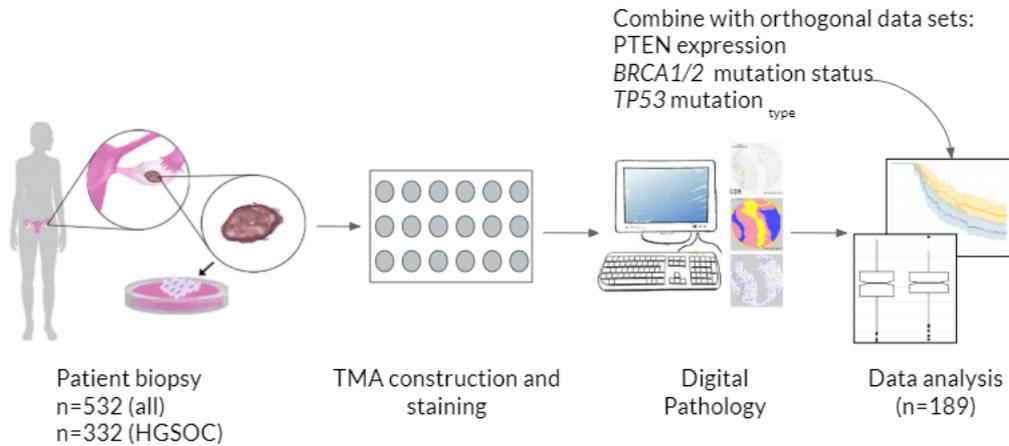


Fig. 3.1 Visual abstract - The project

varying functional forms of infiltrate density, combined multiple infiltrates with principal component analysis for the first time to investigate the underlying patterns of biology and did further analyses including quantifying and assessing the immune exclusion in samples. I also applied some similar analyses to the other subtypes of ovarian cancer in the cohort.

Overall I aimed to address and understand the relative importance of the infiltration in different compartments in the tissue, to understand the impact of both the stromal and epithelial infiltration and to quantify and interpret the relationships between them.

3.1.1 History of the project

This particular part of the work began from collaboration with Anne Montfort(AM) who approached me with raw files from Definiens™ image analysis software. Images of CD8⁺, CD68⁺ and CD45RO⁺ stained slides had been segmented and regions classified as tumour and stroma. The specific contributions of others to this chapter are outlined in Methods.

3.2 SEARCH Cohort

Samples from 570 patients from the prospective SEARCH ovarian cancer population-based study had been used to construct tissue microarrays. Ethical approval for the SEARCH study had previously been granted by the Eastern Multicenter Research Ethics Committee (Reference 99/5/007) and the larger cohort is discussed in detail by Song *et al.*[152]. Among the samples from 570 patients with primary epithelial ovarian tumours, 332 were high grade serous ovarian cancer patients. Patients were staged as having localized, regional

or distant disease (L/R/D); Walters *et al.* describe the mapping between LRD staging and FIGO[153, 154].

3.3 Methodology

3.3.1 Experimental methodology

Immunohistochemistry

Immunohistochemistry on the SEARCH samples was carried out upon the tissue sections by Anne Montfort (AM) according to the protocol in section 2.3.2. Previously published PTEN immunostaining data was used where high PTEN expression was considered to be positive staining and low expression to be weak, heterogeneous or negative staining respectively[87].

TP53 mutation data

Sequencing and mutation analysis was carried out prior to this PhD by Anna Piskorz (AP). The coding regions of *TP53* were sequenced by tagged-amplicon next generation sequencing (TamSeq) as previously described[155] and confirmed by immunohistochemical analysis using a 4-tier core system in which expression is scored as overexpression, complete absence, cytoplasmic or wild type[156]. Sequencing of germline mutations in the *BRCA1* and *BRCA2* genes was performed as previously described[157]. Some samples were resequenced using TamSeq by SAK following my initial exploratory analysis and datacleaning steps. The results of resequencing were integrated into the final analysis.

Tissue classification

The initial utilization of Definiens™ software was carried out by AM. Definiens™ was used to segment and classify the tissue in images of H-DAB stained tissue as epithelium or stroma; classifications were reviewed by a gynaecological pathologist (JMcD). The same software was used to quantify the immune populations in these tissues.

Immune cell quantification

The number of CD8⁺ and CD45RO⁺ cells in epithelial and stroma areas, the area of epithelial and stromal areas covered by CD68 staining and the total area of epithelium and stroma in each core, had been digitally determined using the Tissue Studio software (Definiens™).

Figure 3.2 shows examples of the cell segmentation and tissue classification.

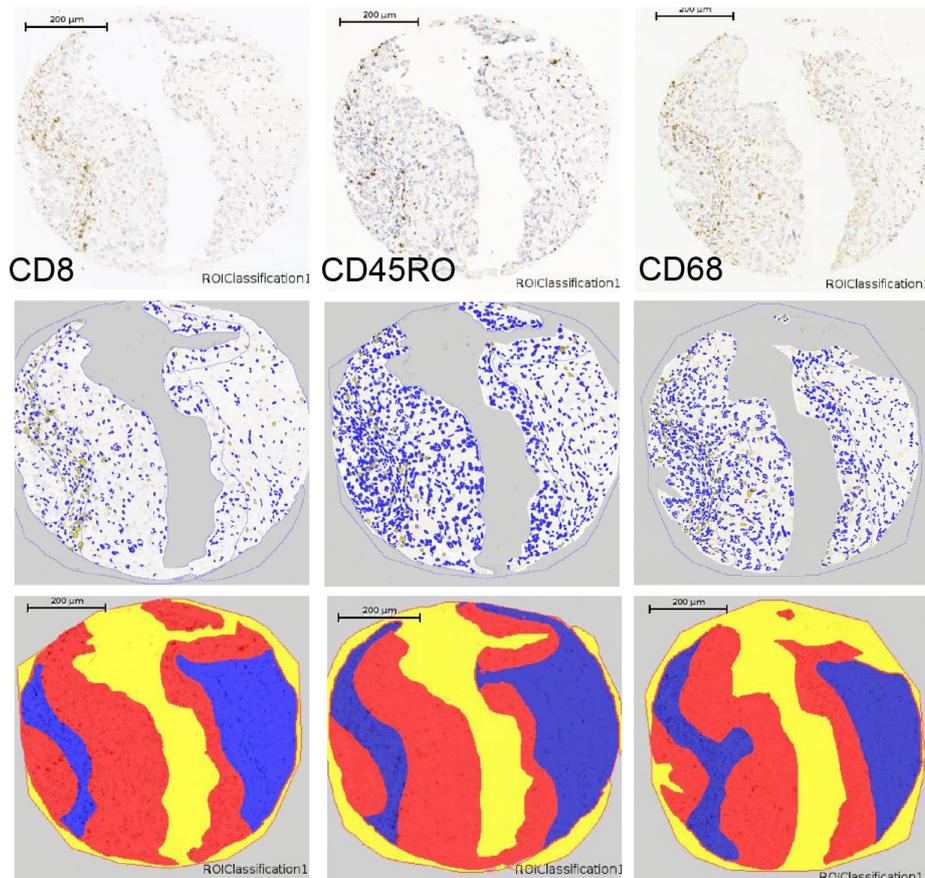


Fig. 3.2 Examples of the segmentation of tissues and identification of cells in the Definiens software.

3.3.2 Statistical analyses

The clinical variables of age at diagnosis, menopause status and stage were available for the cohort and I included them in the analysis. I used univariable Cox regressions to identify best-fitting variables for the final multivariable Cox regression model. The refined model was compared with a combined multivariable Cox regression model including all immune infiltrates. Hazard ratios (HR) refer to a single unit increase in continuous variables. The proportional hazards assumption was tested and satisfied in all cases using Schoenfeld residuals. The Kaplan–Meier analysis (with log-rank test) was applied to illustrate survival differences graphically. Two-sided P-values <0.05 were used to indicate statistical significance.

Principal component analysis (PCA) using the R package **prcomp** was used to extract the independent components of variance between patients. The package **prcomp** uses singular value decomposition and the variables were scaled to have unit variance before creating composite linear independent variables. These were then passed forward to the survival

modelling. The Akaike Information Criterion (AIC), equation 2.2 was used to compare the performance of survival models.

Bonferroni p-value corrections were carried out for all multiple testing. $P < 0.05$ was considered significant for all analyses.

3.4 Results

The SEARCH cohort included patients with HGSOC, Clear Cell, Endometrioid, Mucinous and Low Grade Serous Ovarian Cancer. These subtypes of Ovarian Cancer have vastly different cells of origin, different patterns of infiltration and different genomic drivers and consequently different survival. As I was interested in survival and primarily interested in understanding the differences within populations rather than across subtypes, I extracted the HGSOC subpopulation for the initial analysis.

3.4.1 Patient characteristics

Figure 3.3 shows the REMARK diagram for this study and Table 3.1 shows the clinical characteristics of the 332 HGSOC patients from the study cohort. Immunohistochemical analyses on tissue microarrays (TMAs) were performed to detect CD8⁺, CD45RO⁺ and CD68⁺ cells in tissue cores from primary ovarian specimens. Existing genomic data for somatic *TP53* and germline *BRCA* mutations was available for some cases and was incorporated to investigate the relationship between these genomic markers and immune infiltration. PTEN expression had also been previously assessed for a subset of SEARCH patients and had been found to be prognostic[21]. PTEN expression data was incorporated in order to investigate whether this association with prognosis was due to changes in immune infiltration in HGSOC as seen in melanoma[158].

152 HGSOC cases were available for analysis after quality assurance, data cleaning and the reduction of the data set to only cases with complete results for CD8, CD45RO and CD68 staining in both epithelium and stroma as well as survival data. The characteristics for the 152 patients with epithelial and stromal tissue for all immune markers are shown in Table 3.2. The distribution of clinical characteristics does not appear to vary between the original 332 patients and the complete subset of 152.

Tagged amplicon sequencing data was available for 248 cases and *TP53* mutation was detected in 231 samples (93%) (Table 3.1). The number of patients with loss of function(LOF) or gain of function(GOF) mutations to the *TP53* gene are given in the table. Previously

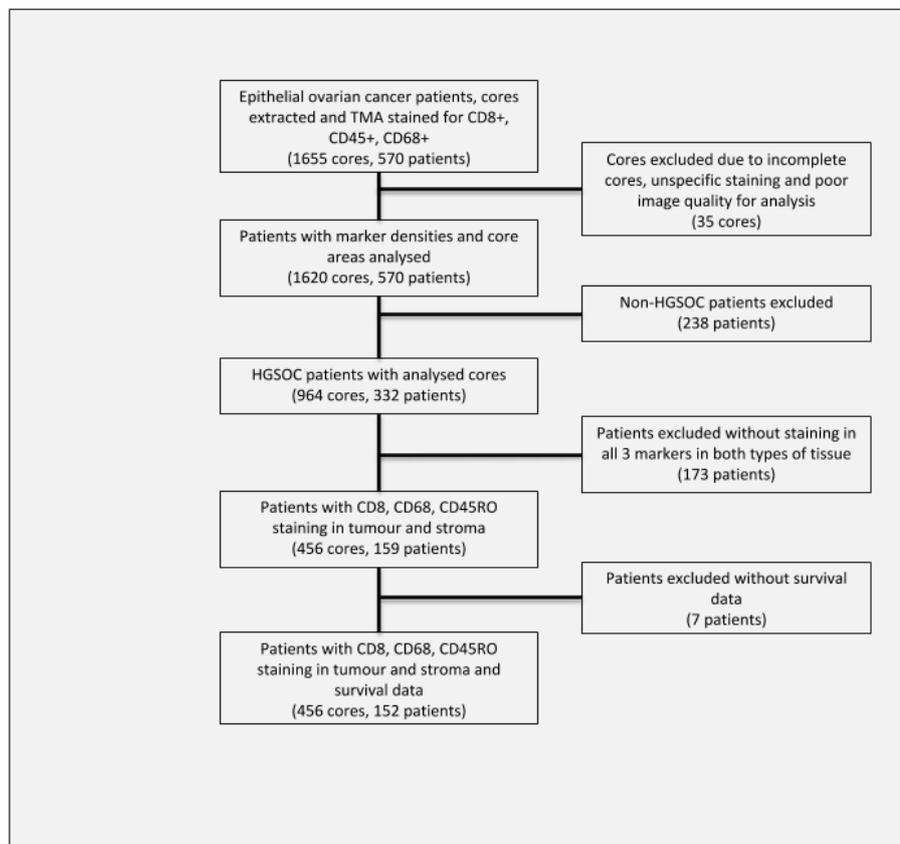


Fig. 3.3 REMARK diagram for the SEARCH cohort.

published data for germline *BRCA1* and *BRCA2* mutation and PTEN expression were available for 297 and 155 cases respectively[87, 157].

3.4.2 Data cleaning, quality assurance and spatial bias

Analysing a large number of digitally imaged and segmented TMA cores gave me the ability to carry out quality checking for bias across TMAs and to evaluate the variation in properties of the tissue core such as tissue area, something typically not measured by pathologists when assessing cores manually.

I examined the tissue area and assessed the TMAs for bias. I did this visually using heatmaps (Figure 3.4) and statistically using Shapiro-Wilk tests ($p > 0.05$). The distribution

N		332
Median Age (IQR)		58.0 (51.0-64.0)
Stage	localized	64 (19.3%)
	regional	42 (12.7%)
	distant	202 (60.8%)
	unstaged	24 (7.2%)
<i>TP53</i> mutation	GOF	137 (55.2%)
	LOF	94 (37.9%)
	wild type	17 (6.9%)
	Not assessed	84
PTEN IF	High	28 (18.1%)
	Low	127 (81.9%)
	Not available	177
Germline <i>BRCA</i> status	wild type	256 (86.2%)
	<i>BRCA1</i>	18 (6.1%)
	<i>BRCA2</i>	23 (7.7%)
	Not available	35

Table 3.1 Patient characteristics for the HGSOc subset of the SEARCH cohort. Table shows the number of patients with gain of function(GOF) or loss of function(LOF) *TP53* mutations, germline *BRCA* mutations of each type and high and low PTEN expression measured by immunofluorescence(IF).

N		152
Median Age (IQR)		58.0 (50.0-64.0)
Stage	localized	27 (17.8%)
	regional	22 (14.5%)
	distant	93 (61.2%)
	unstaged	10 (6.6%)
<i>TP53</i> mutation	GOF	55 (52.4%)
	LOF	42 (40.0%)
	wild type	8 (7.6%)
	Not assessed	47
PTEN IF	High	11 (16.4%)
	Low	56 (83.6%)
	Not available	85
Germline <i>BRCA</i> status	wild type	113 (84.3%)
	<i>BRCA1</i>	10 (7.5%)
	<i>BRCA2</i>	11 (8.2%)
	Not available	18

Table 3.2 Patient characteristics for the subset of HGSOc patients in the SEARCH cohort with complete data (n=152). No statistical difference in clinical characteristics was found between this subset and the full HGSOc cohort.

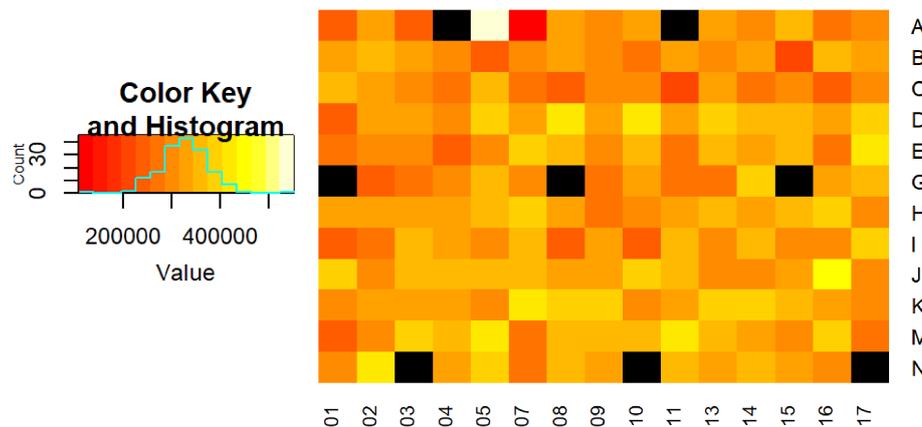


Fig. 3.4 Heatmap of total core area across TMA. The histogram on the left represents the histogram of core areas and the colour assigned to each value. Whiter areas show large areas of tissue, Red show small areas of tissue, black squares illustrate missing or control cores. No obvious spatial bias is seen.

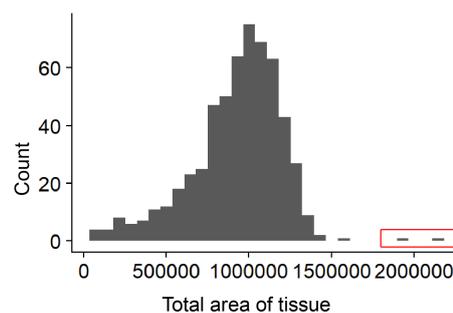


Fig. 3.5 Histogram of the distribution of total tissue area for each core in the SEARCH cohort. Distribution is skewed as parts of cores are easily lost but once tissue is extracted area cannot increase. Outlying cores with significantly larger areas are highlighted in red.

of tissue areas of tumour and stroma are shown in Figure 3.5. As expected these are negatively correlated.

As mentioned, $CD8^+$ and $CD45RO^+$ counts were initially calculated as number of cells in stroma and epithelium and $CD68^+$ scoring was given as tissue area stained. I normalised this raw data; converting cell counts into cells per unit area (μmm^2 and $CD68^+$ coverage into percentage area. This normalization to area allows for comparison between patients whose tissue is made up of varying quantities of stroma and epithelium. I also confirmed using residuals that transforming the count densities to base 10 would normalize them.

Given the large variation in areas of cores I extracted outlying examples of cores for inspection. It was important to verify that core area and position within the TMA was independent of immune density. Figure 3.6 shows a heatmap comparing infiltration density

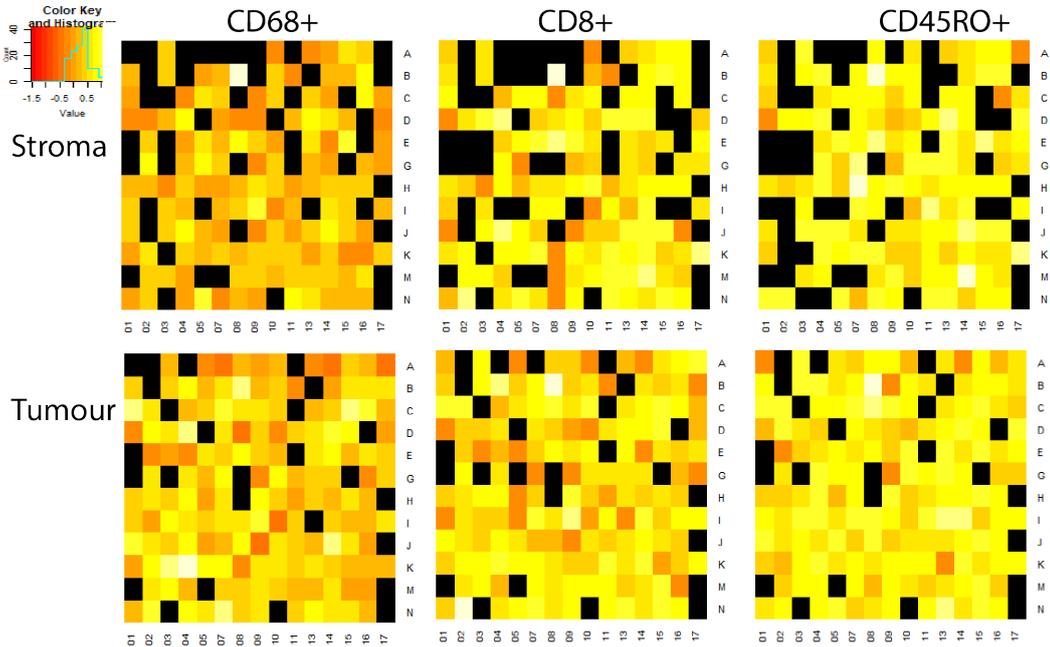


Fig. 3.6 Heatmap of the density of each infiltrate for each position in the SEARCH TMA. The histogram in the top left represents the distribution over percentage areas stained with CD68 and the colour assigned to each value. White areas show large areas of tissue. Red squares represent small areas of tissue. Black squares are either missing data for that tissue type in a core, entirely missing cores from the TMA or control cores. No spatial bias is seen in the individual TMAs.

across the $CD8^+$, $CD68^+$ and $CD45RO^+$ TMAs. No spatial relationship was observed between the relative location in each TMA of a core and the density of immune cells or the tissue area.

3.4.3 Stroma and tumour composition of cores

TMA blocks are built by pathologists who outline a region on H&E blocks for the extraction of sample tissue. Cores are taken from FFPE blocks in these positions for the construction of TMAs. Other than typically aiming to sample regions with high percentages of epithelium, this sampling is otherwise left to the pathologist. Digital pathology and tissue segmentation provides a unique opportunity to evaluate the sampling of stromal and epithelial areas and the makeup of sampled tissue cores.

I wanted to investigate the variation in epithelial and stromal tissue that was sampled in the SEARCH TMAs. The distribution of epithelial and stromal areas are shown in Figure 3.7. An example of the segmentation in Definiens used to obtain these areas was previously

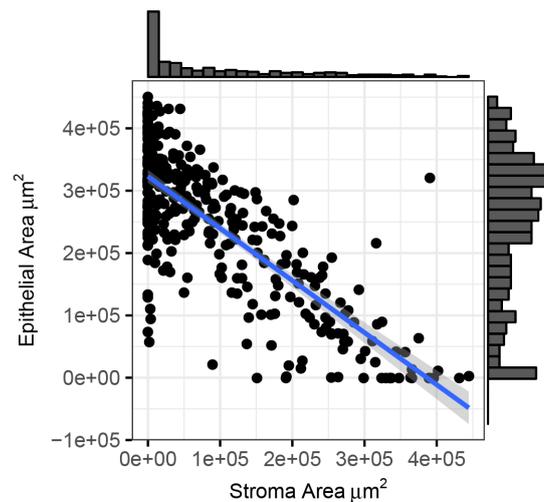


Fig. 3.7 Scatter plot and linear fit for the epithelial and stromal areas of each core. Histograms of the distribution are shown on the axes. Stroma and epithelial area are negatively correlated. Stroma area is strongly negatively skewed, epithelial area is a normal distribution with zero inflation.

shown in Figure 3.2.

As a validation I compared stromal and epithelial area. These were negatively correlated as expected. Stromal area is strongly negatively skewed, epithelial area follows a normal distribution with zero inflation. I found that of 964 images representing 332 HGSOC patients, 69 patients (20.8%) had images which contained malignant epithelium but no stroma; 250 patients (75.3%) had images which contained epithelium and >1% adjacent stroma and 13 patients (3.9%) had images containing no tumour epithelium (Fig 3.7). The median proportion of epithelium and stroma was 85.1% (IQR 51–100%) and 14.9% (IQR 0–49%) respectively.

3.4.4 Mutant allele fraction and epithelial content

Mutant allele fraction (MAF) is the percentage of cells sequenced with a *TP53* mutation. Mutant allele fraction had been previously generated for all samples. The distribution of mutant allele fraction is plotted as a histogram in Figure 3.8. The distribution had some zero inflation. I used this distribution to ensure that the sequencing of all samples classified as HGSOC was of high enough quality. Samples with mutant allele fraction of zero were resequenced and if no mutation was found were excluded from HGSOC analysis.

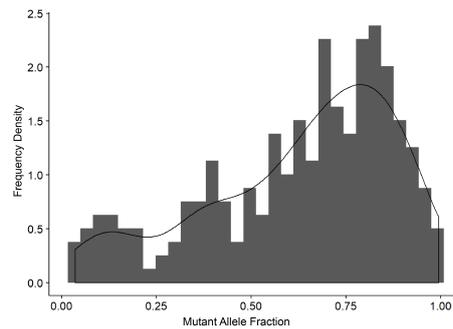


Fig. 3.8 Histogram of counts of samples against mutant allele fraction (MAF). Density of distribution plotted on top of the histogram. MAF is normalised such that a MAF of 1 means 100% of cells in the sample contain a *TP53* mutation.

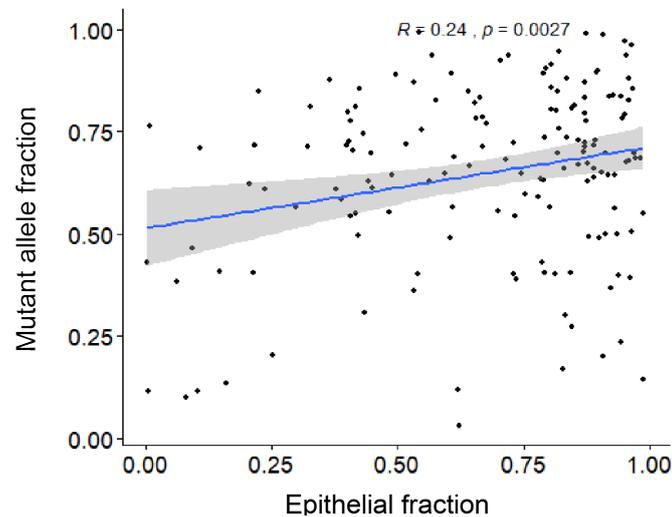


Fig. 3.9 *TP53* mutant allele fraction against epithelial area of a core. A weak correlation is observed.

As *TP53* mutations are found ubiquitously in HGSOC epithelium we would expect that the proportion of epithelial cells in a sample is correlated with *TP53* mutant allele fraction. The mutant allele fraction (MAF) and area of epithelium are shown in Figure 3.9. MAF and epithelial area were positively correlated ($R = 0.24, p = 0.0027$). The sequencing and imaging are not carried out on exactly the same tumour region and so I would not expect a perfect correlation due to natural noise from sampling. The correlation between technical replicates for allele frequency and sequencing depth is shown in Figure 3.10.

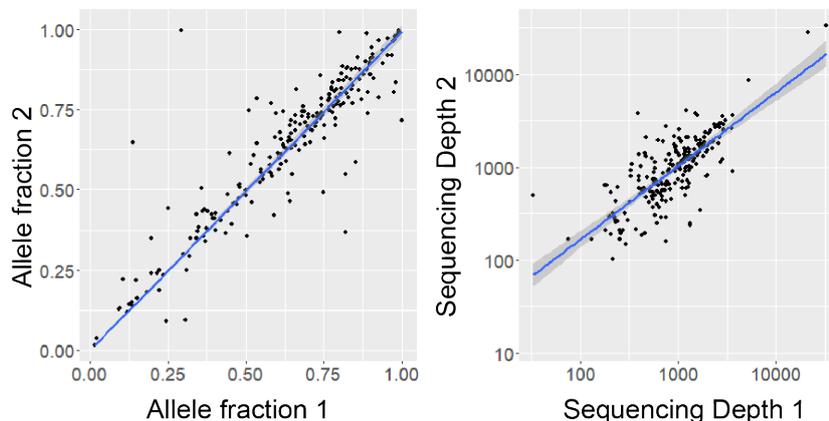


Fig. 3.10 *TP53* allele fraction and sequencing depth for technical replicates.

3.4.5 Immune cell densities

An example of the segmentation of $CD8^+$ T cell, $CD45RO^+$ memory lymphocyte and $CD68^+$ macrophages and the allocation of the stromal and epithelial regions was shown previously in Figure 3.2. These cell counts and tissue areas were the data used to derive densities for each of the immune markers.

As demonstrated, cores sample a wide range of areas of epithelium and stroma and therefore sample different proportions of epithelial and stromal infiltrate. In order to understand the impact of this sampling it was important to know how the density of infiltration varies between epithelial and stromal compartments. The distribution of densities of immune populations within tumour epithelium and stromal areas were compared (Fig. 3.11). I found that the density of $CD8^+$ and $CD45RO^+$ cells were significantly higher in stroma than tumour epithelium ($p = 0.005$ and $p = 0.004$ respectively; Welch's t-test) but not significantly different for $CD68^+$ cells.

The predominant population of immune cells studied in the literature is epithelial $CD8^+$ infiltrate which has been demonstrated to have a log-linear relationship with survival. In order to evaluate the importance of this infiltrate in the multi-immune infiltrated microenvironmental context I investigated whether this infiltrate was independent of others. If this infiltrate is not independent, the survival impact may be an indirect readout of the presence of other infiltrates. The correlation between infiltrates is also important to understand whether there exist spatial dynamics of infiltration and the relationship and movement between adjacent tissue compartments. The quantitative measure of immune density in our samples allowed me to investigate the Pearson correlation coefficient between quantities of immune infiltrate. The quantity of the three immune populations in this cohort showed moderate to strong

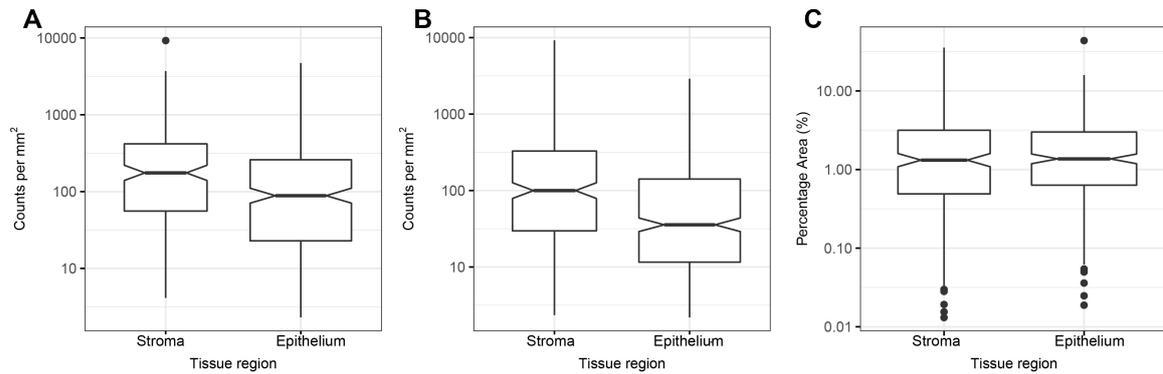


Fig. 3.11 (A), (B) and (C) show the distribution of CD8⁺, CD45RO⁺ and CD68⁺ immune cell densities in epithelium and stroma. CD8⁺ and CD45RO⁺ densities were defined as counts per mm². CD68⁺ density is defined as the percentage of tissue stained for this marker (Notches on box plots extend $1.58 \times \text{IQR} / \sqrt{n}$ and approximate the 95% confidence interval for the median. Box plot whiskers extend to $1.5 \times \text{IQR}$.)

correlation between infiltrate in epithelium and stroma and between the three infiltrates (Figure 3.12).

Given that we see varying epithelial cell percentages and increased infiltrate in the stroma as compared with the tumour I was interested in whether the density of immune infiltrate reaching the epithelium was related to the epithelial percentage of the core. In other words, does increasing the quantity of tumour in a sample increase the infiltration, perhaps due to the presence of more tumour antigens and thereby more immune cell recruitment or decrease it due to reduced stromal access and infiltration. I examined the relationship between the fraction of tumour in a core and the density of immune infiltration in the epithelium using Pearson correlation test. Intra-epithelial CD8⁺ and CD45RO⁺ densities were weakly positively correlated with the purity/tumour fraction of the sample ($R^2 = 0.17$, $p = 0.003$ and $R^2 = 0.16$, $p = 0.006$) whereas CD68⁺ epithelial density was not.

3.4.6 Immune exclusion

When considering patterns in immune infiltration, beyond the binary of present and absent, the terms immune-inflamed, immune-desert and immune-excluded have been used to describe varying T-cell infiltration based on histological and transcriptional analyses [23, 102]. Immune-inflamed and immune-desert patterns reflect high positive or negative correlations between all infiltrates but T-cell exclusion describes tumours where CD8⁺ cells are significantly absent from tumour epithelium whilst still being present in the surrounding stroma [23].

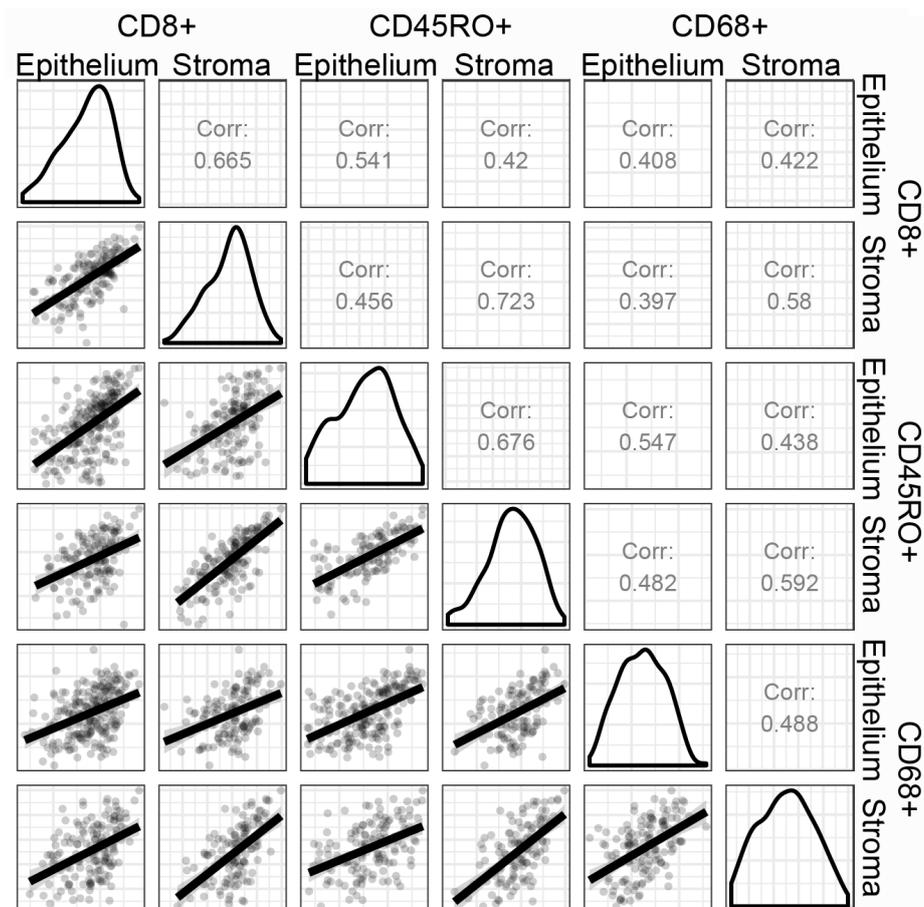


Fig. 3.12 Scatterplots and distributions of CD68⁺, CD45RO⁺ and CD8⁺ infiltrate in Stroma and Epithelium. The quantities of all infiltrates were correlated between matched samples across both tissue region classes.

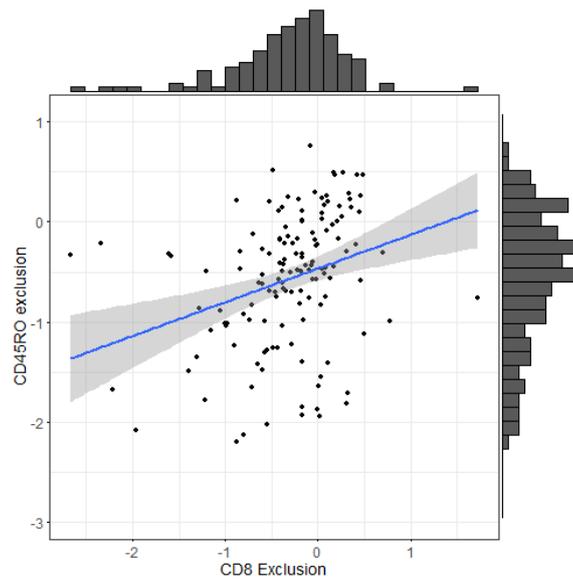


Fig. 3.13 Immune exclusion ratio of CD8 and CD45RO T-cells. Ratios are log-transformed like the counts. Histograms of exclusion are shown on the axes of the graph, both distributions are log-normal.

On average there is more infiltration in the stroma than epithelial compartments as shown in Figure 3.7 but the plot in Figure 3.13 shows that the ratio of epithelial:stromal infiltration is log-normal. As the counts are log-transformed, the ratio of infiltration density is equal to the difference between the logs of the epithelial and stromal infiltration. I defined this ratio as the extent of epithelial exclusion. The standard deviation of the ratio of CD8⁺ and CD45RO⁺ epithelial:stromal infiltration were both 0.68.

Within the patients that had epithelial and stromal tissue and therefore immune staining for all markers 19 (10.1%) cases had 10 × as much stromal CD8⁺ as epithelial CD8⁺ and 39 (20.5%) cases had 10 × as much stromal CD45RO⁺ as epithelial CD45RO⁺. Only 3 cases had 10× as much CD68⁺ stromal infiltrate as epithelial infiltrate. All ratios were normally distributed and the CD8⁺ and CD45RO⁺ exclusion was weakly correlated ($R = 0.2$, $p = 0.009$).

3.4.7 Functional form of infiltrates in building a survival model

The aim of this project was to understand some of the impact of these infiltrate metrics on survival. When modelling survival using Cox regression, one must assess that the assumptions of the Cox model are met and decide upon the functional form of the predictor (Section 2.7.2). I utilized Martingale residuals (Figure 3.14) to confirm that a base 10 transformation addressed the non-linearity of the data. In order to find the best functional

	Linear p-value	Cubic splines p-value	Log(base 10) p-value
Age at diagnosis	0.18	0.36	0.24
CD8 ⁺ epithelium	0.43	0.36	0.25
CD8 ⁺ stroma	0.4	0.17	0.64
CD68 ⁺ epithelium	0.44	0.57	0.63
CD68 ⁺ stroma	0.09	0.06	0.009
CD45RO ⁺ epithelium	0.3	0.09	0.07
CD45RO ⁺ stroma	0.07	0.016	0.002

Table 3.3 P-values for the association of each functional form of the variable with survival using univariable Cox regression analyses. Lowest p-values demonstrate most likely functional form and show that each relationship is approximately log linear

form for each predictor I modelled the univariate relationship between the predictor and survival using Cox regression and plotting the residuals of the model fit against the expected form. The relationship between each of the immune variables and survival was found to be approximately log-linear. The only clinical variables accompanying the cohort were age at diagnosis, stage and menopause status. Age at diagnosis is the only continuous variable of these and the relationship between age and survival was found to be approximately linear (Figure 3.15, Table 3.3).

3.4.8 Stromal CD68⁺ and CD45RO⁺ infiltrate are the strongest individual prognostic markers

Initially I investigated whether survival models of the individual infiltrates in stroma and tumour accurately predicted the data. Univariable analysis showed improved survival with increasing stromal density of CD45RO⁺ (HR 0.76 95% CI 0.65–0.90, p=0.001) and CD68⁺ (HR 0.53 95% CI 0.34–0.81, p=0.003) (Table 3.4). Modelling each immune variable with stage showed improved predictive value for epithelial CD8⁺ density (HR=0.83, p-value=0.027) as well as stromal CD68⁺ and CD45RO⁺ density and epithelial CD45RO⁺ density (Table 3.4). Kaplan-Meier curves require an arbitrary cutpoint for continuous data that impacts reported significance, as such I used these for illustrative purposes only. Figure 3.16 shows illustrative Kaplan-Meier survival curves for high and low stromal and epithelial CD68⁺, CD45RO⁺ and CD8⁺ densities.

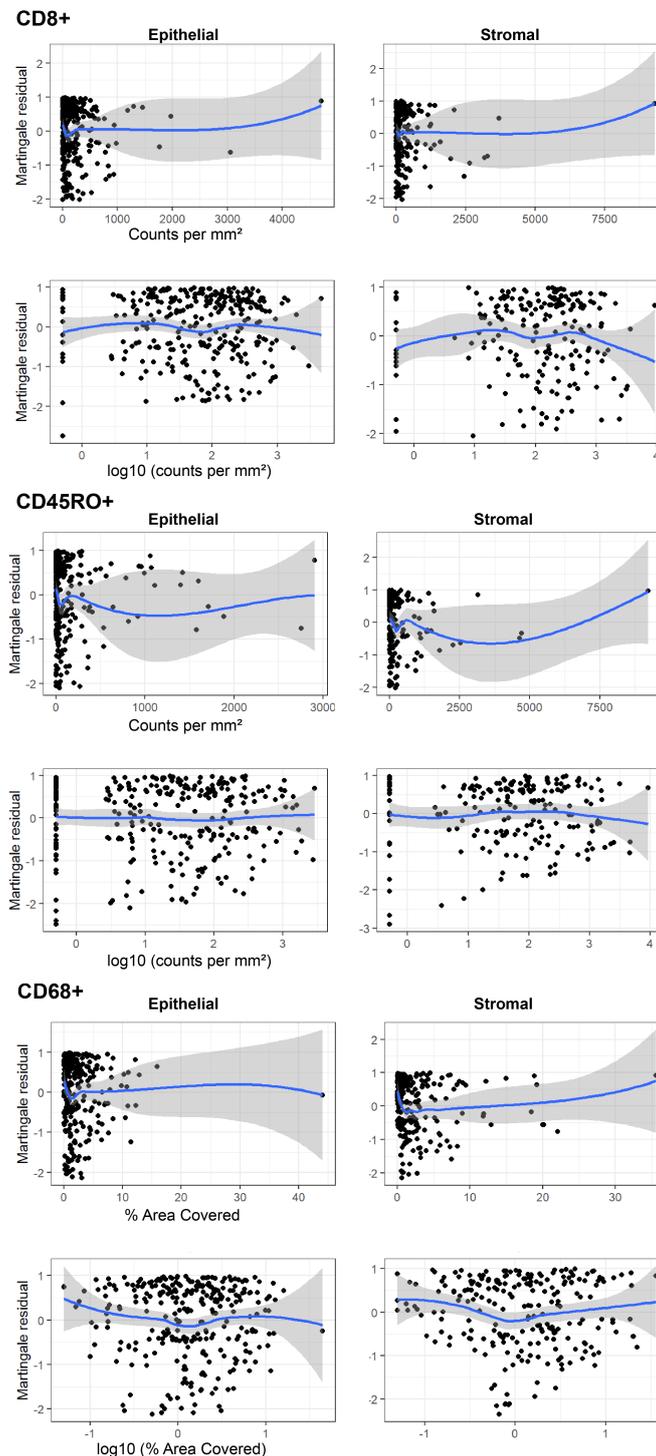


Fig. 3.14 Martingale residuals of each infiltrate assesses the linearity of the variables. Residuals are the deviance of data from a perfectly fitted model. If the data is linear the residuals will be symmetrically distributed. The log10 transformation is seen to reduce the distance and improve symmetry of residuals from the perfect fit and therefore the linearity in all cases.

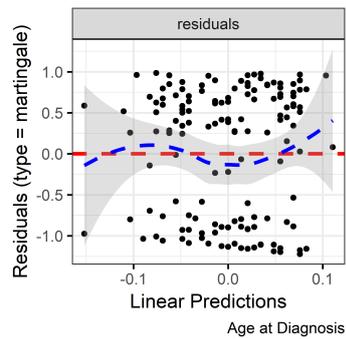


Fig. 3.15 Residuals of model fit plotted against predicted value for age at diagnosis, the relationship was found to be approximately linear.

3.4.9 Averaging immune infiltrate over a core

In clinical reporting, quantifying immune populations in exclusively tumour epithelium is technically challenging and time consuming. I was therefore interested in examining the average density of each marker averaged across both tumour and stromal regions from each core (Table 3.4). Averaging the tissue density of CD8⁺ increased the strength of the associated hazard ratio and significance of the model (HR=0.79, p-value=0.010) indicating increased prognostic value for CD8⁺ infiltrate over quantitation of individual epithelial and stromal infiltrates. The significance did not increase for CD45RO⁺ and CD68⁺ infiltrates. Figure 3.16 shows illustrative Kaplan-Meier survival curves for high and low CD68⁺, CD45RO⁺ and CD8⁺ densities over combined epithelium and stroma compartments.

	Functional Form	Evaluable cases	Tissue compartment	Univariable		Multivariable* (adjusted for stage)	
				HR	p-value	HR	p-value
CD8 ⁺	log10	301	Epithelium	0.89	0.15	0.83	0.027
	log10	202	Stroma	0.97	0.74	0.93	0.40
	log10	315	Average	0.79	0.010	0.72	0.0006
CD45RO ⁺	log10	290	Epithelium	0.86	0.033	0.85	0.022
	log10	196	Stroma	0.76	0.001	0.76	0.0007
	log10	306	Average	0.82	0.006	0.80	0.003
CD68 ⁺	linear	293	Epithelium	0.99	0.67	0.99	0.43
	log10	226	Stroma	0.53	0.003	0.44	0.0003
	log10	308	Average	0.67	0.042	0.62	0.017
Stage1	-	312	Localised	1	0	1	0
			Regional	1.47	0.26	1.15	0.25
			Distant	3.96	«0.001	5.58	«0.001
			Unstaged	3.35	<0.001	3.34	«0.001

Table 3.4 Cox proportional hazards survival analysis for individual infiltrates. Multivariable analysis includes infiltrate and stage. The chosen functional form for modelling and the number of patients for whom this survival can be modelled are given. The infiltrate is assessed in each tissue compartment, epithelial, stromal or averaged over the whole core.

	Cases	Univariable		Multivariable*	
		HR	p-value	HR	p-value
CD8 ⁺	111	0.84	0.27	0.7	0.047
CD45RO ⁺	110	0.98	0.89	0.96	0.78
CD68 ⁺	80	1.21	0.47	1.27	0.39

Table 3.5 Cox proportional hazard regression for cores with malignant epithelial tissue only. Multivariable analysis includes stage.

3.4.10 Epithelial core subset survival analysis

As mentioned, the sampling of stroma and epithelium varied across samples. In order to elucidate whether there were further effects of tumour/core composition I reduced the set to purely epithelial cores and I asked whether CD45RO⁺ and CD8⁺ epithelial infiltrate were still significant predictors in this predominantly epithelial environment. In the separate subset of cores with <1% stroma, epithelial CD8⁺ malignant epithelial infiltrate remained an independent prognostic factor but epithelial CD45RO⁺ density was not significant (n=110, p=0.78) (Table 3.5).

3.4.11 Combined model

Given that individually, all infiltrates were prognostic in at least one compartment, the next step I took was to model the survival of patients based on all infiltrates together in an attempt to account for the potentially different impacts of each infiltrate. I carried out multivariable Cox regression analysis including all infiltrates and stage on patients with complete data for all infiltrates (n=152) (Table 3.6). In this model only stage and CD68⁺ stromal infiltrate were significant predictors of survival.

I then refined the model by removing the least significant elements (defined as those with p>0.1) (Table 3.6, n=152). Interestingly, I found that the p-values for CD68⁺ and CD45RO⁺ stromal infiltrates in the refined model become less significant and the hazard ratios are attenuated in comparison to both the univariable regression and the full combined model. This is likely due to the inability of Cox regression to distinguish with confidence whether stromal CD45RO⁺ or CD68⁺ density is the most significant predictor when there is a strong correlation between all the immune variables.

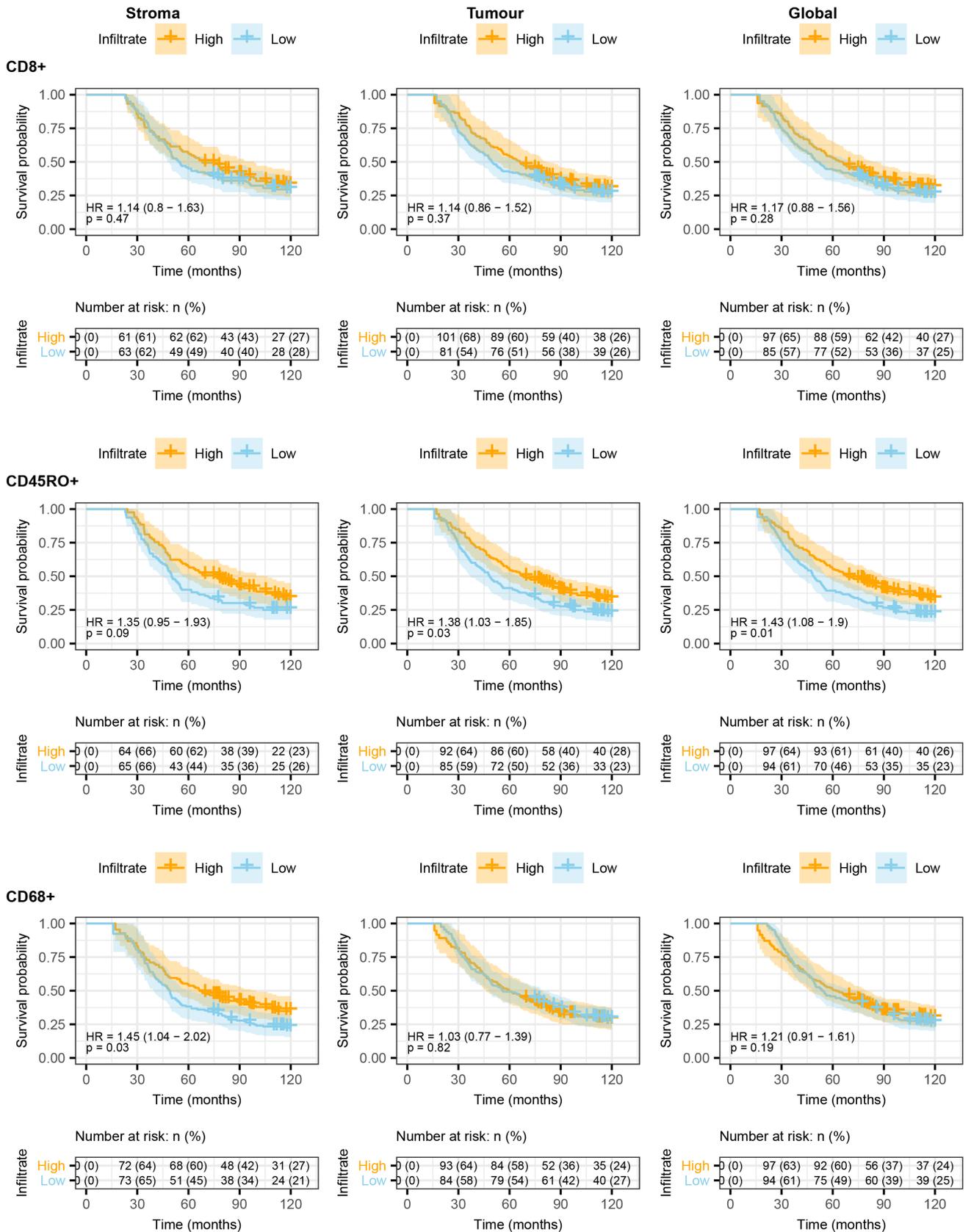


Fig. 3.16 Kaplan–Meier survival curves using cut point of median density of stromal CD68⁺ macrophages and CD45RO⁺ memory T cells using left truncation and right censoring. Median entry to the study for all patients after diagnosis was 26.4 months. Median follow up time from diagnosis to exit or death was 105.1 months.

		Multivariable (all combined)		Refined model	
		HR	p-value	HR	p-value
CD8 ⁺	Epithelium	0.96	0.81	-	-
	Stroma	1.07	0.63	-	-
CD45RO ⁺	Epithelium	1.12	0.37	-	-
	Stroma	0.83	0.09	0.68	0.11
CD68 ⁺	Epithelium	1.16	0.63	-	-
	Stroma	0.53	0.038	0.88	0.17
Stage	Localised	1	0	1	0
	Regional	2.00	0.16	2.03	0.14
	Distant	4.82	«0.001	4.70	0.0001
	Unstaged	8.15	«0.001	8.25	0.0001

Table 3.6 Cox proportional hazard regression HR for combined and reduced models. The Refined model is reduced down from all infiltrates to attempt to improve the model. P-values associated with selected predictors increases. Stage is included in both models.

3.4.12 Is the immune exclusion ratio of intra-epithelial infiltrate to stromal infiltrate prognostic?

The immune exclusion ratio or the ratio of tumour to stromal infiltrate is a quantitative measure of the immune phenotypes mentioned earlier and may indirectly measure properties of the ECM that allow for infiltration into the tumour core or measure the extent of the interaction of the immune cells with epithelial cells. As shown previously, stromal and intra-epithelial immune cell densities are correlated. I investigated whether the ratio of epithelial infiltrate density to stromal immune infiltrate density was a significant predictor of risk in HGSOc using the Cox proportional hazards model.

	n	Univariable		Multivariable	
		HR	p-value	HR	p-value
CD8	188	0.94	0.65	0.95	0.69
CD45RO	180	1.13	0.36	1.28	0.20
CD68	211	1.53	0.09	1.54	0.018

Table 3.7 Univariable and Multivariable (adjusted for stage) Cox proportional hazard tests with ratios of epithelial to stromal infiltrate as predictors. Epithelial:Stromal exclusion ratio is predictive for CD68⁺ alone.

There is no significant effect on survival of epithelial:stromal ratio for the T-cell subsets. The ratio of epithelial:stromal macrophages however has an impact upon survival, with a

higher epithelial:stromal infiltration ratio having a negative impact upon survival (Table 3.7, $p = 0.018$, HR = 1.54).

3.4.13 Principal components and the combined immunospace

A problem that I encountered in analysing the effect of multiple immune populations upon survival was that the correlation of multiple immune infiltrates hampered modelling (see Figure 3.12, Table 3.6) and conceptually the combination of infiltrates may have additive or suppressive effects. In order to assess the multi-dimensional nature of the immune response in more detail I proposed that as the three types of immune infiltrate vary continuously across epithelium and stroma these variables can be regarded as six dimensions of an ‘immunospace’ (three infiltrates, 2 localizations). Given the strong correlations between infiltrates (Figure 3.12), we can safely say these immune variables are not independent. I used principal component analysis (PCA) to determine the independent patterns across these dimensions, using the 152 patients for whom complete data were available.

PCA transformed the six correlated infiltrate variables into six independent axes with the first component containing the largest proportion of variance (60%) in the data set^{3.8}. Patients are plotted by their PC1 and PC2 values in Figure 3.17. In Principal Component 1 (PC1), the weightings of all immune infiltrates are positive and similar in magnitude (Table 3.8). This indicates that as one infiltrate increases so do all the others and represents the degree of coordination and extent of the immune response. The remaining principal components characterize additional patterns across immune infiltrates independent of PC1. The additive contribution of PC2 characterises negative correlation between CD8⁺ infiltrates and CD68⁺ macrophages and CD45RO⁺ memory cells. PC3 characterizes additional variation where epithelial and stromal infiltrates are negatively correlated, the most positive values of PC3 correspond to high infiltration in tumour epithelium compared to stroma and the most negative values of PC3 correspond to the opposite, the aforementioned immune exclusion.

Having calculated the principal components from the data which help visualise and understand variation, it is possible to extract the patients who lie at the most extreme ends of these axes of variation. Figure 3.19 shows representative images with the largest magnitudes of PC1, PC2 and PC3 to visually illustrate the patterns described above. This figure demonstrates especially clearly the coordinated immune infiltration and the immune exclusion that are described by PC1 and PC3 respectively.

The variance in the remaining principal components (4-6) is smaller and less informative. Variance in PC4 is predominantly from CD45RO⁺ stromal density, PC5 is from CD45RO⁺ epithelial density and PC6 is from CD68⁺ epithelial density.

	PC1 (60%)	PC2 (13%)	PC3 (9%)	PC4 (8%)	PC5 (5.5%)	PC6 (4.5%)
CD8 ⁺ tumour density	0.38	-0.56	0.58	-0.1	0.4	-0.21
CD8 ⁺ stromal density	0.38	-0.59	-0.47	0.27	-0.21	0.43
CD68 ⁺ tumour density	0.41	0.47	0.37	0.11	0.09	0.67
CD68 ⁺ stromal density	0.42	0.30	-0.29	0.56	0.36	-0.45
CD45RO ⁺ tumour density	0.45	0.11	0.21	-0.07	-0.78	-0.36
CD45RO ⁺ stromal density	0.41	0.15	-0.42	-0.77	0.22	-0.03

Table 3.8 Contributions of normalized infiltrates to all principal components. Figures in brackets indicate proportion of total variance

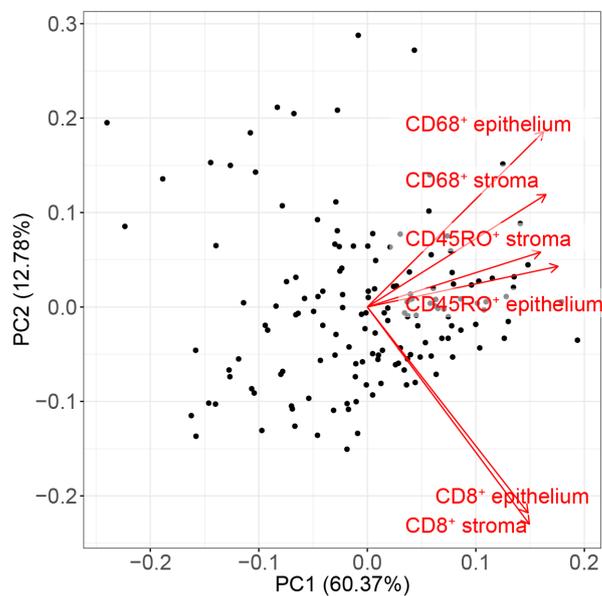


Fig. 3.17 Patients plotted on axes giving their values of Principal Component 1 and 2. A continuous distribution of patients over the principal components is shown.

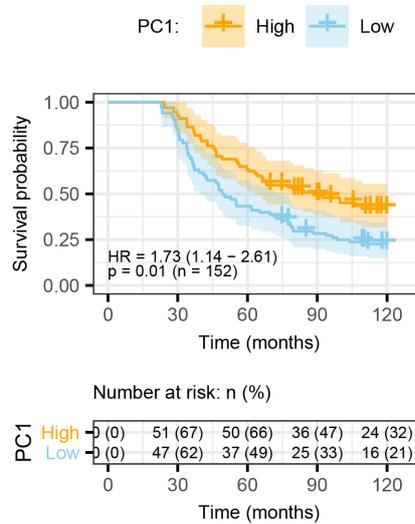


Fig. 3.18 Kaplan Meier illustrative survival curve. Values of Principal Component 1 are split at the median. The survival of high and low values of Principal Component 1 are plotted against time.

In order to evaluate the prognostic impact of these patterns of infiltration, I used Cox proportional hazard regression to assess whether these principal components were predictive of survival. Cox regression survival models were also calculated on all combinations of principal components and stage. Only PC1 was an independent predictor of survival in this cohort and was associated with improved survival (Univariate: HR=0.89, $p = 0.024$, PC1+Stage: HR: 0.88, $p = 0.016$) (Table 3.9) reflecting the good prognosis of a strong coordinated immune response. The association of this principal component with survival is also illustrated graphically using Kaplan Meier curves in Fig. 3.18.

	Univariable		Multivariable	
	HR	p-value	HR	p-value
PC1	0.89	0.024	0.88	0.016
PC2	0.94	0.61	0.92	0.52
PC3	1.2	0.2	1.23	0.18
PC4	1.14	0.43	1.07	0.69
PC5	0.82	0.31	0.74	0.11
PC6	1.25	0.23	1.22	0.33

Table 3.9 Cox proportional hazard regression for principal components as predictors. Multi-variable analysis includes stage. Only PC1 is a significant predictor of survival across both uni and multivariable models.

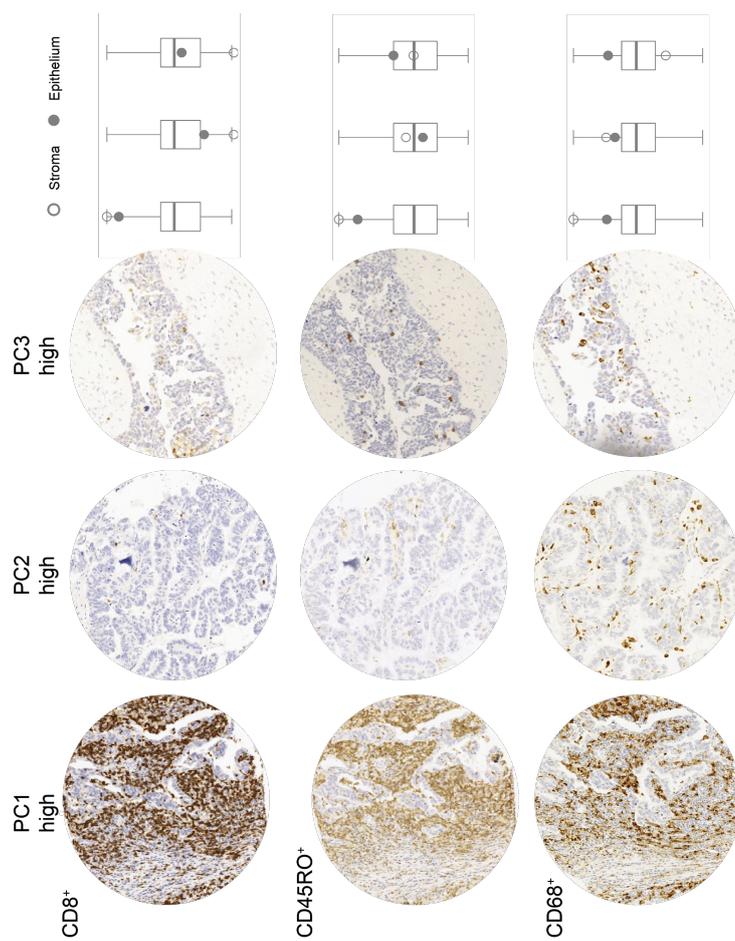


Fig. 3.19 Images of the sections of patients with the highest values of principal components 1, 2 and 3. High PC1 shows high infiltration in all regions, High PC2 shows high CD45RO⁺ and CD68⁺ infiltration when compared with CD8⁺ infiltration. PC3 shows high intra-epithelial infiltration compared with stroma. The values of the immune infiltrates from these images are shown on the boxplots demonstrating their position in the distribution across all samples.

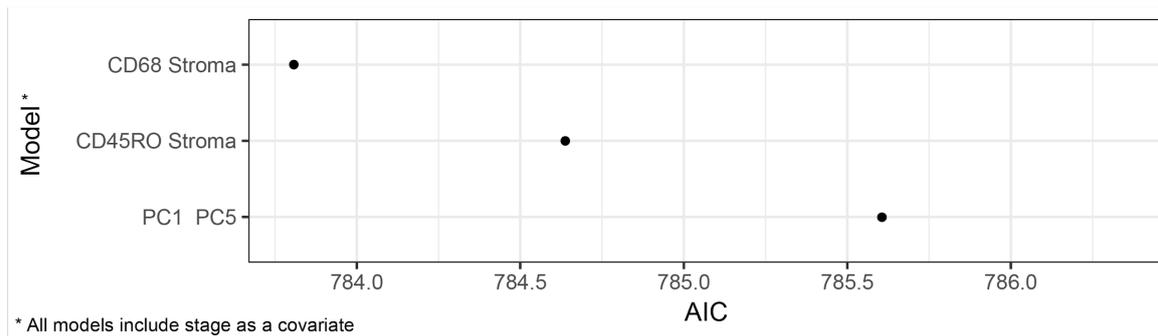


Fig. 3.20 AIC for most significant models. The best fitting model is the CD68⁺ stromal model and has the minimum AIC value. The subset of models whose fit is considered statistically undiscernable from the best fit model have a $\Delta AIC < 2$ away from the minimum. The models within this confidence range were CD45RO⁺ Stroma and PC1+PC5 and were included in the plot.

3.4.14 Comparing Survival Models

It is possible to assess all combinations of survival models including the combination of Principal components. This was carried out and the Akaike Information Criterion (AIC) was used to compare the performance of these survival models and includes a penalty on the number of terms to reduce over-fitting. The AIC values are shown in Figure 3.20. The model combining stage, PC1 and PC5 had the best performance for predicting overall survival. The improvement with the addition of PC5 shows that the addition of this principal component has a suppressor effect in the model, increasing the significance of other variables when included. This demonstrates that survival is predominantly determined by the coordinated immune response and further variation in survival from this trend can be encoded by variation in the epithelial CD45RO⁺ infiltrate from the general trend (higher CD45RO⁺ than expected when CD8⁺ is low and vice versa). Interestingly, the models that contained stage and either stromal CD45RO⁺ or CD68⁺ infiltrate contained a similar amount of information about patient survival as the one that contained stage and principal components 1 and 5. In our cohort, the density of CD68⁺ and CD45RO⁺ stromal infiltrates are therefore the best single infiltrates for survival modelling.

3.4.15 Are genetic defects associated with HGSOC driving individual infiltrates or the coordinated immune response in tumours?

Infiltrate density is not significantly associated with *BRCA* status

To investigate the interactions between the TME and the genome, one can ask whether a germline mutation in either of the *BRCA1/2* genes results in changes in the TME and specifically the immune infiltrate. I investigated whether tumour, stroma or full core measures of CD8⁺, CD45RO⁺ or CD68⁺ densities were related to *BRCA1/2* mutations using the Kruskal-Wallis rank sum test. I compared the infiltrate density distributions between patients with no mutation, a mutation in *BRCA1* and a mutation in *BRCA2*. I found that the distributions of CD8⁺, CD45RO⁺ and CD68⁺ infiltrates and the principal components were not significantly associated with *BRCA* mutation status (Figure 3.21). This may seem unexpected as previous results have found increases in CD8⁺ infiltrate with a *BRCA1* mutation[159] but the study referenced had data on both germline and somatic *BRCA* mutations and also analysed methylation status.

Infiltrate density is not significantly associated with type of *TP53* mutation.

Mutant p53 proteins have been linked to multiple micro-environmental changes[160]. I investigated whether the type of *TP53* mutation produced measurable changes in the TME. Tumour, stroma or full core measures of CD8⁺, CD45RO⁺ or CD68⁺ densities were compared between gain or loss of function mutations in *TP53*. I used the Kruskal-Wallis rank sum test and found that the distribution of immune infiltrate was not significantly associated with type of *TP53* mutation.

Infiltrate density is not significantly associated with change in PTEN expression.

Changes in PTEN expression are cell-intrinsic and levels of PTEN expression are prognostic in HGSOC[87, 161]. Tumour, stroma or full core measures of CD8⁺, CD45RO⁺ or CD68⁺ densities were compared between patients with high or low PTEN expression as measured by both IHC and IF. Using the Kruskal-Wallis rank sum test I found that change in PTEN expression was not significantly associated with any change in the density of immune infiltrate. This particular cell-intrinsic change is not associated with a change in immune infiltration in our cohort.

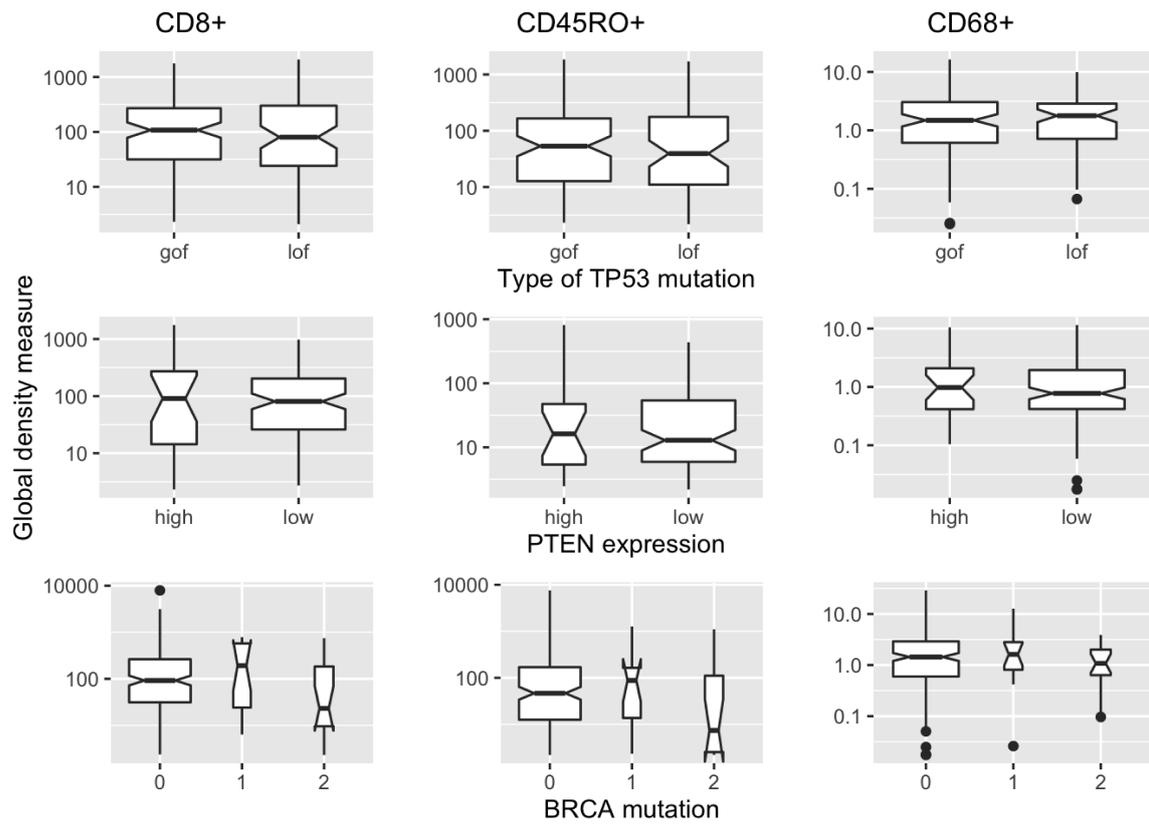


Fig. 3.21 Boxplots of infiltrate distribution against genomic phenotype. Overlapping notches show that no significant differences are observed between the distributions of these populations. (Notches on box plots extend $1.58 \times \text{IQR} / \sqrt{n}$ and approximate the 95% confidence interval for the median. Box plot whiskers extend to $1.5 \times \text{IQR}$.)

		<i>BRCA1/BRCA2</i>	<i>TP53</i> GOF/LOF	PTEN
CD8 ⁺	Epithelium	0.04	0.54	0.67
	Stroma	0.17	0.29	0.07
	Average	0.17	0.53	0.84
CD45RO ⁺	Epithelium	0.2	0.42	0.92
	Stroma	0.46	0.84	0.82
	Average	0.1	0.36	0.87
CD68 ⁺	Epithelium	0.43	0.72	0.79
	Stroma	0.78	0.85	0.65
	Average	0.68	0.66	0.41
Principal Component	1	0.81	0.8	0.55
	2	0.1	0.26	0.17
	3	0.27	0.58	0.51
	4	0.58	0.99	0.17
	5	0.76	0.88	0.79
	6	0.97	0.44	0.18

Table 3.10 P-values associated with Kruskal Wallis test for detecting differences in mean ranks of immune infiltrate in patients grouped by mutation in *BRCA1*, *BRCA2* or not-detected, *TP53* GOF or LOF and PTEN high or low in HGSOc.

3.4.16 Other ovarian cancer subtypes

The automated and reproducible nature of this workflow means similar analyses are easily transposed across to the same immune populations assessed in the other subtypes of Ovarian cancer. Figure 3.22 shows the number of patients with each OC subtype in the SEARCH cohort. Due to the smaller numbers of patients with these subtypes, even before incomplete data removal, I limited this secondary analysis to just examine the stromal and epithelial areas and comparing correlations between infiltrates. Figure 3.23 shows the areas of epithelium and stroma across all OC subtypes. Stromal area sampled is significantly lower than epithelial area in all subtypes except Low Grade Serous Ovarian Cancer (LGSC), the LGSC group is likely too small to see significant results. Endometrioid samples have a significantly lower median stromal area than HGSOc.

Figure 3.24 shows the distribution of CD8⁺, CD45RO⁺, CD68⁺ cells in the stroma and epithelium across the different OC subtypes. CD8⁺ cell infiltration is significantly higher in HGSOc than all other subtypes across both epithelium and stroma. CD45RO⁺ infiltration in HGSOc stroma and epithelium is significantly higher than clear cell. The three infiltrates show similar relative patterns across the subtypes.

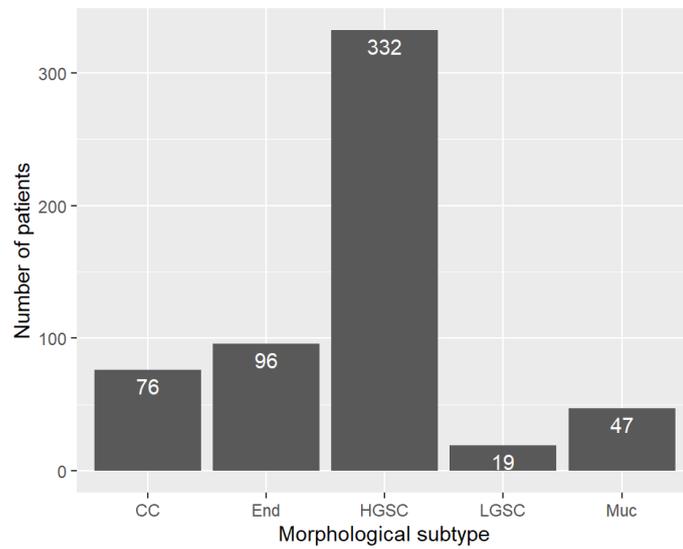


Fig. 3.22 Histogram showing number of patients with each morphological subtype of OC in the SEARCH cohort. HGSC represents 58.2% of the cases in this cohort.

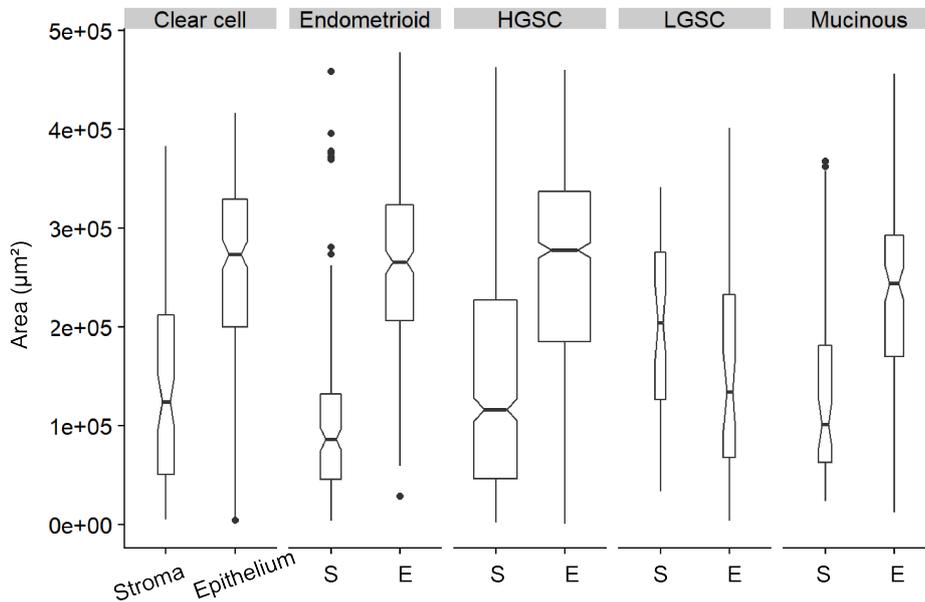


Fig. 3.23 Area of epithelium and stroma across morphological subtypes. Stromal area is significantly lower than epithelial in all but the LGSC subtype. In LGSC the confidence intervals overlap, likely due to the small number of cases. (Notches on box plots extend $1.58 \times \text{IQR} / \sqrt{n}$ and approximate the 95% confidence interval for the median. Box plot whiskers extend to $1.5 \times \text{IQR}$.)

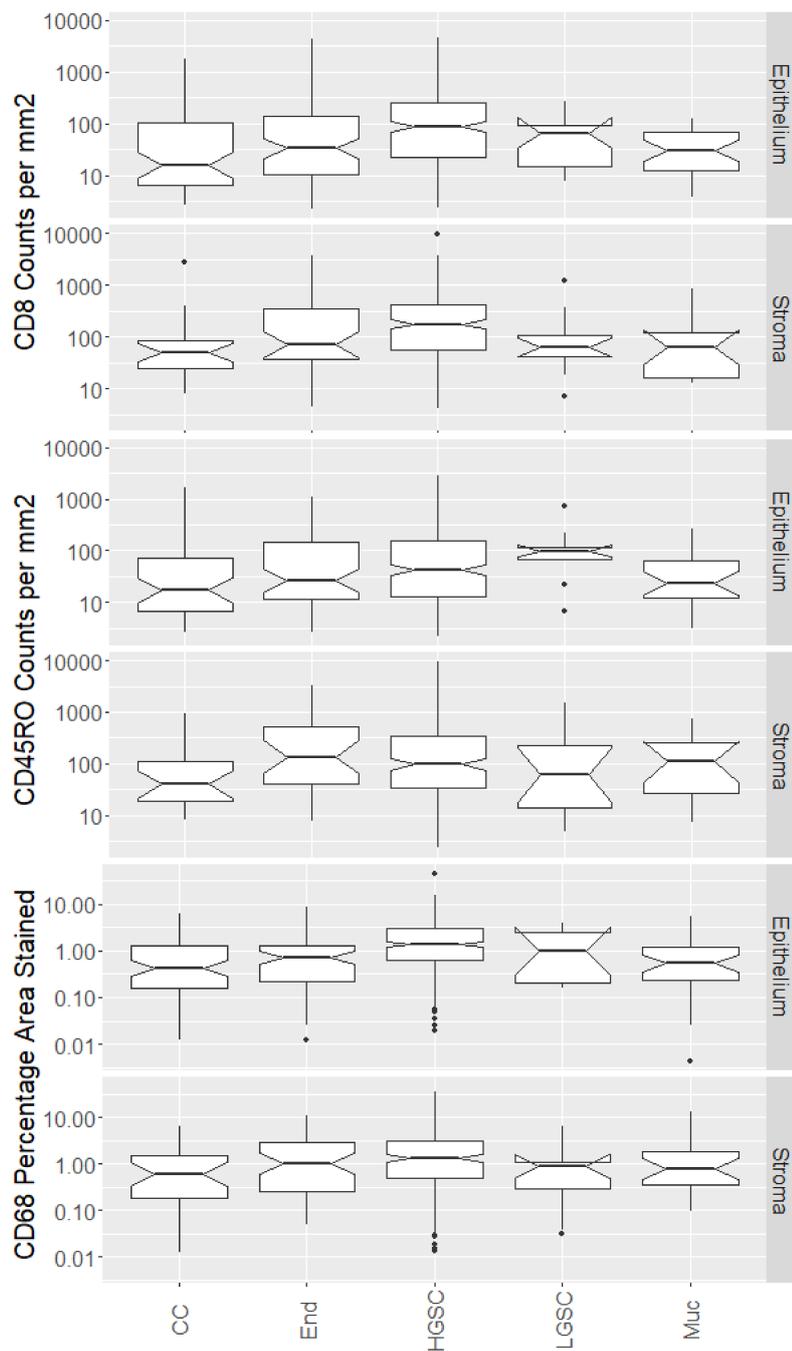


Fig. 3.24 Distribution of $CD8^+$, $CD45RO^+$ and $CD68^+$ in stroma and epithelium for different morphological subtypes of ovarian cancer. Log base 10 scale is used for counts. Endometrioid and HGSC/HGSC have significantly higher $CD45RO$ and $CD68$ infiltration than Clear Cell(CC). Notches on box plots extend $1.58 \times IQR / \sqrt{n}$ and approximate the 95% confidence interval for the median. Box plot whiskers extend to $1.5 \times IQR$.

3.5 Discussion

This reproducible analysis of quantitative immune data allowed for a rigorous examination of the distributions of immune infiltrates within the SEARCH population and of the potential impacts of localisation of CD8⁺, CD68⁺ and CD45RO⁺ in HGSOE tumours.

I observed that stromal and epithelial populations and all infiltrates were somewhat positively correlated. This means that samples with low density of stromal immune populations generally had low density of epithelial infiltrate and vice versa. It also meant that patients with high quantities of one infiltrate type often had high infiltration of the others. I observed that CD8⁺, CD45RO⁺ and CD68⁺ infiltrates predominantly exist on a continuum without clear justification for cutpoints. The implication of this is that morphological subtypes based on immune infiltration have no clear cutpoint for classification and gene expression signatures will just distinguish higher infiltration tumours from lower ones.

It is worth noting that the CD8⁺ and CD45RO⁺ subpopulations are not biologically mutually exclusive and as such, we expect to see some degree of correlation. The analysis was, however, carried out on serial sections and as such the analysis is still on distinct cell populations. A weakness of this data set for analysis was that with single slide, single marker IHC staining there was no possibility of analysing subpopulations of CD8⁺CD45RO⁺ immune cells. Another weakness is that only single cores were taken for each patient; using replicates through multiple sampling of the tissue block and across areas with different epithelial and stromal composition would allow the differing impact of the infiltrates in different tissue contexts to be explored.

One of the key and novel findings of this work was that the global measure of CD8⁺ infiltrate was a stronger and more significant predictor of survival over CD8⁺ infiltrate alone. The implication of the average infiltration being a stronger predictor of survival is that averaging infiltration over the intra-tumour stroma as well as the epithelium allows for a better measure of survival. This is possible as the infiltration in the margins close to the epithelium reflect long term infiltration dynamics and measures the presence of a reservoir of immune cells, poised to infiltrate the epithelium. I also demonstrated here that increasing quantity of epithelial and stromal CD45RO⁺ is a also positive prognostic feature, a result which also confirms the positive incremental benefit of a long term and mature immune response.

As discussed in the introduction, many pieces of research have illustrated that macrophage function is micro-environment dependent and highly variable[162, 163, 42, 164], demonstrating that localisation effects macrophage function. The predominant literature associates macrophages with poor prognosis but Tanaka et. al showed that the expression of Bikunin, a protease expressed by CD68⁺ macrophages, was associated with improved survival[44].

My work is also in agreement with the observation by Li et al. that stromal regions are most informative in which to evaluate macrophages. I find however, in contrast to their work in lung cancer, that stromal CD68⁺ macrophages are a positive prognostic feature. Their work is however, also further evidence for distinct phenotypes of macrophages between tissue regions. Interestingly, if CD68⁺ macrophages in the epithelial region are predominantly also CD163⁺, then the negative prognostic epithelial to stromal ratio of CD68⁺ cells I derived would loosely correlate with and be in agreement with the CD163⁺ ratio discussed by Yafei et al. [43].

The impact of specific immune cells can vary between cancer types and so further elucidation of the nature of macrophages in HGSOE and their interaction with the TME is required.

I investigated the concept of immune exclusion across this quantitative data set and found that the ratio of infiltrates and thereby the extent of exclusion, was normally distributed. Such a distribution implies that there is not a very distinct subpopulation of samples with exclusion. I found that this ratio of epithelial:stromal infiltrate was prognostic for macrophages. Increased epithelial macrophage:stroma macrophage ratio had a negative impact upon survival. Prior to this work I would have expected the quantity of T-cell infiltration from the stromal to intra-epithelial regions to affect tumour progression. This data however suggests, as the ratio is not prognostic and as the average density of the T-cell subsets over the core is prognostic, that there is no explicit long term exclusion of cells. The implication is that infiltration at the tumour-stroma interface is dynamic and random and that the overall quantity in the larger region is reflective of impact. Specific immune exclusion, if it is an active process, was not observed to any extreme in this cohort[23].

Using quantitative data also allowed me to observe that there was no clear threshold for immune exclusion. A single standard deviation was equivalent to a 10 fold difference in immune infiltration; this threshold was used for comparisons. This, along with the observation that average infiltration is higher in stroma than tumour, is important when considering that in the literature tumours have been classified into immune-hot, immune-cold and immune-excluded groups without reasonable justification[99, 165, 166], and unsupervised clustering does not identify such groups[102]. In fact, given the log-normal distribution it is likely that attempting to classify the immune content of patients into three groups merely results in identifying the first third, middle and last third of log-normally distributed counts. Other work attempting to compare immune infiltrates quantitatively have sometimes forgone the analysis of the distribution and normalisation.

The focus of many studies in immune oncology is the intra-epithelial immune infiltration; likely due to evidence of direct contact and the cytotoxic interaction between immune and

tumour cells. The results in this work suggest that stromal infiltrate should also be evaluated. Even disregarding the observed survival modelling benefit, it is statistically beneficial to evaluate a larger area of tissue. Furthermore, if tissue segmentation is not required, one can obtain more accurate infiltrate counts more easily. For macrophages it appears that stromal infiltration plays a distinct role and phenotype as compared to epithelial macrophages in the immune response. In the case of the T-cell subsets evaluated here, the addition of the stromal region also appears to provide more information about the epithelial infiltration as we can infer some element of the longer term immune infiltration dynamics.

In this chapter I also developed the concept of the immunospace. The immunospace views the quantity of each immune infiltration as part of a multidimensional immune landscape for a patient. This approach allowed me to analyse patterns across different immune populations and to derive the principal components as measures of variation. I observed that the correlation between infiltrates amounts to general coordinated immune response which is a positive prognostic. It also allowed me to examine the independence of other phenomena such as the immune exclusion from the epithelium which was obtained as an axis of variation in PC3 in an unbiased manner. Such high dimensional methods will be of importance when assessing higher dimensional data sets such as IMC.

Having reduced the survival analysis to cores that contained epithelium and stroma in order to compare survival impact, the limitation of this analysis is that it assesses the impact of these immune cells within a stromally infiltrated environment. In order to address this I compared the survival impact of CD8 and CD45RO⁺ in the subset of cores with >99% epithelium. CD45RO⁺ was no longer associated with survival which may imply a variation in the functionality of these cells which depends upon their spatial location (within tumour nest).

The presence of germline *BRCA2* mutations was significantly associated with lower CD8⁺ cell density than patients with a germline *BRCA1* mutation after multiple testing correction (Figure 3.21). There was no significant association between the quantity of CD45RO⁺ or CD68⁺ infiltrate and the mutational status of either *BRCA1* or *BRCA2* genes (Figure 3.21). No significant association was detected between *TP53* GOF and LOF mutations or PTEN expression and the densities of CD8⁺, CD45RO⁺ and CD68⁺ cells in epithelium or stroma (Table 3.10). Similarly, changes in the principal components were not significantly associated with PTEN expression, *TP53* GOF or LOF mutation or germline *BRCA1/BRCA2* mutation status (Table 3.10). The lack of a result with respect to *BRCA* is likely in part due to effects of cohort size. Studies that have found a positive result have sometimes had hundreds of patients with a mutation and still observe p-values of approximately 0.05 [82, 27, 159]. These studies also frequently included somatic *BRCA* mutations and also analysed methylation status,

neither of which we had information on. In this cohort, With approximately 20 patients with mutations, I am underpowered to ask multiple questions about the differences in immune infiltration between non-*BRCA* and *BRCA*-mutated patients.

I observed the differences in tissue sampling between ovarian cancer subtypes. The observation that the distribution of tissue area sampling remains constant either demonstrates that the distribution of epithelium and stroma across patients in all populations is similar or that the sampling by the pathologist is consistent in extracting quantities of epithelium. Given that the range of the epithelial tissue percentage is high I would propose that the distributions are similar across the population. I also examined the infiltration density between subtypes of ovarian cancer. HGSOC has significantly higher CD8⁺ infiltration than other subtypes and endometrioid OC was observed to have higher infiltration than clear cell. This is in line with other publications which identify HGSOC as the subtype with the highest immune infiltration[167]. This data also shows that this analysis can easily be extended to other ovarian subtypes.

Having demonstrated the importance of spatial location and the inclusion of neighbouring stroma in terms of the impact of infiltration, questions are raised about the exact spatial nature of the immune infiltration. It is also evident that in current approaches there is no accounting for tumour structure as a factor in the infiltration. The differing quantities of infiltration between epithelial and stromal regions also implies that averaged infiltration quantities could be an indirect readout of the structure of the tumour. In addition, the difference in prognostic impact in stroma infiltrated tumours further suggests a role for tumour structure in prognosis.

Chapter 4

Quantifying and classifying tissue structure

4.1 Introduction

Having demonstrated in Chapter 3 that the specific stromal or epithelial localisation of several immune infiltrates was an important factor in the prognosis of patients, I wanted to investigate in more detail the structure of epithelium in tissue sections and assess the relationship of the epithelial and overall structure with immune infiltrate.

In order to demonstrate the accuracy of the structure classification and its wider applicability, I opted to use both H&E and pan-cytokeratin (CK) stained images to derive this structure. Pan-cytokeratin (a standardised mix of cytokeratin markers) was used as the epithelial marker and provided a gold standard for epithelial cell identification for structure analysis. This allowed me to validate tissue classification of H&E such that structures can be compared across multiple sections from the same core.

It is clear from observing images of tumour sections that the structure of the epithelium and stromal regions varies dramatically. Examples of these morphological differences have been discussed by Lisio et al. [13] amongst others. The morphologies discussed by Lisio et al. are solid architecture, glandular architecture with slit-like spaces, papillary architecture and cribriform and pseudoendometroid architecture[13]. The automated distinction of these has not been carried out but I aimed to at least initially build up a series of tissue metrics for which segmentation of epithelial and stroma cells must first be acquired.

This section of the thesis aims to set out an automated classification of epithelial and stromal cells, to derive metrics for tissue structure and to investigate whether these metrics are conserved across a tumour and across tumour sites in a patient. I also wanted to analyse

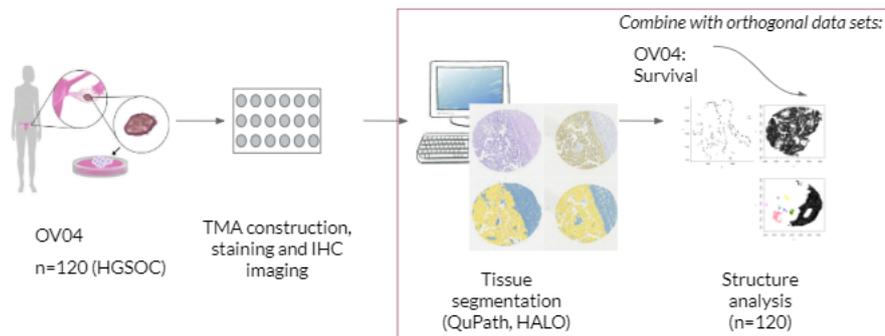


Fig. 4.1 Project outline for this Chapter. Red highlighted section was carried out by myself.

whether these metrics could be related to tissue architecture morphologies. Finally I aimed to investigate whether the structure defined this way is related to the infiltration of the tumour by particular types of immune infiltrate. The workflow of this Chapter is illustrated in Figures 4.1 and 4.2.

4.2 Methodology

4.2.1 Patient Cohorts

I utilised the OV04 Cohort for this analysis as it had existing TMA sections that had H&E and CK staining, it also had samples from multiple sites. This cohort also had unstained sections available for further analysis.

4.2.2 IHC

Pan-cytokeratin, CD8 and FOXP3 staining

Tissue sections from the OV04 cohort were stained with Pan-cytokeratin, CD8 and FOXP3 markers by Sarwah Al-Khalidi (SAK) and the method of staining and imaging is laid out in section 2.3.2.

4.2.3 H&E

Automated Haematoxylin and Eosin staining was carried out by SAK and is described in section 2.3.2.

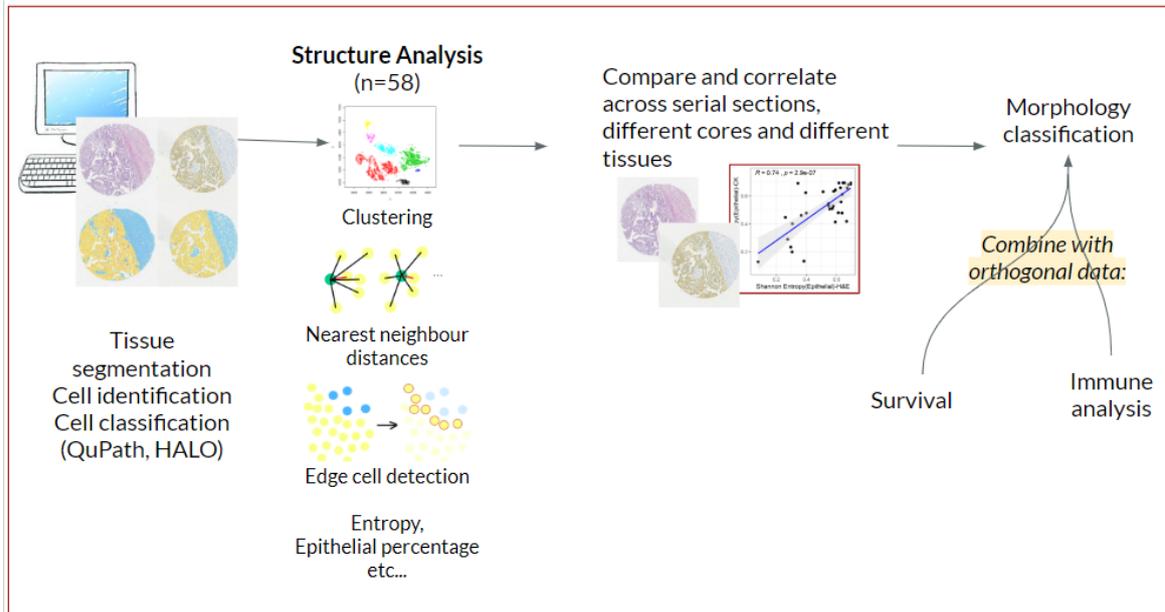


Fig. 4.2 Further details of the analysis in the project outline.

4.2.4 Statistical Analysis

The package **dbscan** in R was used for density based clustering analysis. **spatstat** was used for the creation of point patterns and estimation of K and L functions. **SpatEntropy** was used to analyse Shannon and Batty entropy and **GGally** used for correlation plots.

4.3 Results

4.3.1 Patient Characteristics

TMAAs had been previously constructed from the CTCR-OV04 clinical studies, which were designed to collect imaging, blood, and tissue samples for exploratory biomarker studies. All patients provided written, informed consent for participation in these studies and for the use of their donated tissue, blood specimens, and anonymized data for the laboratory studies carried out. The CTCR-OV04 studies were approved by the Suffolk Local Research Ethics Committee (reference 05/Q0102/160) and Cambridgeshire Research Ethics Committee (reference 08/H0306/61)[168].

Samples were collected and tissue blocks constructed where possible for each patient from normal fallopian tube, malignant ovary, omentum and non-omental metastases. Two cores were extracted per tissue block where possible and two TMA blocks were constructed

N	58
Age	77 (60-90)
Stage (FIGO)	N
1	1
2	2
3	37
4	9
Unstaged	2
Treatment Hospital	
Addenbrooke's	26
Bedford	6
Other	22
Unknown	4

Table 4.1 OV04 study patient characteristics.

from these cores per tissue type. The subset of the cohort that I analysed contained 58 patients with HGSOV. Clinical data for this cohort included FIGO stage[153].

I analysed the following subsets of samples: a single TMA slide for H&E of the Ovary, 2 matched TMA slides from different cores within the same tissue block with CK staining of the Ovary, a single TMA slide with CK staining of Omentum. In an ideal scenario with no missing data this results in a serial H&E and CK stained section from core 1 from the ovary block, a CK stained section from core 2 from the ovary block and a CK stained section from an omental metastasis block. Figure 4.3 shows the subsets and numbers of patients that have the different sample types.

4.3.2 Cell classifier performance and comparison between H&E and cytokeratin stained images

Images of serial tissue sections stained with both H&E and a pan-cytokeratin marker were analysed. Images of sections with cytokeratin staining are referred to in this document as CK images. QuPath was used to segment nuclei based upon their optical density. Cells were defined as regions at most 5 micron from and including a nuclei. These segmented nuclei and their surrounding area will be henceforth referred to as "cells" in the image. I utilised the cytokeratin staining in the CK images and examined the serial H&E sections to label epithelial and stromal regions. Over 2000 cells from each TMA slide were labelled and they were split randomly into a 50:50 ratio, 50% were used for training and the other 50% were used as a validation test set. The training subset was used to train a random forest based classifier. Cells were then classified as tumour or stroma based upon their size,

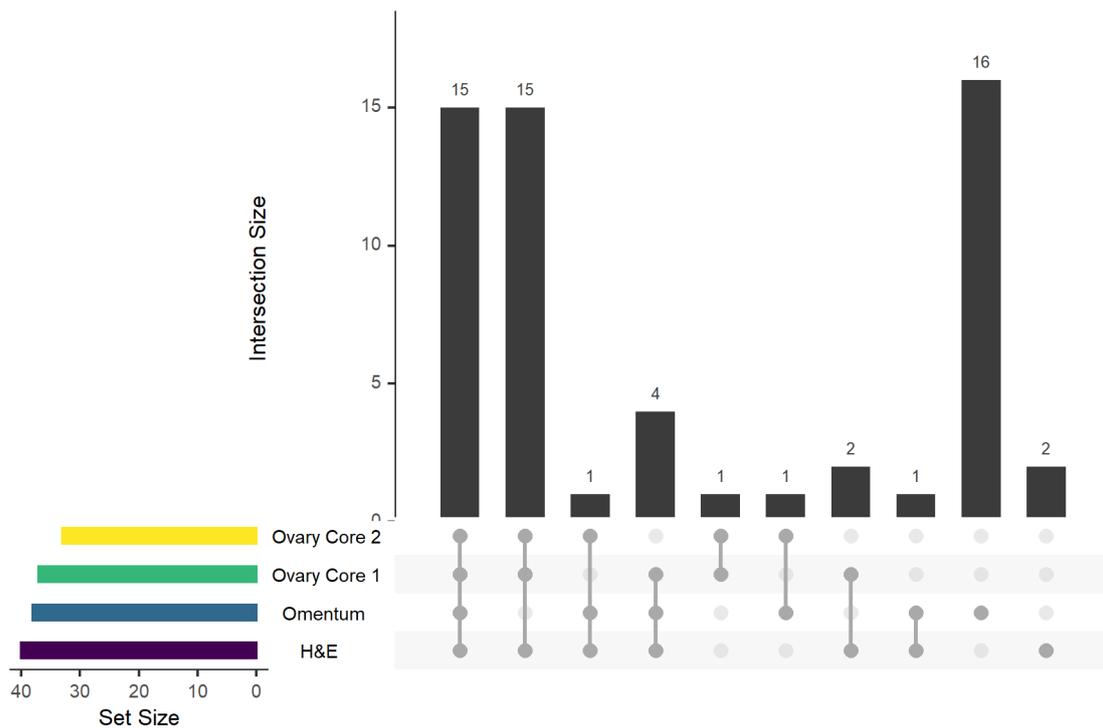


Fig. 4.3 UpSetR plot of the subsets of patients with images for each sample type. Numbers above the bar plots signify the number of patients in each subset. For example 16 patients have images of Omentum samples only, 15 patients have images from all slide types, 15 patients have samples across all 3 Ovary slides but no sample from the Omentum. H&E refers to serial sections stained with H&E of the Ovary Core 1 sample.

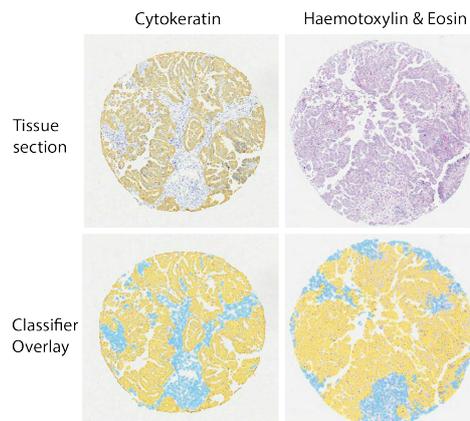


Fig. 4.4 An example of the stroma and tumour classifier built using QuPath software. Tumour is highlighted in yellow and stroma in blue. Classifier is trained on both H&E and Pan-cytokeratin(CK) stained images.

CK		Label	
		Stroma	Tumor
Classification	Stroma	1263	2
	Tumor	2	346
H&E		Label	
		Stroma	Tumor
Classification	Stroma	2653	40
	Tumor	11	2165

Table 4.2 Confusion matrix for QuPath H&E and CK based classifiers. Percentage of correctly classified objects in CK test set: 99.8% (n=1613). Percentage of correctly classified objects in HE test set: 98.95% (n=4869).

shape and smoothed features using this random forest based classifier. Examples of tissue classification on H&E and CK stained images are shown in Figure 4.4. Confusion matrices for the H&E and CK classifiers are shown in Table 4.2. These show a very good performance by both classifiers and show that such analysis could be carried across to H&E images. It is important to note that in this analysis, lymphocytes were just classified as their nearest cell type (epithelial or stromal).

The classifiers both perform well and as further validation across the image types, Figure 4.6 shows that the percentage of epithelial cells in each image are strongly correlated, as would be expected from serial sections. I observed in training that at least 1000 examples of classified cells were required in the CK based images to get an over 99% classification

accuracy across all images which naturally vary slightly in staining intensity and tissue texture.

4.3.3 Metrics for tissue structure analysis

I was interested in being able to quantify the tissue structure of a sample and in order to define a tissue structure of each core I selected several possible metrics to define, examine and measure;

- Stroma/Tumour percentage
- Cell packing (Nearest Neighbour Epithelial Distance)
- Surface Area:Volume ratio of Epithelial regions
- Shannon and Batty entropy
- Number of clusters in tissue

I aimed to derive these values and then compare these metrics between serial sections, samples from a different region of the same tissue block and between different tissues from the same patient. As cores contained a minimum of 1000 cells and these analyses require at least 25 epithelial cells for epithelial clusters to be identified, sections which were over 97.5% stromal cells were excluded from further structural analysis.

4.3.4 Stroma/Epithelium percentage

This most basic measure of tissue composition was utilized in the previous chapter. In this chapter I measured this proportion as a percentage of the total cells segmented in the image and examined whether it was conserved across intra-patient samples.

Distribution of epithelial fractions.

The distribution of epithelial cell percentages as derived from CK and H&E images are shown in Figure 4.5. Box-plots show the distribution of epithelial percentages in the OV04 cohort based upon H&E Ovary, CK stained matched tumour block cores and Omentum core. Boxplots display overlapping notches at the median demonstrating no significant difference between the distributions of epithelial percentage in the different sample types.

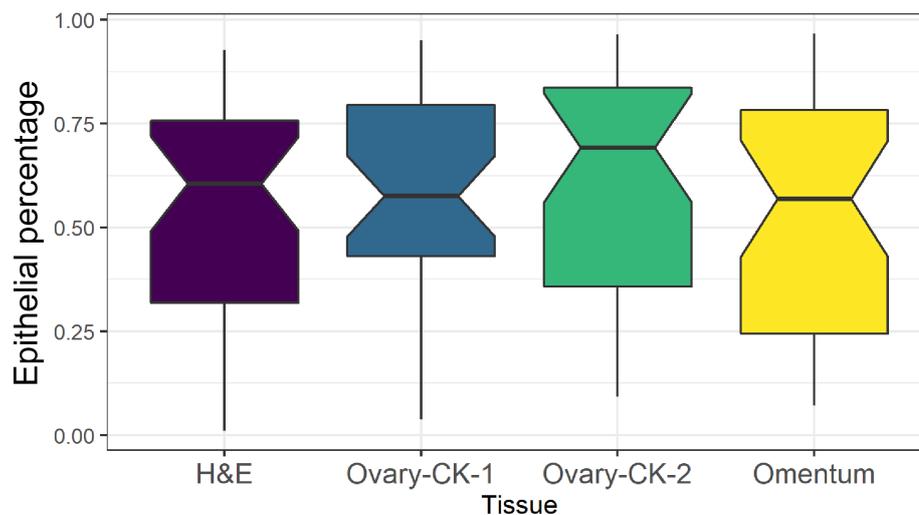


Fig. 4.5 Boxplots of the distribution of epithelial percentages in the OV04 cohort based upon H&E Ovary, CK stained matched tumour block cores and Omentum TMA. Boxplots display overlapping notches at the median demonstrating no significant difference between epithelial percentage distributions. Notches estimate the 95% confidence interval around the median. Whiskers extend to $1.5 \times \text{IQR}$

Comparison across images from different samples

In order to understand how intrinsic the percentage epithelium was to a patients tumour and whether it varied with the location of that tumour I analysed the correlation between epithelial percentage across serial sections. To do this I analysed the epithelial fraction in cores from the same tumour block and cores from different tumour sites (Ovary and Omentum). Figure 4.6 shows these correlations and the associated Pearson correlation coefficients and p-values. I found that epithelial fractions from serial sections and sections from the same tumour block were significantly correlated but epithelial fractions in matched samples from different sites were not.

4.3.5 Epithelial Nearest Neighbour Distance

Epithelial cells are not randomly dispersed through tissue and tend to cluster, ergo the average distance between nearest neighbour epithelial cells is a product of both the average epithelial cell size in the sample and a function of the organization and packing of the cells. Figure 4.7 shows the L function for one of the spatial point pattern of epithelial cells. Above the minimum distance between cells ($5\mu\text{m}$), the distance from the centre of the nuclei to the edge, the L function is above the theoretical line. This demonstrates increased clustering; cells are closer together than they would be if randomly dispersed. This pattern is to be

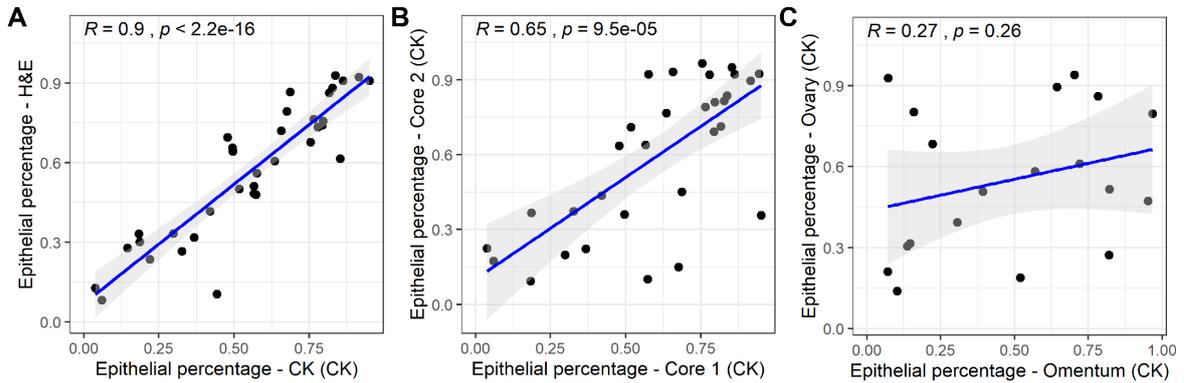


Fig. 4.6 Epithelial percentage is compared between samples. (A) Correlation between percentage epithelium measured via H&E versus CK stained serial section. (B) Correlation between percentage epithelium for different cores from the same tumour block. (C) Correlation between percentage epithelium for cores from different tumour blocks (Ovary and Omentum). All R and p-values are Pearson correlation.

expected as tissue cells clump together and is seen across all cores. (See Methods for details on Ripley K and L functions).

Comparing NNdist for different matched samples

In order to understand how intrinsic the cell packing distance of the epithelium was to a patient's tumour and whether it varied between tissue sites I analysed the correlation between epithelial NN distance across serial sections, in cores from the same tumour block and cores from different tumour sites (Ovary and Omentum). Figure 4.8 shows these correlations and the associated Pearson correlation coefficients and p-values. I found that median epithelial nearest neighbour densities from serial sections ($R = 0.65, p = 2.7 \times 10^{-5}$), sections from different cores from the same tissue block ($R = 0.58, p = 0.00064$) and sections from differing tumour sites were significantly correlated ($R = 0.78, p = 7.8 \times 10^{-5}$).

4.3.6 Quantifying border cells and the invasive front

In a purely diffusive model of immune infiltration where immune infiltrate travels from the stroma into the epithelial nests, epithelial infiltrate would be proportional to the surface area of the epithelial nodules that are exposed to the surrounding stroma.

The ratio of circumference to the area of a circle and the surface area to volume ratio of a sphere are both proportional to $\frac{1}{r}$. As such the surface area of a tumour nodule to its volume can be estimated from a 2D section as the following:

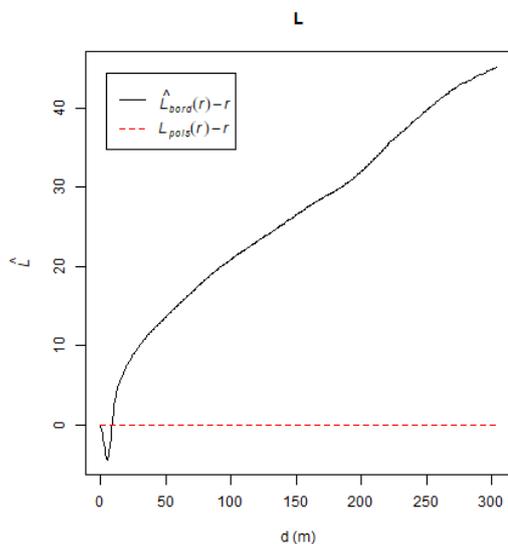


Fig. 4.7 L-function plotted against distance for a single core epithelial cell point pattern. Within the radius of the cell, epithelial cells are less clustered than expected but beyond the limit of physical proximity, cells are more clustered than at random.

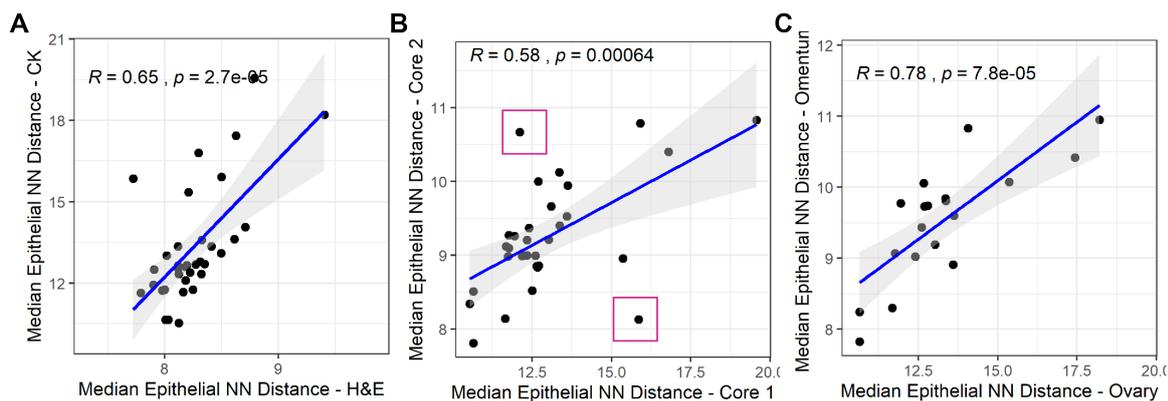


Fig. 4.8 Correlation between median Epithelial NNdist for (A) H&E versus CK stained serial sections (B) different cores from the same tumour block (C) cores from different sites and tumour blocks (Ovary and Omentum).

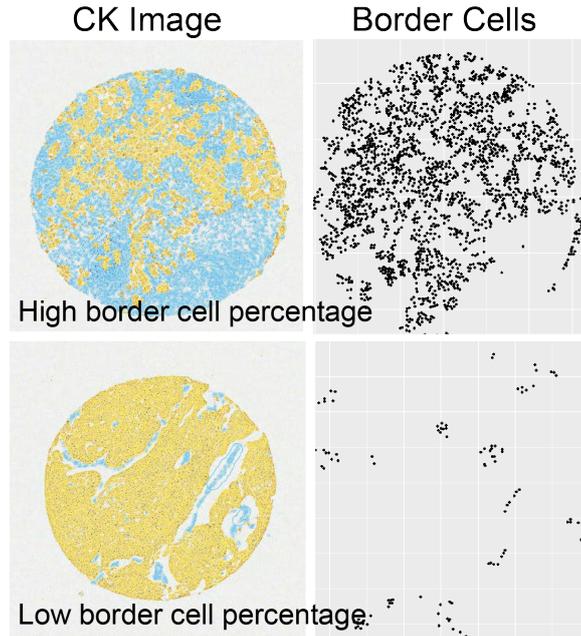


Fig. 4.9 Example of images with very high and very low border cell percentages. Image is classified as stroma(blue) and epithelium(yellow) and the locations of cells in the image designated as border cells are plotted.

$$\frac{SA}{Vol} \propto \frac{(N_{edge})}{(N_{epi})}$$

Where N_{border} is the number of cells on the border between an epithelial tissue compartment and stromal one and N_{epi} is the number of cells making up the epithelial cluster.

In order to gain an approximation of this surface area, I defined border cells as those epithelial cells which had a close stromal neighbour. Samples with no epithelial cells were excluded from this analysis.

Tumour cells are on average clustered closer to each other than stromal cells are. As such I used this domain knowledge and visual inspection to define the neighbour distance cutoff. I defined this as a stromal cell within twice the distance of the nearest epithelial cell.

Figure 4.9 shows an example of two tissue sections and the cells classified as border cells. As shown, these located cells are visually a good approximation to the edges of the epithelium in contact with stroma in the tissue section. Figure 4.10 shows the distribution of percentage of edge cells across samples.

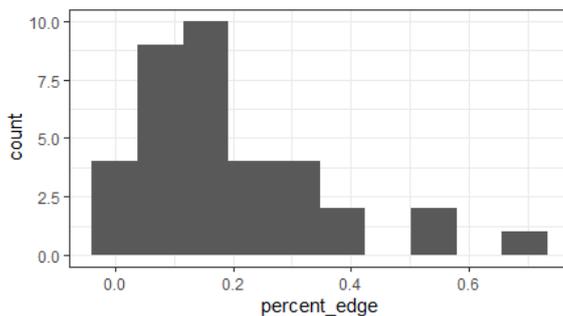


Fig. 4.10 Histogram showing the distribution of cells in an image classed as border cells as a percentage of total epithelial cells. Samples with no epithelial cells were excluded from this analysis.

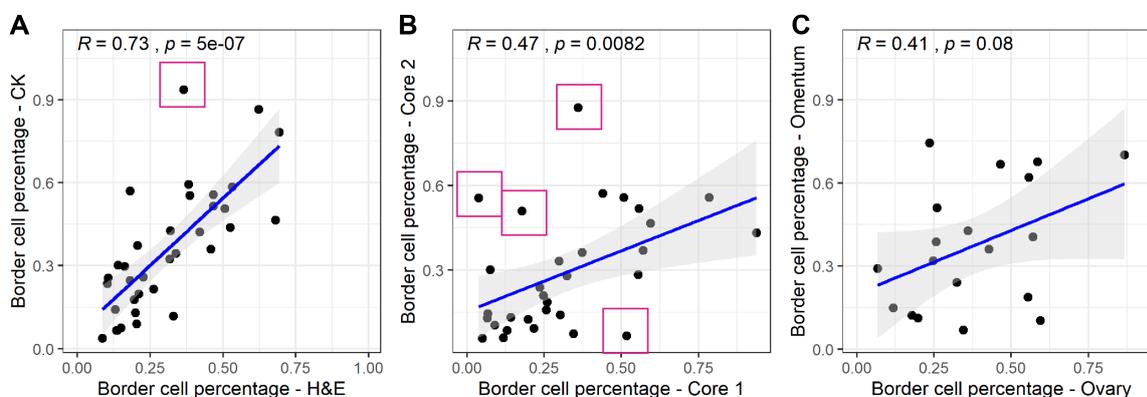


Fig. 4.11 Correlation between percentage of epithelial cells classified as border cells for (A) matched serial sections (B) matched cores from same tumour block (C) different blocks from different sites (Ovary-Omentum) from the same patient. Pearson R and p-value shown. Outliers are highlighted with purple squares as points with largest residual from linear model of the general trend.

Comparison across different samples

I used the OV04 cohort to examine the number of epithelial cells classified as border cells and compare this between cores extracted from the same tumour block. In order to assess whether this element of tissue structure is conserved across serial sections, cores from the same tumour block and different sites of tumours I compared the percent of edge cells in each. Figure 4.11 shows the scatter plots and Pearson correlations. The percentage of border cells is strongly correlated between H&E and CK on serial sections but less correlated within spatially separated cores from the same tumour block and not significantly correlated between matched samples from ovarian and omental sites.

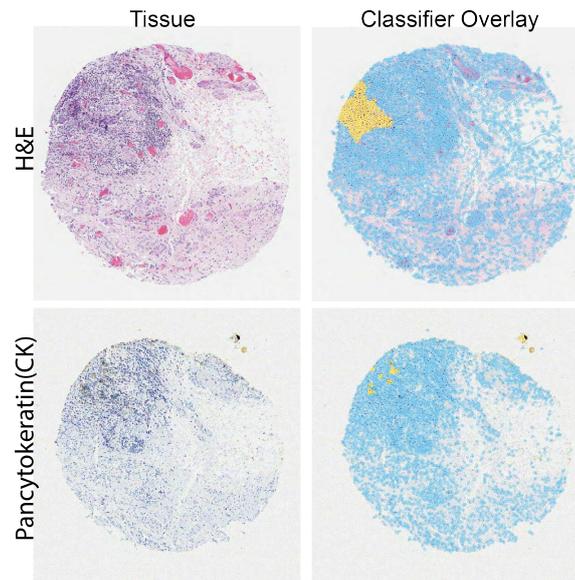


Fig. 4.12 Serial sections in H&E and CK staining with significantly different border cell percentages. Despite visually similar structure and whilst both sections have only a small epithelial region, these samples have significantly different border cell percentages. This is because the cytokeratin stained section has an epithelial area of less than 10% of the size of the one in the H&E(both labelled yellow). The region in the CK image is also split into multiple regions. This extremely small area results in a much higher percentage of total epithelial cells which are touching the adjacent stroma (blue) and hence a significant difference between the serial sections.

Outlying cores

As shown there are a few cores whose structural metric seems to vary significantly between cores from the same tissue block. As this metric is very visually understandable and in order to confirm that these data points were accurate, I wanted to assess whether these outlying sections were visually dissimilar. I extracted the single outlier in Figure 4.11A for examination and the sections from this patient are shown in 4.12.

To extract multiple outliers from Figure 4.11B I used a linear model (**lm**) in R (see Table 4.3) to generate residuals. The model of Cores 1 and 2 from the same block had residuals of median -0.03 and range (-0.29-0.55). I extracted the images of the sections for the four pairs of cores with the largest residuals (highlighted with purple squares in Figure 4.11. Images of the four examples of pairs of cores with significantly different percentages of border cells but taken from the same tissue blocks are shown in Figure 4.13. These examples demonstrate that differences in this structural metric map to very distinct morphological differences.

Coefficients:				
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.13	0.053	2.38	0.024 *
Border Cell Percentage	0.328	0.15	2.18	0.0378 *

Table 4.3 Coefficients for the linear model relating border edges between the two tumour blocks. Residuals calculated for a linear model ($y = mx + c$) with y equal to Core 1 Border Cells and x equal to Core 2 Border Cells. Where p-values associated with the fit are less than 0.05, the result is marked with a '*’.

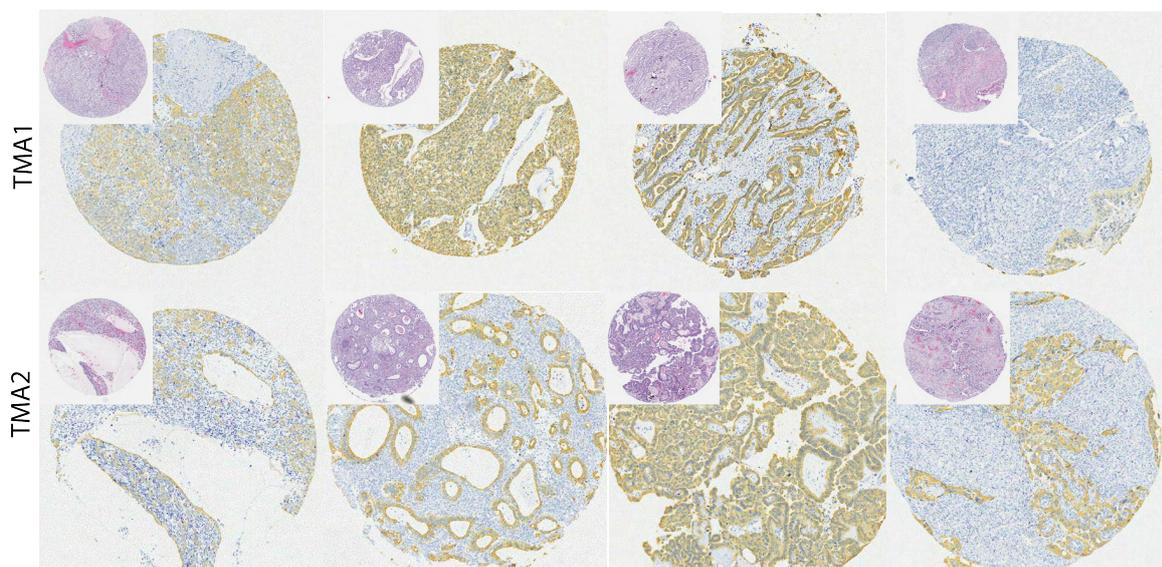


Fig. 4.13 Outlying Cores. Matched samples that have the largest differences for the percentage of border cells between images of sections of different cores from the same tissue block. Accompanying serial section H&E images in top left of each section verify that no random mispositioning or mislabelling of cores during staining or image analysis was responsible.

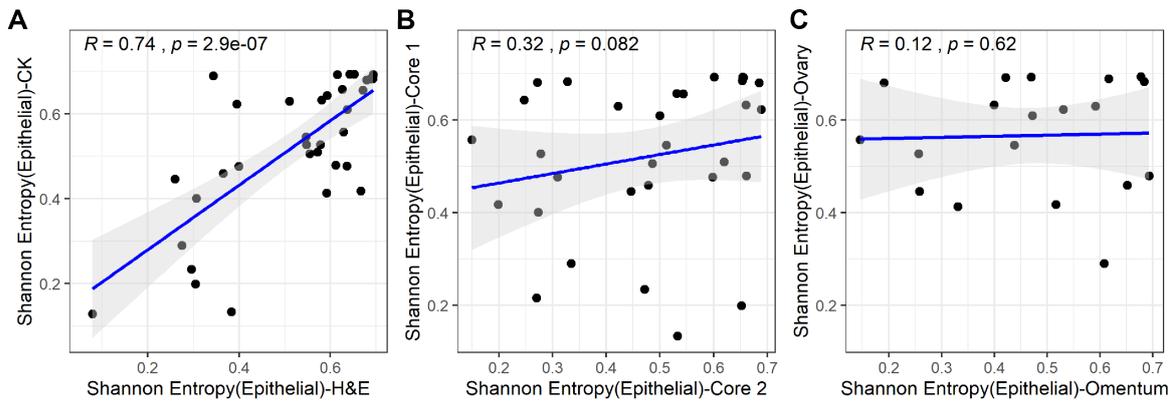


Fig. 4.14 Correlation between Shannon entropy for (A) matched serial sections (B) matched cores from same tumour block (C) different blocks from different sites (Ovary-Omentum) from the same patient. Pearson R and p-value shown.

4.3.7 Entropy

Entropy measures heterogeneity and information content and I proposed the use of Entropy measures in order to assess the heterogeneity of samples. The Shannon Diversity Index is one such measure that has been used in other work in the field [169, 139]. For each tissue image I calculated the Shannon and Batty entropy for the core.

Comparing Shannon Entropy across different tissue samples

The simplest calculation for entropy is the Shannon entropy, the formula for which is given in Methods 2.8. This calculation is also referred to as the Shannon Diversity Index. Shannon entropy is a function of cell type frequency alone. I analysed this across the multiple sample types to investigate whether this was conserved.

Figure 4.14A shows the correlation between the Shannon entropy of images of serial H&E and CK stained sections. There was a moderate correlation between these samples.

Figure 4.14 also compares the measures of entropy across tissue samples from the same tissue block but different cores (B) and the same patient but different tissues (C). The Shannon entropy is not correlated between these, only the Shannon entropy of serial sections in H&E and CK are correlated ($R = 0.74$, $p = 2.9 \times 10^{-7}$).

Given that Shannon entropy is a non-linear function of epithelial fraction and not of spatial organisation, the third result is to be somewhat expected given the previous results in this Chapter.

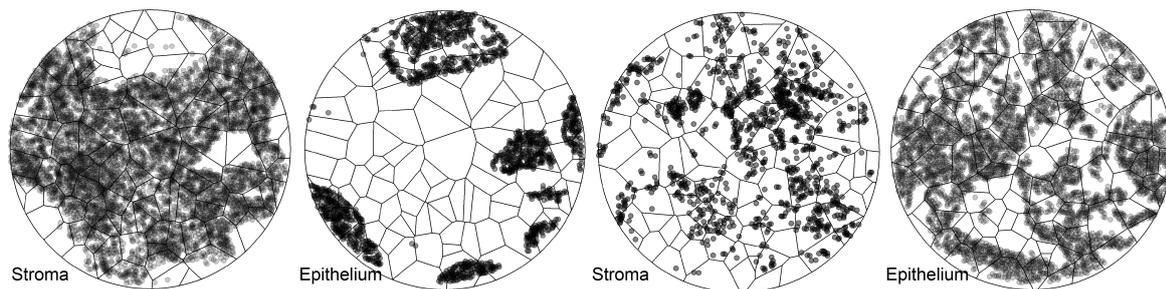


Fig. 4.15 Examples of plotted point pattern of cells and the division of the region into 100 parts for the calculation of Batty entropy.

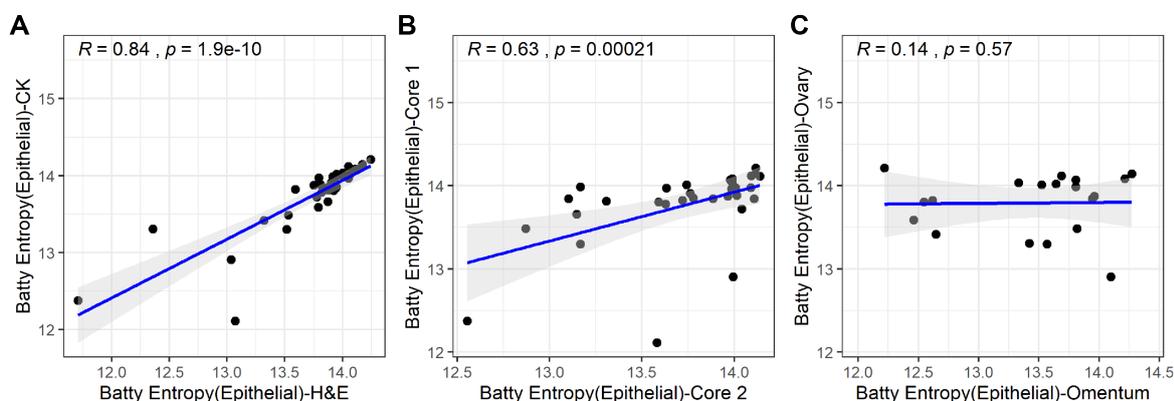


Fig. 4.16 Correlation between Batty entropy of point patterns split into 100 random regions for (A) matched serial sections (B) matched cores from same tumour block (C) different blocks from different sites (Ovary-Omentum) from the same patient. Pearson R and p-value shown.

Comparing Batty Entropy across different tissue samples

There are many other types of entropy measure, some of which incorporate spatial features, Batty entropy is described mathematically in the Methods. Batty entropy effectively splits a spatial distribution into subregions, counts the points within each and calculates the entropy of the arrangement. As the structure of the epithelium is the predominant assessment of morphology, I reduced the point patterns to epithelial cells only and calculated the Batty entropy over 100 regions. Examples of the plotted cell point pattern and the splitting of the region are shown in Figure 4.15. Figure 4.16 shows the correlation plots between the different samples.

Batty Entropy was significantly correlated over serial sections ($R = 0.84$, $p = 1.910 \times 10^{-10}$) and sections from the same tissue block ($R = 0.63$, $p = 0.00021$) but was not correlated between cores from different tissues.

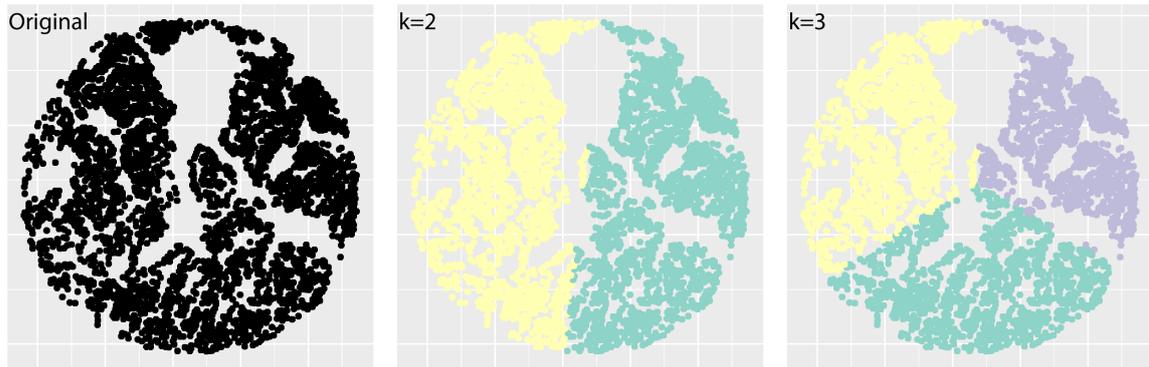


Fig. 4.17 k-means neighbour clustering of epithelial cells identified in a tissue section. Original point pattern of epithelial cells, $k=2$, $k=3$. Densely packed cells are split evenly geometrically into pre-specified clusters with no biological meaning.

4.3.8 Cell Clustering

In observing tissues we see varying clustering of cells; some tissues have solid epithelial sections separate from stroma, others have stromal islands surrounded by epithelium, other are highly mixed. In order to quantify this I aimed to identify cells clustered within a tissue, and use this to understand the overall composition of the core. I assessed multiple methods for clustering cells in the tissue and highlight k-means and DBSCAN methods here.

k-means clustering

An example of a result the of k-means clustering method is shown in Figure 4.17. k-means clustering, as an example, requires a pre-requisite number of clusters, a parameter that varies dramatically between tissue sections. k-means is also an unsuitable method as when tissue sections comprise varying numbers within groups and varying spatial densities of cells, the method tends to split the cells equally into the pre-specified number of clusters allocating by a distance from the center of a cluster, rather than recognising different sized clusters within the data.

Density based clustering

I found that it was better to use DBSCAN, a density based clustering method. The DBSCAN method requires two parameters, epsilon (*eps*), the distance within which two points are considered neighbours and *minPts*, the minimum number of points in a cluster. Both of these are domain knowledge that is more universal than the number of clusters.

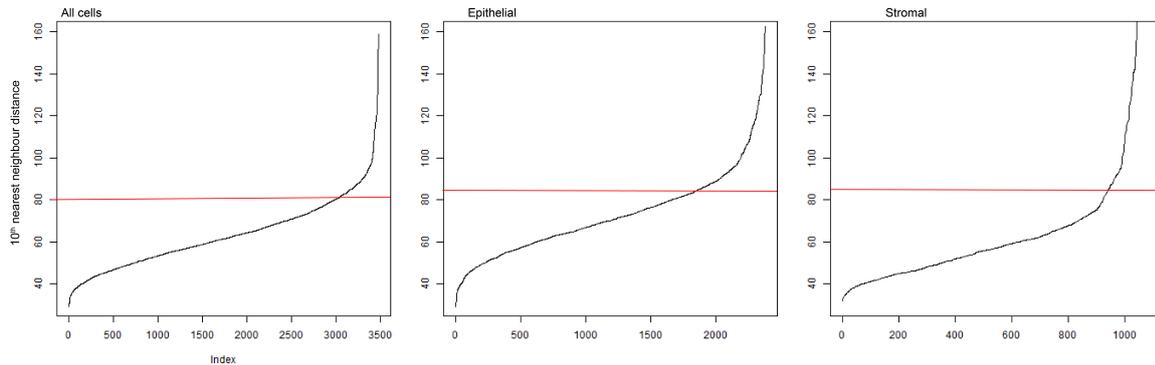


Fig. 4.18 k nearest neighbour (knn) elbow plot to find optimal eps . 10th nearest neighbour plotted against index, the optimal value for the eps is at the elbow of the curve. The eps was approximately 80 for All cells, 82 for Tumour and 85 for Stroma.

I decided to only call clusters of at least 15 cells; given that the number of cells in an image is of the order of 1000 cells, I thought each cluster should contain more than 1% of the cells, keeping the maximum number of clusters to the order of 100. To derive the optimum eps distance for clustering groups of cells on the tissue structure level, I plotted the k^{th} neighbour distribution and located the "elbow" of the curve as discussed in Rahmah et.al[170]. I used $k = 10$. The eps value obtained from these was $80\mu\text{m}$ for Epithelium and $85\mu\text{m}$ for Stromal cells. The larger value was used for clustering all cells. These plots for tumour and stroma and the eps cutoff from each are shown in Figure 4.18.

An example of DBSCAN clustering on serial H&E and CK sections is shown in Figure 4.19. These show visually accurate identifications of cell clusters across H&E and CK images. Boxplots of the distributions of the number of epithelial cell clusters are shown in figure 4.21.

Correlations across structure measures

Figure 4.21 shows the distribution of number of epithelial clusters across the different tissues. The notches represent the confidence interval around the median and the overlapping distributions show that there is no significant difference in the medians of the number of stromal or epithelial clusters in samples.

As shown in Figure 4.22 I analysed whether the number of epithelial clusters was correlated between matched serial sections, cores from the same tumour block and different tissues in the same patient.

The number of epithelial clusters is correlated in serial sections ($R = 0.81, p = 4.8 \times 10^{-10}$) and in cores from the same tissue block ($R = 0.53, p = 0.00054$) but not across sections from different tissues.

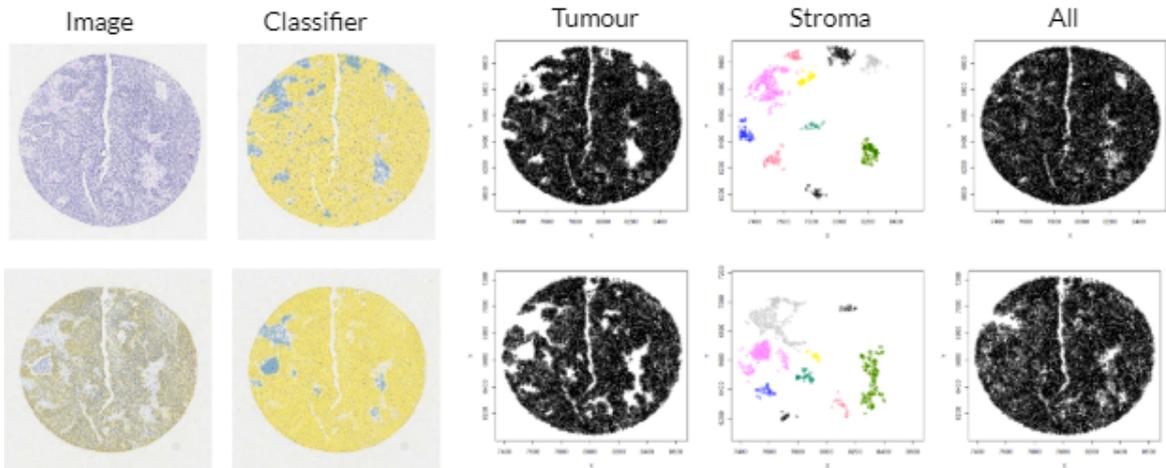


Fig. 4.19 Examples of clusters of cells generated with DBSCAN algorithm. Clusters generated from images of tissue sections when epithelial, stromal cells are viewed separately and when all cells are input.

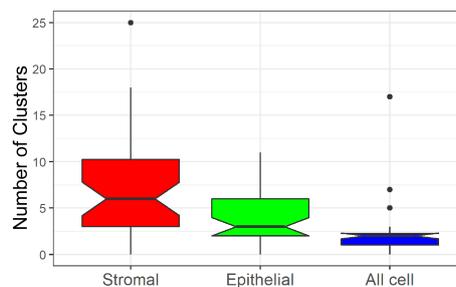


Fig. 4.20 Boxplots for distribution of number of stromal, epithelial and all cell types. Significant differences are seen between distributions showing more stromal clusters in samples than epithelial ones. Clusters over "all cell types" reflect how separated the overall core is. As many cores are one single piece of connected stromal and epithelial tissue, the number of clusters over the whole section is, as expected, much lower than when these tissues are viewed separately. Notches estimate the 95% confidence interval around the median. Whiskers extend to $1.5 \times \text{IQR}$

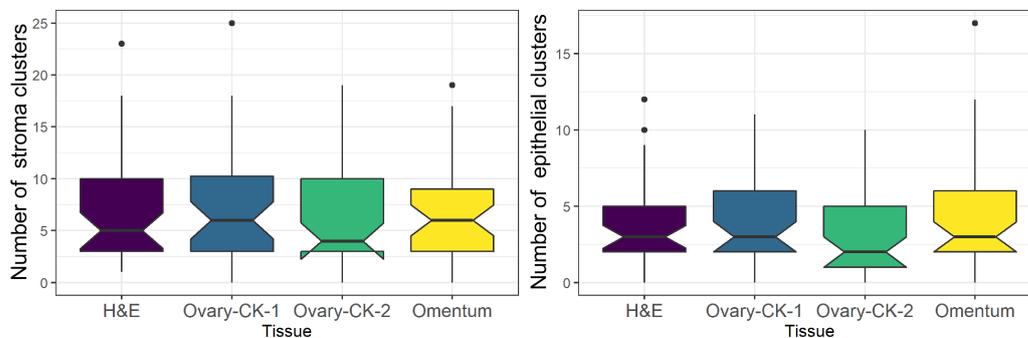


Fig. 4.21 Boxplots for distribution of number of clusters across different sample and staining types. No significant difference is seen in the number of stromal clusters or epithelial clusters across the tissue types. Notches estimate the 95% confidence interval around the median. Whiskers extend to $1.5 \times \text{IQR}$.

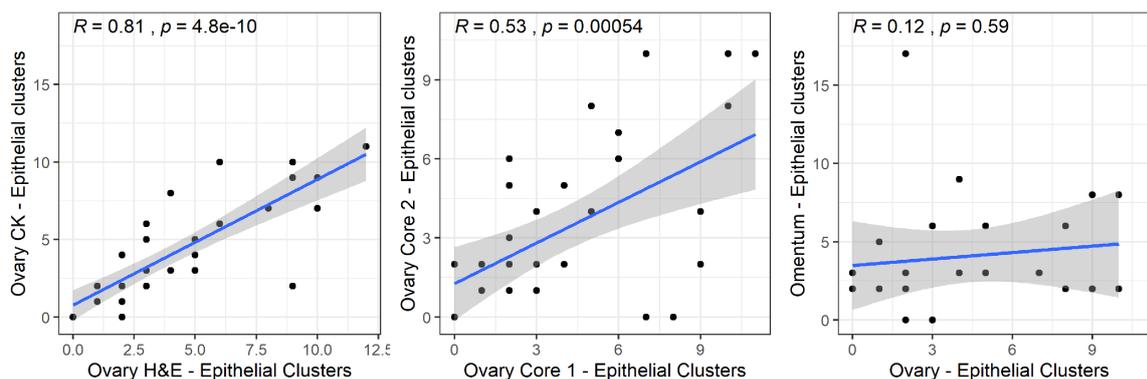


Fig. 4.22 Correlation between number of epithelial clusters for (A) matched serial sections (B) matched cores from same tumour block (C) different blocks from different sites (Ovary-Omentum) from the same patient. Pearson R and p-value shown.

4.3.9 Correlations between structural metrics

Having defined a series of structural metrics, many of which were conserved over different cores of the tissue block, I was interested before utilising these together for classification, how the metrics were related to each other. Figure 4.23 shows the correlations between Epithelial percentage, Percentage of border cells, Shannon Entropy, Batty Epithelial Entropy, Median Epithelial NN dist.



Fig. 4.23 Correlations between structural metrics, there are many positive and negative correlations across metrics and Shannon entropy is a non-linear function of Epithelial percentage.

	PC1 (54.3%)	PC2 (22.3%)	PC3 (11.3%)	PC4 (6.1%)	PC5 (3.8%)	PC6 (1.2%)	PC7 (1.0%)
Batty Entropy	0.33	-0.49	0.05	-0.64	-0.21	-0.04	0.43
Epithelial Nndist	-0.39	0.27	0.30	-0.64	0.47	-0.17	-0.18
Border cell %	-0.46	-0.07	-0.39	-0.23	-0.12	0.75	0.01
Epithelial %	0.48	-0.12	-0.16	-0.23	0.00	0.14	-0.81
Epithelial-Stroma NN	0.40	0.25	0.55	0.07	0.24	0.61	0.19
N Epithelial Clust	-0.34	-0.31	0.65	0.07	-0.52	0.05	-0.30
Shannon entropy	-0.13	-0.72	0.05	0.25	0.63	0.07	-0.03

Table 4.4 Loadings for the principal components of structural metrics selected from analysis.

The most significant correlations in Figure 4.23 are that epithelial percentage is negatively correlated with Epithelial NN dist ($R = -0.74$, $p = 4.7 \times 10^{-7}$) and number of epithelial clusters. It is positively correlated with Batty Entropy ($R = 0.74$, $p = 4.1 \times 10^{-7}$) and negatively correlated with the number of border cells ($R = -0.75$, $p = 2.1 \times 10^{-7}$). A curve is seen between Shannon entropy and Epithelial % as it is a function of Epithelial percentage.

4.3.10 Principal components

Given that these metrics are correlated in complex manners, I performed a principal component analysis to understand if there were patterns in the variation between patients. I performed the principal component analysis on the following metrics extracted from CK stained TMA1: Epithelial percentage, Percentage of border cells, Shannon Entropy, Batty Epithelial Entropy, Median Epithelial NN dist. The loadings of the principal components are shown in Table 4.4.

The first principal component is dominated by the border cell percentage and the epithelial percentage which are negatively correlated. The second principal component is dominated by Shannon entropy. The third principal component is dominated by the number of epithelial clusters.

4.3.11 Morphology classification

Having reproducibly obtained all the properties above, the quantity of stromal and epithelial cells, identified subclusters, entropy, density of tissue and surface area to volume ratio of the epithelium, I aimed to then investigate whether these metrics or the principal components containing them could distinguish the morphological subtypes mentioned earlier.

To do this I extracted the cores at the extremes of the principal components to visualise the differences. Solid and desmoplasia subtypes seemed to be separated at the extremes of PC1. This axis of variation separates disparate small clusters of epithelium with large surface area to volume ratio from large solid masses of epithelium with stromal exclusion.

I also compared individually the metrics measured earlier between samples identified as solid, papillary/glandular and desmoplasia tissue morphologies as shown in Figure 4.24

4.3.12 Relationship between Structure and Immune infiltration

I utilised a classifier trained by Sarwah Al-Khalidi for CD8⁺ and FOXP3⁺ classification to extract the number of CD8⁺ and FOXP3⁺ cells in these samples using HALO. In order to keep analyses between chapters as similar as possible, these were defined as number of cells per unit area (cells per μm^2).

4.3.13 CD8⁺ Correlation

Table 4.5 shows the correlation between structural metrics and the density of epithelial CD8⁺ cells. Before multiple testing correction CD8⁺ density is correlated with multiple features including PC1 and PC3 and the average nuclear epithelial size. After correction these are no longer significant. The results that remain significant are that CD8⁺ infiltrate is positively correlated with Epithelial NN distance ($R = 0.91$, $p = 1.7 \times 10^{-13}$) and less so with the percentage of border cells ($R = 0.54$, $p = 0.009$).

4.3.14 FOXP3⁺ Correlation

Table ?? shows the correlation between structural metrics and the density of epithelial FOXP3⁺ cells. Intraepithelial FOXP3⁺ density is positively correlated with the number of clusters when analysed with DBSCAN over all cell types and more weakly correlated with structural Principal Component 3. After multiple testing correction FOXP3⁺ infiltrate is only correlated with the number of clusters over the whole tissue.

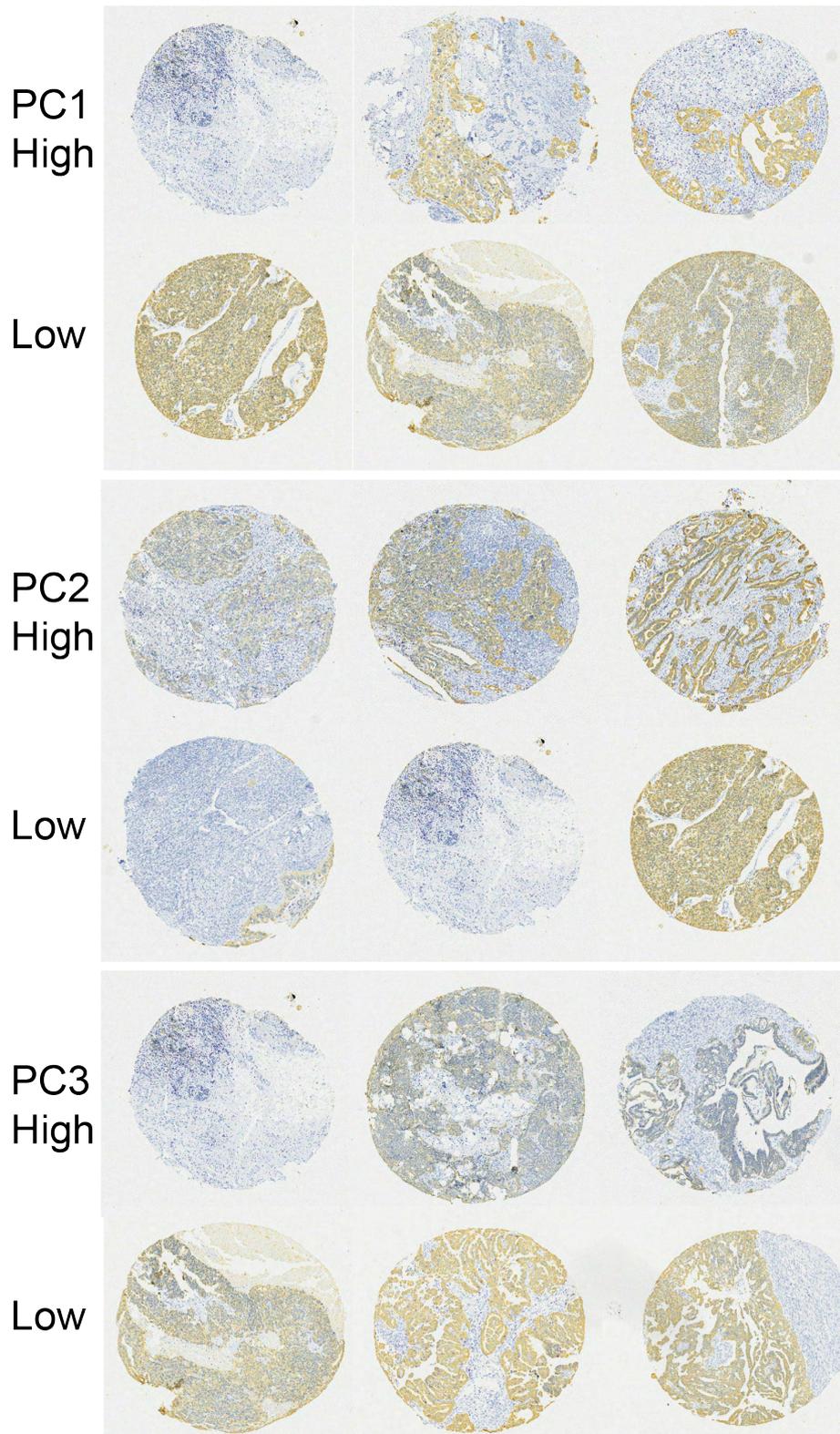


Fig. 4.24 Images of cores stained with pan-cytokeratin are shown. Cores with high and low PC1, 2 and 3 values are displayed. PC1 appears to separate solid from desmoplastic structures. PC2 is dominated by entropy features as seen by mixed and non-mixed populations. PC3 is dominated by epithelial clusters and epithelial and stromal contact, with low PC3 potentially identifying papillary structures which have minimal stroma.

Variable	Correlation	p-value
PC1	-0.53	1.1E-03 **
PC2	-0.14	4.3E-01
PC3	-0.36	3.2E-02 *
PC4	-0.09	5.9E-01
PC5	-0.06	7.2E-01
PC6	0.34	4.3E-02 *
PC7	0.03	8.5E-01
Tumour Clusters	-0.05	7.4E-01
Stroma Clusters	-0.14	3.8E-01
All cell Clusters	0.07	6.8E-01
Batty Epithelial Entropy	-0.09	5.9E-01
Batty Stromal Entropy	0.18	2.8E-01
Epithelial Nucleus Size	-0.35	3.2E-02 *
Border Cells	0.54	5.4E-04 ***
Shannon Entropy	-0.31	5.8E-02
Epithelial NN Dist	0.91	1.0E-14 ****
Stromal NN Dist	-0.29	8.3E-02

Table 4.5 Correlations between spatial metrics and CD8⁺ Epithelial Density. * indicates $p < 0.05$, ** indicates $p < 0.01$, *** indicates $p < 0.001$, **** indicates $p < 0.0001$. After multiple testing correction, values marked *** and above remain significant.

Variable	Correlation	p-value
PC1	-0.03	8.52E-01
PC2	-0.06	7.31E-01
PC3	-0.34	4.84E-02 *
PC4	0.10	5.74E-01
PC5	-0.08	6.33E-01
PC6	0.04	8.03E-01
PC7	0.08	6.30E-01
Tumour Clusters	-0.17	2.91E-01
Stroma Clusters	0.29	6.98E-02
All cell Clusters	0.54	2.91E-04 ***
Batty Epithelial Entropy	-0.03	8.42E-01
Batty Stromal Entropy	0.04	8.18E-01
Epithelial Nucleus Size	-0.21	2.08E-01
Border Cells	0.02	8.90E-01
Shannon Entropy	0.10	5.49E-01
Epithelial NN Dist	-0.20	2.38E-01
Stromal NN Dist	-0.08	6.50E-01

Table 4.6 Pearson Correlation and associated p-values for FOXP3⁺ epithelial infiltrate density and spatial metrics. * indicates p<0.05, ** indicates p<0.01, *** indicates p<0.001, **** indicates p<0.0001. After multiple testing correction, values marked *** and above remain significant.

4.4 Discussion

Attempts like those of Murakami et. al. [97] have been made to robustly classify morphologies of HGSOC tissue. Such a method would be useful in potentially subtyping HGSOC. It is a varied and heterogeneous cancer and automating the classification of structures would allow research to investigate the structural diversity between patients and its impact on prognosis. There have also been observations that patterns differ between *BRCA* and non-*BRCA* mutant tumours and so structural features may be a driver behind the difference in immune infiltration between these groups. Murakami et al. achieved inter-pathologist agreement in their classification study but an automated digital analysis that successfully categorises tissues within HGSOC has not been published. Methods such as deep learning are being attempted but even if successful will not necessarily provide clear insights into the physical features that define the morphologies or a physical understanding of the natural structural variation in samples.

A key problem with current subtyping methods for both morphology and gene expression is that they frequently do not separate the presence of TILs from the classification process. Given the widely accepted positive impact of TILs this often leads them to classify prognostic morphologies and signatures which are purely reflective of immune infiltration[98]. As such I made an effort to create metrics that were generated independently of TIL infiltration.

I generated structural metrics from point patterns of epithelial and stromal cells and examined whether these metrics were maintained across multiple patient samples.

To do this I successfully built classifiers of stroma and epithelium across both Cytokeratin stained and H&E stained slides. I validated these with training and test sets and obtained >98% classifier accuracy for both. The fidelity of the H&E classifier demonstrates a key finding that the methods laid out here need not require new staining of sections and could be applied to already available H&E images that are routinely collected and part of larger public cohorts.

I utilised the point patterns I derived from the nuclei of the segmented and classified cells to obtain metrics of tissue structure, some of which are examined here for the first time. I was able to examine whether these structural metrics were correlated across different patient samples, between different cores and sites. A key finding of this work was that there was a high and significant degree of correlation across all metrics between H&E and CK stained serial sections, further cementing the evidence here that these methods could be transposed to existing H&E data sets.

The metric measuring epithelial cell percentage was significantly correlated across samples from the same tumour block but not across separate sites. This demonstrates that in the

main part the epithelial percentage is consistent across a single tumour block and reinforces that single core sampling is still useful in understanding the development of a tumour.

I defined a metric approximating a surface area to volume ratio of epithelial regions. This was defined as the percentage of epithelial cells which were next to stromal cells. This "border cell percentage" metric separated highly mixed tumours from ones with separate epithelium and stroma. I was also able to use this metric to pick out examples of patients with highly heterogeneous structure between matched cores from the same block. This metric has not been previously utilised but provides a good insight into visual tissue structure and a measure of exposure of the epithelium to the stroma.

I also compared metrics based upon Shannon and Batty entropy, demonstrating that the spatial nature of Batty Entropy increased the strength of correlations across cores from the same tumour block. Batty entropy was significantly correlated across different cores from the same block but Shannon entropy was not. This is likely as Shannon entropy is unable to distinguish a tumour that is 90% epithelial from one which is 90% stromal. The Batty entropy however in this usage incorporates spatial entropy information and is consequently more informative of structure. Extreme values of the Batty entropy correspond to evenly spread and tightly localised epithelium.

I explored cluster based metrics and found DBSCAN to be optimal. Analysing the point patterns for clustering allowed me to observe strong correlations in the number of epithelial and stromal clusters, although the absolute value of this is scale dependent, this could be normalised in future to the tissue area to provide a better scale free metric.

I observed that tumour cell packing, the epithelial NN distance, was correlated across all samples, even between samples in different tissue sites, something that suggests the cell packing aspect of tumour growth is cell-intrinsic, and independent of the microenvironment.

Having derived these metrics I investigated the correlations between them, of interest was the inverse relationship between the positively correlated epithelial cell percentage and Batty entropy and the number of tumour clusters and percentage of border cells. This is quantitative evidence that supports that the expansion of epithelial regions does not result in the even distribution of epithelium but in the merging of epithelium to the exclusion of stromal infiltration. This also is quantitative evidence of tumour cells tending to cluster and clump rather than disperse. The negative correlation therefore between Batty epithelial entropy and Epithelial % is reflective of the observation that the structure is dependent upon the proportion of epithelium the core contains.

I have yet to see other work analysing the structure of tissue separately to the immune infiltration. The lack of such an approach means that the true nature of region-specific features(eg. epithelial infiltrate) cannot be assessed. Research currently fails to address the

covariate of the epithelial and stromal structure that may influence the access of immune cells to the epithelium and the rate at which they infiltrate tumours.

I hypothesised that more stromally infiltrated tumours, with higher border cell percentages, would see higher infiltration of immune cells into the epithelium.

In order to address this question I analysed the correlation between the area density $\log_{10}(\text{cells per } \mu\text{m}^2)$ of FOXP3⁺ and CD8⁺ cells in Epithelium and Stroma and the structural metrics and their Principal Components.

I found that CD8⁺ epithelial infiltrate was correlated with PC1 ($R = -0.53$, $p = 1 \times 10^{-3}$), the percentage of border cells ($R = 0.54$, $p = 5.8 \times 10^{-4}$) and Epithelial NN Distance ($R = 0.91$, $p = 1 \times 10^{-14}$). This finding means that the CD8⁺ density in the epithelium is positively correlated with the density of the cell packing, it is also significantly positively correlated with the percentage of cells classified as border cells. This implies that the epithelial infiltration is a function of border length and cellular packing and unless these factors affect the collagen structure in a tumour, CD8⁺ infiltration is barely affected by structure within the stroma.

FOXP3⁺ cell density in the epithelium was significantly correlated with the total number of clusters in the core ($R = 0.54$, $p = 2.9 \times 10^{-4}$). An increase in clusters over all cell types was associated with more infiltrate. FOXP3⁺ cells are T-regulatory cells and are immunosuppressive[171, 172]. Increased numbers of FOXP3⁺ cells may allow the tumour to evade immune surveillance[173]. Increasing FOXP3⁺ infiltration is associated with increased stage, metastases and worse survival in HGSOc[174, 175]. Increased stage is associated with larger ascites volume and increased ascites volume is associated with reduced immune infiltration and metastasis[176, 171]. Given that gaps in the tissue section may correspond to regions of fluid it may be that the gaps in the solid tumour structure correspond to ascites fluid build up and hence the presence of FOXP3⁺ cells. Future work could verify this by comparing both structure and the number of FOXP3⁺ cells directly with clinical measures such as ascites volume.

Principal components were not much more effective than the individual variables in this case but would be useful for combining even more metrics or infiltrates and were useful for outlining the potential links of structural metrics to morphologies.

In summary I have defined and extracted a novel combination of spatial metrics on point patterns of epithelial and stromal cells. I have found that some structural properties of epithelium correlate with FOXP3⁺ and CD8⁺ infiltrates. Neither the physical border nor the chemo-attractant hypotheses of infiltration are specifically validated by these findings but this work allows for important structural information to be incorporated that should be taken into account when assessing results based upon these hypotheses.

This data set also allows for easy calculation of other structural metrics that can be carried out upon point patterns such as persistent homology[83]. Future work could utilise differences in the metrics here or in other structural metrics across samples and across sites to create a measure for heterogeneity within a patient.

Chapter 5

Collagen as a potential structural driver and influence on immune infiltration

5.1 Introduction

Having explored the tissue structure in Chapter 4 and its relation to CD8 and FOXP3 infiltrate, I found that my measure of surface area to volume ratio of epithelium was positively correlated with quantity of CD8⁺ infiltrate and the distance between neighbouring epithelium, a tissue packing metric, was negatively correlated. FOXP3⁺ on the other hand was weakly correlated with the number of clusters as a whole. This data implies that the structural layout and composition of a tissue is related to the density of immune cells within the epithelium; this motivated me to understand the influence of structure on more immune populations.

I also wanted to assess collagen for numerous reasons, it is a key structural component of tumours and a key component of desmoplastic tissue morphology[177], something that was identified by my structural metrics. Collagen is also associated with metastasis and with movement of cells and immune infiltrate. I wanted to investigate how collagen fits into the picture of structural and immune infiltration differences and investigate the physical border hypothesis in a similar way to Li et al., assessing whether increased collagen deposition decreased access to epithelium for immune cells[178]

In order to investigate the relationship between collagen, tissue structure and a wider immune population at the same time I came up with two experimental plans:

- To quantify collagen through marker free Second Harmonic Generation imaging and simultaneously image CD8⁺ and CD68⁺ infiltration with dual IF. Utilise the OV04 cohort for this to compare these properties with the structural metrics calculated in Chapter 4.

- To expand the assessment of links between structure and collagen to other immune populations by measuring many more markers on a single section than standard IHC and IF.

IMC is one of the best multiplex methods, allowing for measuring more than 30 markers on a single tissue section by conjugating antibodies with metal isotopes. I chose to work with Sarwah Al-Khalidi(SAK) to use IMC to assess additional markers for microenvironmental features such as collagen and hypoxia alongside a panel of immune markers that included the ones that had been investigated in previous chapters. I aimed to carry out the IMC analysis to investigate links between collagen, the structural micro-environment and more immune populations simultaneously in both epithelium and the tumour-adjacent stroma using IMC data from the BriTROC cohort and to carry out new IMC staining and imaging on the ICON7 cohort.

5.1.1 History of the project and Collaborator roles

Jodi Miller (JM - Core facilities) worked with me to optimise the protocol for Dual IF that would be compatible with simultaneous SHG imaging. Sarwah Al-Khalidi was in the process of optimising an IMC panel for immune markers alone, I worked with SAK to include collagen and hypoxia markers in this panel in order to both better delineate stromal tissues from epithelium and investigate the structure of collagen and its interaction with immune cells. Sarwah Al-Khalidi (SAK) carried out the panel optimization and staining on the BriTROC cohort. Fatime Cosaj(FC) carried out the staining for me with the same panel on the ICON7 cohort using the same protocols as optimized by SAK. Richard Grenfell (RG) of the Institute core carried out the imaging of the stained ICON7 slides for me on the same machine as the BRITROC cohort had been done.

From the markers included in the IMC panel, the subset of these of markers that I was interested in were;

- **CD8, CD68, CD45RO, FOXP3** - Key immune cell populations for comparison with other datasets
- **Pancytokeratin and Cytokeratin 7** - Key structural markers, gold standard for comparison of epithelial structure and quantity with other datasets
- **Collagen1** - Marker for the most abundant form of collagen, allowing for the analysis of collagen structure and identification of some stromal areas within tissue
- **CA9** - Hypoxia marker to investigate the relationship between tissue structure and hypoxia

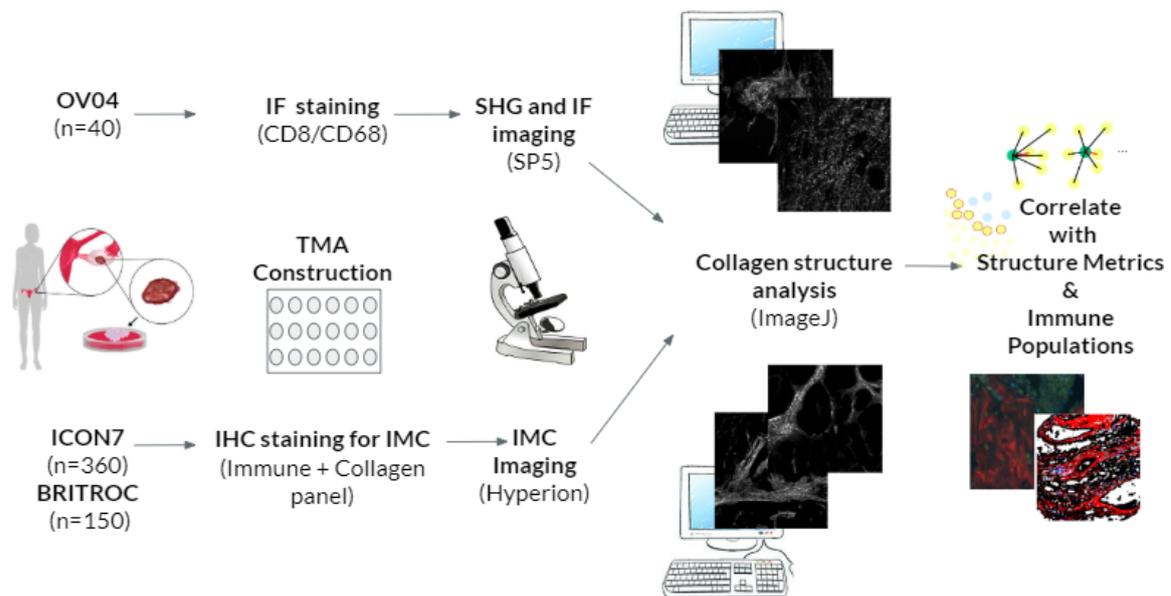


Fig. 5.1 Visual abstract for this chapter. IMC images are processed and collagen channels as well as multi-immune channels are cleaned and then analysed for structural and spatial features.

- **Ki-67** - Proliferation marker to investigate the relationship between tissue structure and proliferation

5.2 Methodology

5.2.1 Cohort Summary

BriTROC

Samples were collected previous to this PhD under the British Translational Research Ovarian Cancer Collaborative (BriTROC), a non-randomised prospective study enrolling patients with recurrent HGSOV[142]. Ethics/IRB approval for this study was given by Cambridge Central Research Ethics Committee (Reference 12/EE/0349). Figure 5.2 shows the REMARK diagram for this cohort. A total of 446 diagnostic samples were collected from 276 patients with relapsed HGSOV. 247 samples from 172 patients had enough material to generate a tissue microarray (TMA) for immunohistochemistry staining. Three 1mm cores from the tumour area of each sample were marked on H&E-stained slides by Luiza Moore and Michelle Loxley and used by Darren Ennis to construct the tissue microarray (TMA) of the samples.

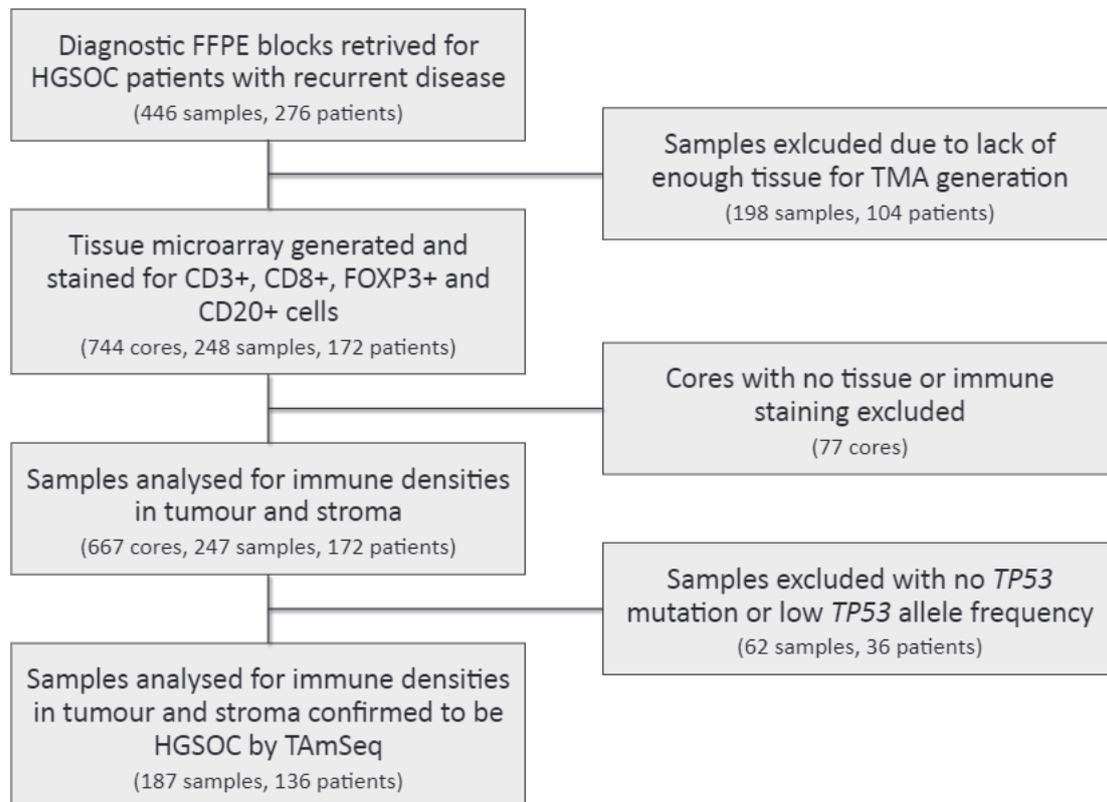


Fig. 5.2 REMARK diagram for the BriTROC cohort.

ICON7

ICON7 was an international, phase 3, open-label, randomised trial undertaken at 263 centres in 11 countries across Europe, Canada, Australia and New Zealand. Eligible adult women with newly diagnosed ovarian cancer that was either high-risk early-stage disease (International Federation of Gynecology and Obstetrics (FIGO) stage I–IIa, grade 3 or clear cell histology) or more advanced disease (FIGO stage IIb–IV), and the study was powered to detect both progression free survival and a difference in overall survival [145, 179]. Samples used for this work were obtained prior to the chemotherapy being administered in the trial. This trial was registered as an International Standard Randomised Controlled Trial, number ISRCTN91273375 [180]. The TMAs for this analysis were prepared by the Medical Research Council (MRC).

ICON7 had cores sampled specifically from epithelium and adjacent stroma, making it ideal for measurement of the microenvironment and measurement of both tumour and stroma in a single section which had been limited to a subset of the patients in previous cohorts.

5.2.2 SHG

SHG microscopy excited at wavelength 920nm and detected at 450–470nm was performed on the Leica SP5 microscope and imaged using a 20x Dry objective. DAKO mounting media and a coverslip of thickness 150 μ m was used.

5.2.3 Dual IF

Dual IF staining for CD8+ and CD68+ cells was performed by JM as described in Section 2.5. DAPI, DRAQ5, FITC, Cy3 and Cy5 dyes were all tested for optimisation. The spectra of the optimal dyes are shown in Figures 5.4, 5.3 and 5.5. These spectra demonstrated the least overlap.

5.2.4 Immuno Metal Conjugation Marker Panel

The marker panel is specified in Table ?? and was optimised by SAK. I worked with SAK to include collagen in this panel for further investigation alongside the immune populations.

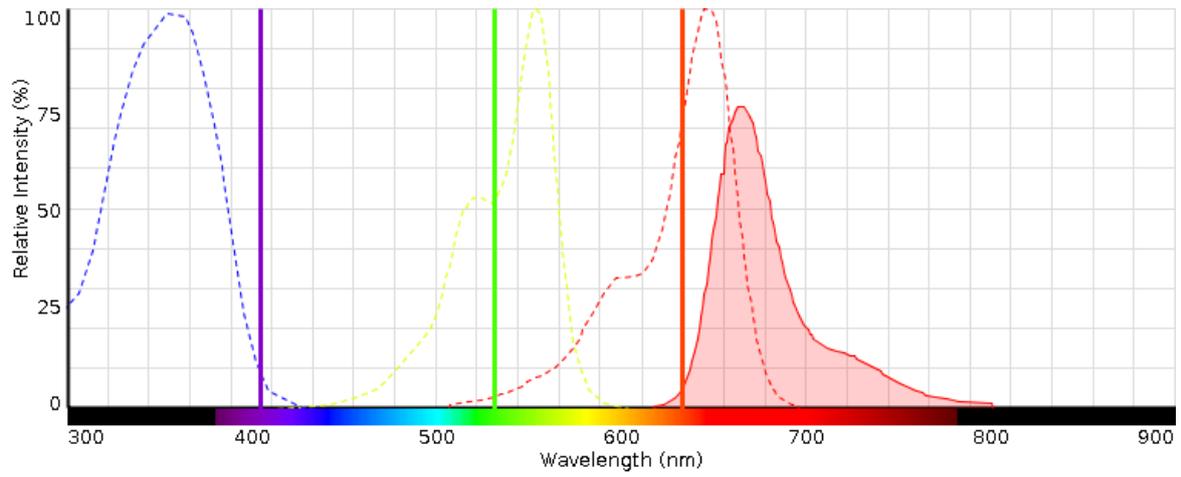


Fig. 5.3 Emission spectrum for Cy5 with excitation at 633nm

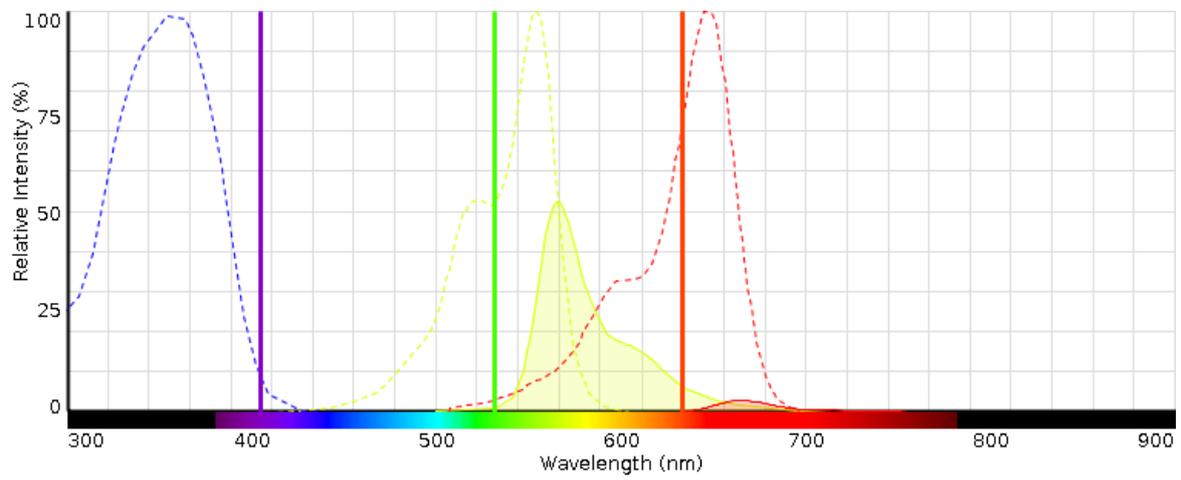


Fig. 5.4 Emission spectrum for Cy3 with excitation at 542nm

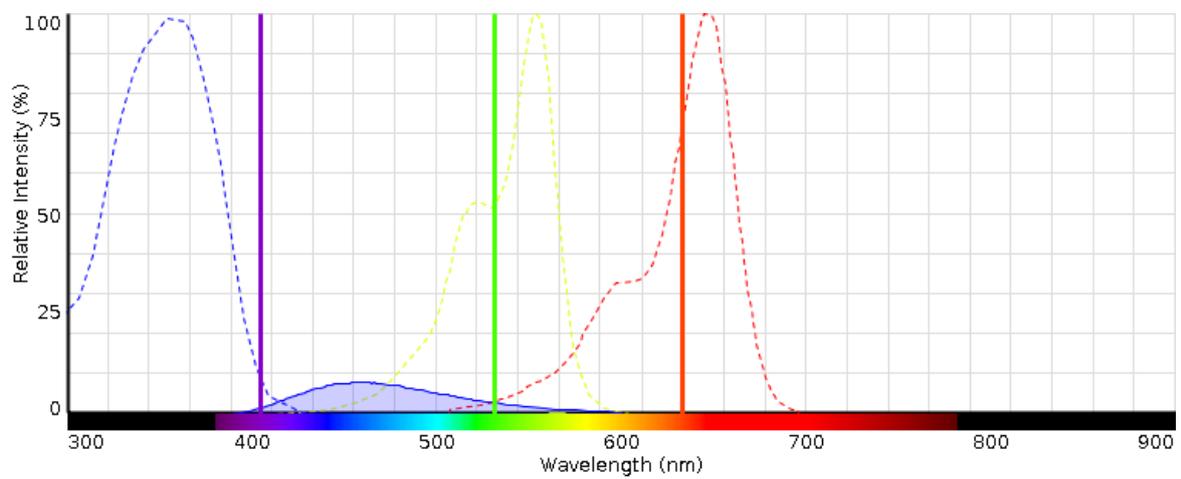


Fig. 5.5 Emission spectrum for DAPI with excitation at 405nm

5.2.5 IMC Staining and Imaging

Staining and imaging of the BriTROC panel was carried out by SAK. Staining and imaging of the ICON7 TMAs was carried out by Fatime Cosaj(FC) and Richard Grenfell(RG) respectively.

5.2.6 Image Analysis

Structure analysis of SHG images and IMC signal in the collagen channel was carried out with Imagej and the GLCM, Orientationj and Directionality plugins.

Python was used for IMC data cleaning and channel labelling and Halo was used for the subsequent analysis of the IMC data files and extraction of immune cells.

5.3 Results

5.3.1 Patient cohorts

BriTROC and ICON7 cohorts were used to analyse the interaction between immune cells and structural markers on single tissue sections are shown.

OV04 summary

I carried forward 40 patients from the primary TMA slide for potential analysis in this section between structural features and collagen.

BriTROC summary

IMC data was available for 136 patients across TMAs in triplicate.

ICON7 summary

There were 3×6 TMA blocks created for the ICON7 patient cohort. One set of 6 sampling epithelium, one sampling epithelium and adjacent stroma and one sampling stroma. I utilized the set of 6 TMA slides for this cohort which contained both epithelium and stroma (n=326).

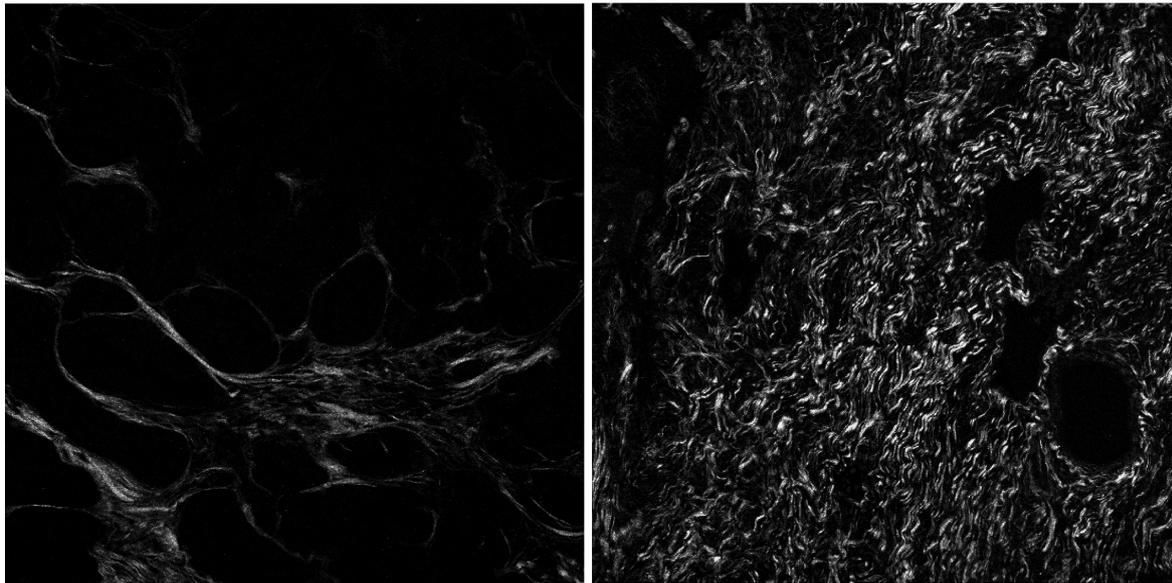


Fig. 5.6 Examples of SHG collagen signal. The left shows thin oriented fibres, the right shows much more collagen distributed across the sample.

5.3.2 SHG Imaging

I imaged the OV04 TMA for SHG signal on the SP5 microscope using the multiphoton laser with excitation at 920nm. Examples of clear collagen fibres showing clear signal and obviously distinct collagen morphologies that were generated are shown in figure 5.6

SHG collagen structure analysis

37 of the 40 OV04 patients included in the OV04 TMA for imaging were successfully imaged for collagen. In order to analyse the collagen structure I used ImageJ and two of its plugins, GLCM and OrientationJ. As coherency and orientation are scale dependent I wrote an ImageJ macro to split the image into 10×10 regions and analyse these features on each one.

I took the median value of the properties in Table 5.1 over the 100 image regions.

To visually validate my collagen analysis I extracted the images with highest coherency, highest energy and highest entropy to confirm they were organised, had a large quantity of collagen and were well mixed respectively.

I also analysed the correlations between these collagen metrics in order to reduce the set to the fundamental properties.

Feature	Definition
Mean	Mean pixel value
Min	Minimum pixel value
Max	Maximum pixel value
IntDen	Sum of all pixels value normalised to area
% Area	Percentage of image with signal
Energy	Uniformity of the image. <i>Constant image returns Energy=1.</i>
Orientation	Orientation of the fibres (degrees)
Coherency	Coherency measures alignment of fibres through structure tensor
Contrast	Returns a measure of the intensity contrast between a pixel and its neighbor over the whole image. <i>Constant image has Contrast=0.</i>
Correlation	Returns a measure of how correlated a pixel is to its neighbor over the whole image.
IDM	Local Homogeneity, Inverse Difference Moment
Entropy	Inhomogeneous images have low entropy, while homogeneous images have a high entropy.

Table 5.1 Definitions of collagen features

5.3.3 Correlation of structural metrics with collagen features

I used the structural metrics generated for these patients in Chapter 3 to analyse the correlation between collagen mean, coherency, correlation and orientation and structural metrics. I split the structural metrics into three categories, nearest neighbour properties(Figure 5.7), entropy(Figure 5.8) and Border cell features (Figure 5.9). Mean collagen was significantly correlated with many variables, an obvious read out of collagen quantity affecting overall structure.

5.3.4 Dual IF and counterstain

In order to assess multiple immune markers on a slide in combination with SHG I decided to use immunofluorescence labelling. I worked with Jodi Miller (JM) from the Institute core to create an optimized panel to measure CD8, CD68 and nuclear stains simultaneously as well as being able to image the collagen in the 450nm channel. We originally attempted to utilize DRAQ5 as a nuclear stain to keep the 450nm channel free. As shown in Figure 5.10 the DRAQ5 successfully stained the nuclei but the FITC and Cy3 channels for the CD8+ and CD68+ marker combinations were not successful on FFPE. This was

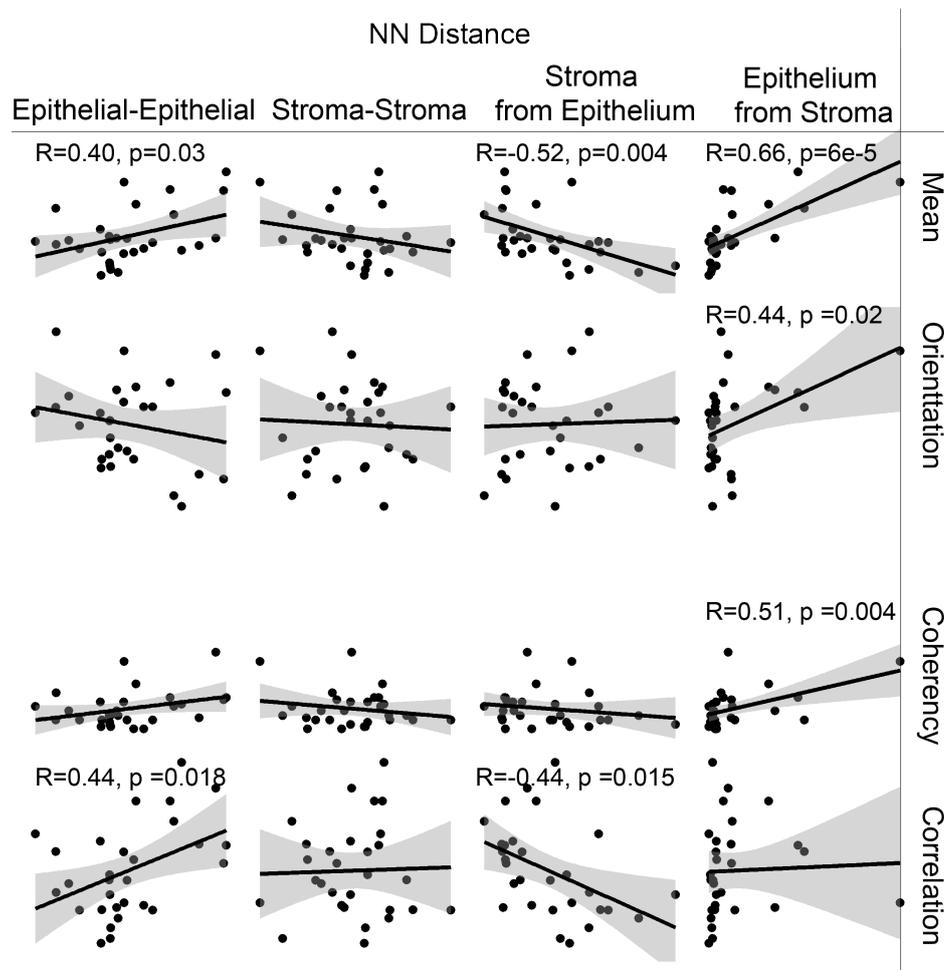


Fig. 5.7 Correlation between structural metrics and collagen. Pearson R and p-values are labelled above the plots for which correlations were significant before multiple testing correction. After multiple testing correction, the correlation between mean collagen signal and the epithelial-stromal and stromal-epithelial nearest neighbour distances were significant, as is the correlation between coherency and the average nearest neighbour distance between epithelium and stromal cells.

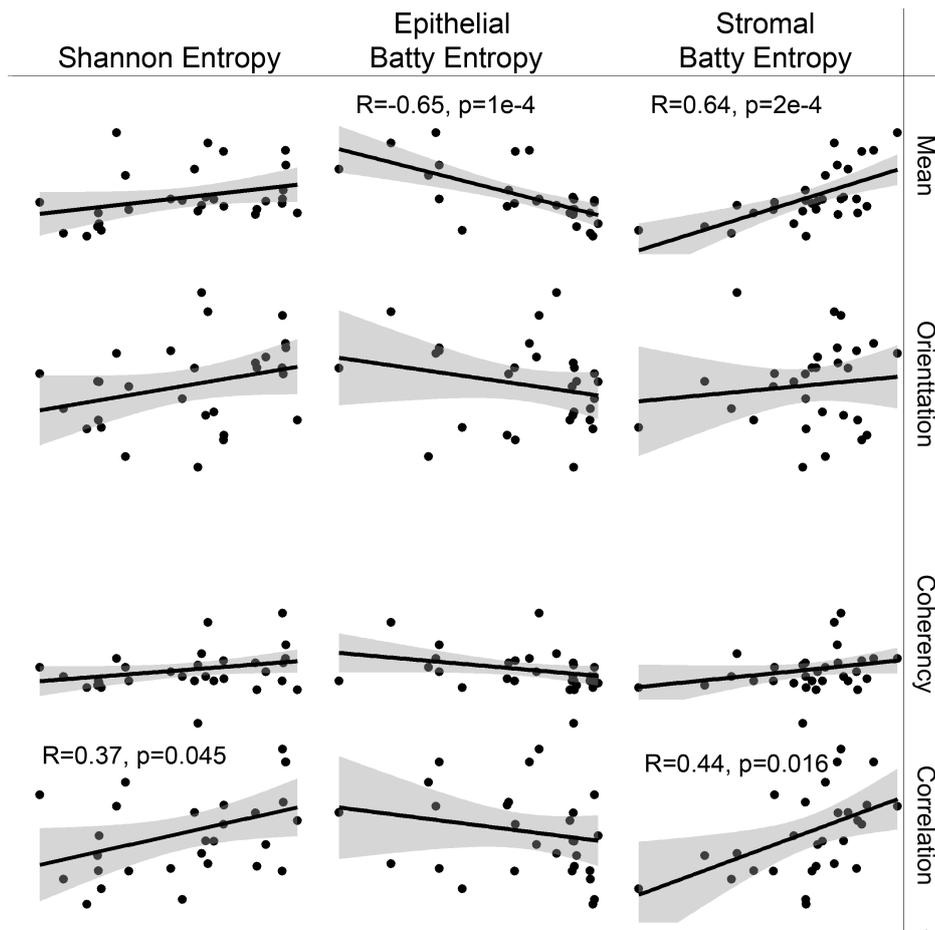


Fig. 5.8 Correlation between entropy metrics and collagen. Pearson R and p-values are labelled above the plots for which correlations were significant before multiple testing correction. After multiple testing correction only the correlation between mean collagen signal and both epithelial Batty entropy and stromal Batty entropy were significant.

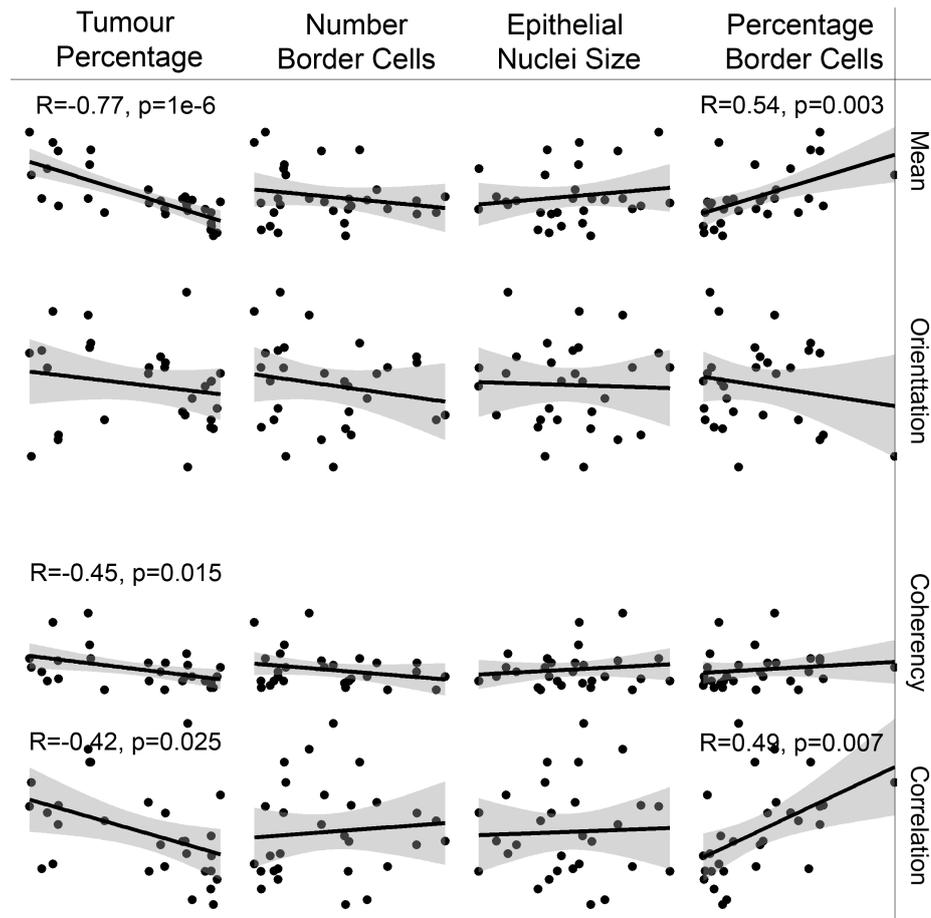


Fig. 5.9 Correlation between Border cell metrics and collagen. Pearson R and p-values are labelled above the plots for which correlations were significant before multiple testing correction. After multiple testing correction only the correlation between mean collagen signal and both tumour percentage and border cells were significant.

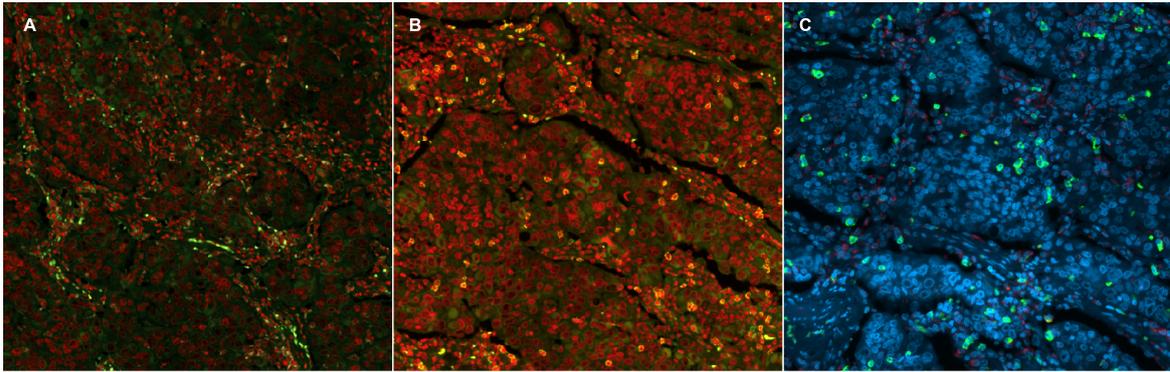


Fig. 5.10 False colour images from IF protocols. (A) DRAQ5(Cy5)-Red, CD8+(FITC)-Green, CD68+(Cy3)-Yellow (B) DRAQ5(Cy5)-Red, CD8+(FITC)- Green, CD68+(Cy3) - Yellow (C) DAPI - Blue, CD8+(Cy3) - Yellow, CD68+(Cy5) - Red. CD68+ Macrophages, CD8+ cells and DAPI stained nuclei are most distinct and visible in protocol C.

partly due to large amounts of autofluorescence from the FFPE section as well as overlapping FITC and Cy3 emission spectra meaning signal could not be easily distinguished. DAPI(Blue)/Cy3(Yellow)/Cy5(Red) staining was very clear and chosen for the rest of the imaging. The final protocol and antibodies used are detailed in Section 2.5.

5.3.5 IF and SHG combined

I tuned the wavelength to excite the SHG at a slightly higher wavelength of 920nm in order to raise the wavelength of the emitted signal, this meant that excitation of and signal from DAPI stained nuclei was minimised such that the collagen signal was clear. Images from this dual staining and SHG are shown in figure 5.11.

5.3.6 Correlation between FOXP3+ and CD8+ infiltration and collagen properties

I assessed the correlation between density of infiltrates and properties of the collagen signal. The Pearson Correlation and p-values for CD8 Density are shown in Table 5.2. There was no significant correlation between the collagen features calculated and CD8+ or FOXP3+ intraepithelial cell densities.

5.3.7 IMC image analysis

IMC images contain hot pixels distributed randomly across the image. In order to clean the data, I wrote a hot-spot removal script in python using a median filter with a 3×3 pixel area.

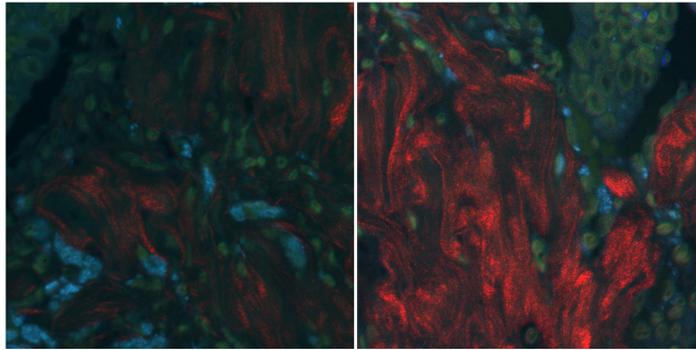


Fig. 5.11 Images measuring IF and SHG signal. False coloured images visualise the following dyes in the following colours: DAPI-green, CD8+(Cy3)-dark blue, CD68+(Cy5)-cyan, SHG-red. CD68+ Macrophages, CD8+ cells and DAPI stained nuclei are very distinct and the fibres of collagen are visible in red. Imaged at 40x on Leica SP5 with tuneable MP, UV, Argon and HeNe Lasers.

	Correlation	P-value
Contrast	0.11	0.53
Energy	0.10	0.57
Coherency	-0.13	0.44
Correlation	-0.06	0.75
Entropy	0.15	0.38

Table 5.2 Correlation between CD8+ epithelial density and collagen properties. No significant relationships were observed.

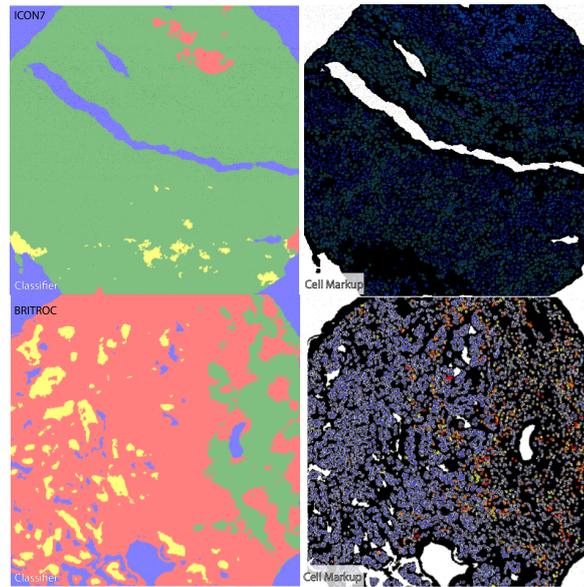


Fig. 5.12 Example of BRITROC and ICON7 classifiers (left) and cell markup (right) from HALO on examples of IMC multichannel images.

I then built a classifier in Halo to distinguish epithelial tissue from low density and high density collagen, specified by the intensity of the Collagen1 marker over small areas. I then built cell segmentation algorithms based on nuclear DNA marker staining to segment nuclei. I then used the IMC marker panel to classify cells based on their nuclear and cytoplasmic staining. An example of tissue segmentation, cell segmentation and classification of ICON7 and BRITROC images is shown in Figure 5.12.

5.3.8 Correlations between more immune populations

The BRITROC cohort comprised 3 TMAs in triplicate, I decided to use the first TMA in order to do an exploratory analysis for hypothesis generation. I would then be better powered to test the statistical significance of such hypotheses on the other two TMAs and the ICON7 cohort.

The first investigation of this dataset was to investigate the correlations within immune populations to validate the results and to build up a picture of which features of the micro-environment are connected.

I first selected a subset of markers relevant to T-cell infiltration and investigated the correlation between numbers of CD3, CD8 and PD-1 positive cell areas in samples (Figure 5.13). I found that these three markers were strongly correlated across samples. I also analysed how three macrophage related markers, CD68, CD163 and PD-1 were associated across samples (Figure 5.14). PD-1, although typically associated with T-cells, has also been

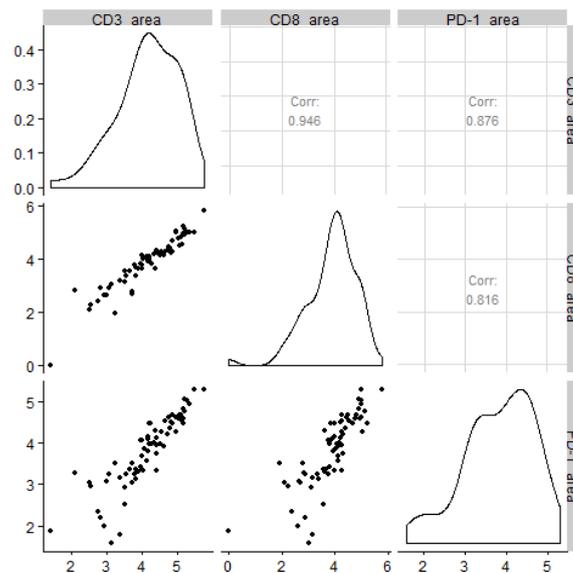


Fig. 5.13 Correlation between markers of T-cell subsets CD3, CD8, PD-1 are shown. Scatter-plots in lower left, distributions (log10 transformed) along the diagonal. Pearson R in top right. Strong correlations are observed between all 3.

found to be expressed on macrophages and is associated with decrease in phagocytosis[47]. These three markers were also correlated across samples.

Figures 5.15 and 5.16 show the Pearson correlations between the number of CD8+, CD68+, CA9, CK7, Collagen1, CD45RO, CD56 and Ki-67+ cells in the BRITROC TMA. CD8, CD68 and CD45RO, the populations investigated in Chapter 1, show the strongest positive correlations. I also see the negative correlation we would expect between the number of collagen and cytokeratin positive cells as these cells are mutually exclusive and together make up the majority of the structure of a tissue. Despite seeing a higher density of infiltrate in stromal regions in earlier chapters, I see a negative correlation between the number of collagen positive cells and the number of immune cells in these samples, potentially due to collagen mediated exclusion.

5.3.9 Collagen deposition and hypoxia

There is evidence of cycles of hypoxia and collagen modification in tumours. CA9 is a hypoxia marker and expression predominantly occurs in epithelial cells and given the negative correlation between epithelial cell quantity and collagen, I normalised the CA9 positive cells as a percentage of total epithelial cells.

I found no correlation between the percentage of epithelium positive for CA9 staining and the area of weak or strong Collagen1 staining. This implies that the quantity of collagen

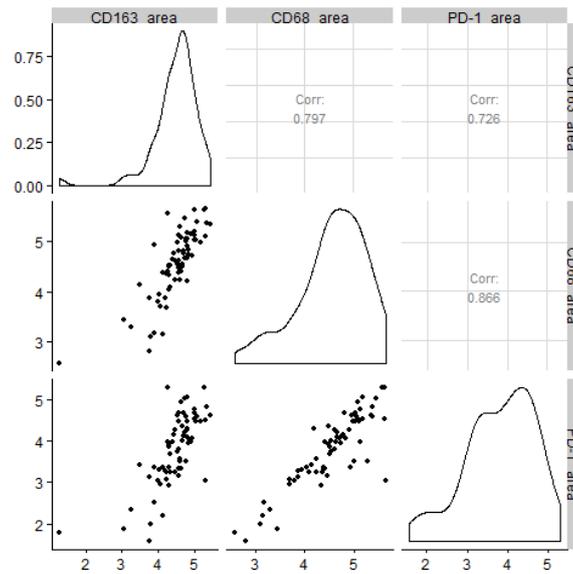


Fig. 5.14 Correlation between markers of macrophage subsets CD68, CD163 and PD-1 are shown. Scatterplots in lower left, distributions (log10 transformed) along the diagonal. Pearson R in top right. Strong correlations are observed between all CD163, CD68, PD-1.

and the stromal infiltration in a sample is not related to the extent of hypoxia, as measured by upregulation of CA9 and hence CA9 staining within the epithelium, at the time the tissue is fixed.

5.3.10 Comparing immune cell densities between epithelium, high density and low density collagen

Macrophages

CD68 and CD163 are both macrophage markers and macrophages are associated with collagen remodelling. I assessed the difference in the quantity of these macrophages across different strengths of collagen staining and the epithelium. I found no difference between the density of macrophage infiltration of collagen and epithelium.

The area density of the CD163+ infiltrate does not vary between regions but CD68+ infiltrate is excluded from dense collagen.

5.4 Discussion

Chapters 3 and 4 discussed how compartmentalised immune infiltrates had differing survival impacts and how structural variation in the tumour epithelium could be linked to changes

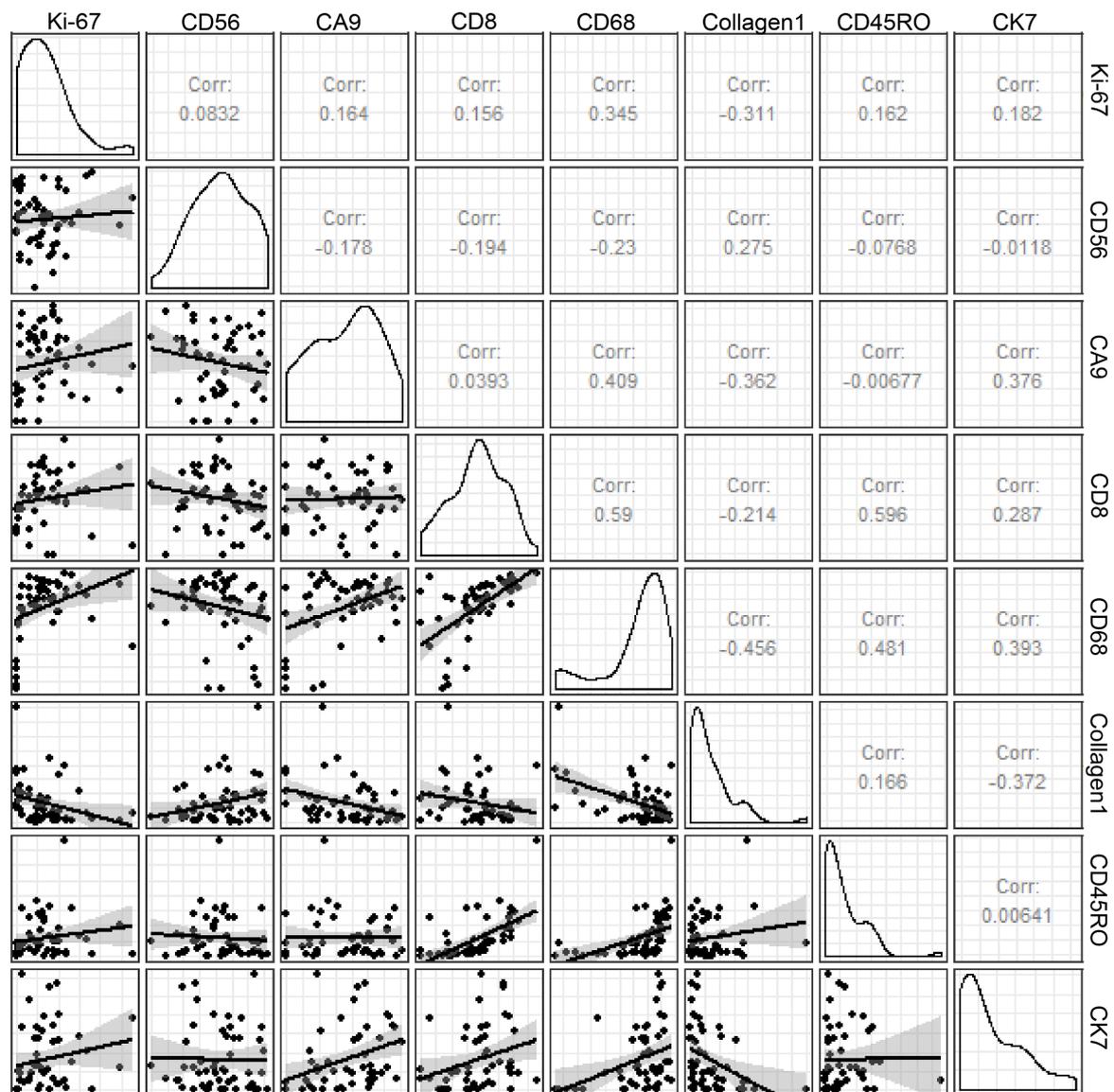


Fig. 5.15 Correlations between cells positive for each marker. Pearson correlation coefficient is given along the top right. Distributions of each infiltrate(log10 transformed) are given down the diagonal. Scatterplots of correlations between number of cells positive for each marker.

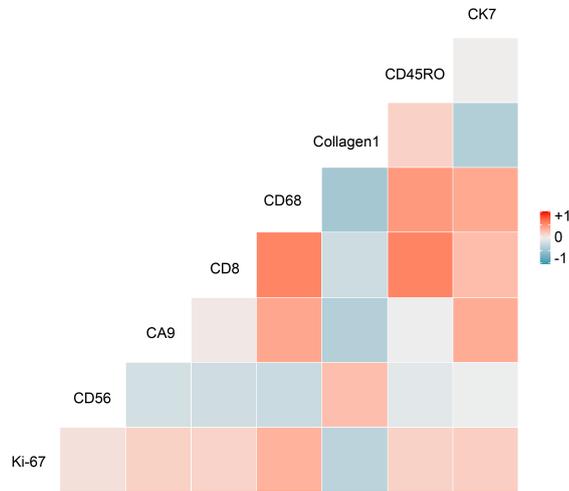


Fig. 5.16 Heatmap of correlations between number of cells positive for each marker. Correlations are seen between CD45RO, CD68 and CD8 cells and negative correlations between collagen and most immune cells.

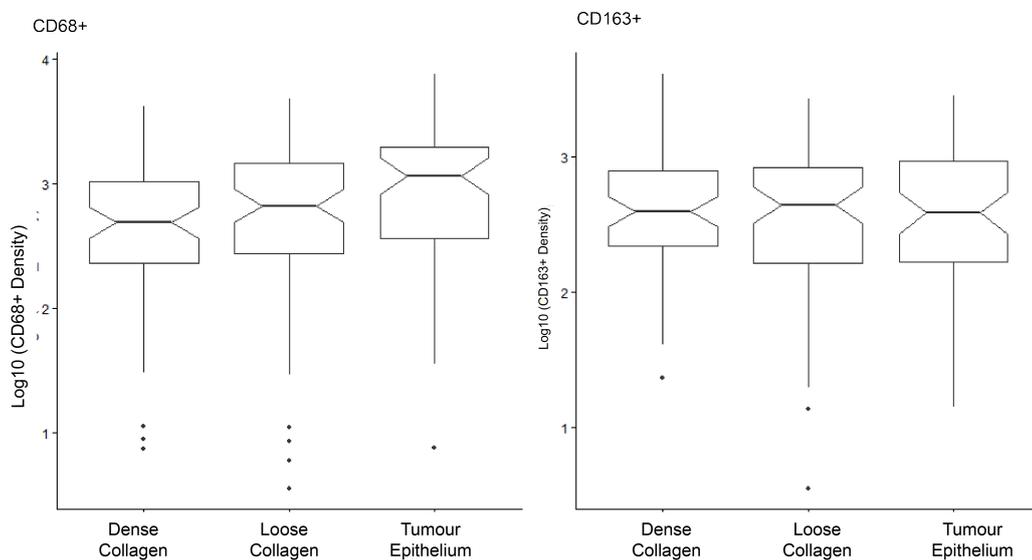


Fig. 5.17 Distribution of densities of CD68+ and CD163+ cells in the dense collagen, loose collagen and epithelial regions of the tumour section. There was no significant difference in CD163+ infiltrate between regions. Significantly higher CD68+ density was seen in Tumour Epithelium than Dense collagen as shown by non-overlapping medians.

in this infiltration. Although this is interesting and a significant finding, this does not provide many mechanistic links to how these structures form or how the structure affects the infiltration and the infiltration occurs. In order to test the hypothesis that a physical collagen border impedes immune cell axis to the epithelium, I analysed collagen on serial sections and in IMC data.

There is a lot of evidence for collagen remodelling in tumours[61]. As mentioned in the literature by Failmezger et. al. amongst others [83] deposition and altering of the structure of collagen could be the important link as to how the structural differences in tumours lead to differing immune infiltration.

I first set out to investigate the relationship of the structure of the tissue with the structure of collagen. In order to investigate the structural properties of collagen I imaged the collagen fibres using a technique called Second Harmonic Generation. I demonstrated that SHG can be imaged simultaneously with dual IF CD68/CD8 and DAPI and showed that the features measured correspond to visual differences in fibre density and alignment.

I was able to investigate the link between these measures of collagen fibre structure and epithelial structure. I found that the collagen energy or mean value, was correlated with a large number of properties, it was strongly negatively correlated with tumour percent and positively correlated with the distance of Epithelial cells from stroma. However, I found no correlation between fibre coherency, energy or correlation and the infiltration of CD8+ or FOXP3+ cells.

This implies that the links with structure of FOXP3+ and CD8+ cells are likely due to something other than corresponding fibre orientation.

In order to test some of these ideas and expand the relationship to more immune populations and alternate relationships to structure such as hypoxia, I utilised the multi-marker method of IMC. I carried out analysis on images of TMAs from BRITROC. I performed preliminary analyses of the relationship between infiltration of immune cells, collagen and the presence of hypoxia. I found that CD68+ cells were excluded from dense stroma whereas CD163+ cells were not. Both CD68 and CD163 are expressed on tumour associated macrophages(TAMs) and the markers are not independent. In some literature CD68+ macrophages are associated with M1 phenotypes whereas CD163+ is associated with the M2a and M2c phenotypes but there is a significant amount of controversy around this phenotyping in the literature[181–183]. CD163+ cells are associated with tissue remodelling so it is perhaps unsurprising they are localised in more dense collagen areas[184]. I found no link between collagen density and either epithelial infiltrate or hypoxia.

I used BRITROC TMA1 to develop an analysis pipeline which can be extended to the other BRITROC TMAs and the ICON7 cohort in future analysis. The ICON7 cohort is ideal

for a full analysis as it has many patients with TMAs sampled from Tumour, Tumour+Stroma and Stroma where possible. It is a unique data set as it is a large and complete dataset for these analyses and specifically samples the tumour-stroma margin.

Chapter 6

Conclusions and Outlook

6.1 Summary

This thesis aimed to advance the understanding of the interplay between structural and the immune response in HGSOE through careful multimodal analyses of tissue samples and the extraction of spatial information from them. In this project I first investigated the nature of sampling and the corresponding compartmental infiltration by various immune cells. Having normalised and examined the distributions and correlations of these infiltrations, I calculated quantitative measures of immune exclusion and combined multiple infiltrations into an immunospace. I assessed the patterns in this multi-dimensional view of the immune infiltration by analysing the principal components within the data and revealed the underlying pattern connecting patients was one of a predominantly coordinated immune response between CD8+, CD68+ and CD45RO+ cells. I highlighted the important confounding factor of the structural makeup of the tumour sample by demonstrating that direct measures of stromal infiltration (CD68+ macrophages) were prognostic and that averaging across epithelium and stroma (CD8+ T-cells) improved CD8 based survival models. I built the first multi-immune survival models in HGSOE and optimised and compared these across the SEARCH cohort.

In part due to the impact and importance of stroma and epithelium highlighted in the initial work, I decided to quantify the tissue structure in the samples in novel ways. To do this I developed methods for the measurement and classification of the structure of tumour sections. I derived accurate and robust structure metrics and observed that some of these metrics were conserved not only across spatially separated samples from ovarian tumours but also across ovarian and metastatic samples in the same patient. I also investigated how these structural features may then be related to immune infiltration and found some were strongly linked to FOXP3+ T-regulatory cell infiltration and CD8+ epithelial infiltration.

My hypothesis was that the link between this structure and the infiltration may be due to collagen deposition. To understand this further I examined collagen as a potential structural driver by carrying out SHG imaging and analysis of orientation and structure. I found that many collagen metrics were correlated with epithelial structural metrics. To then investigate whether the link between structure and infiltration was due to collagen, I developed a method for simultaneous IF and SHG imaging and I analysed correlations between structural collagen metrics and immune infiltrations. In SHG collagen analysis methods combined with IHC staining I observed no link between collagen metrics and epithelial CD8+ and FOXP3+ infiltration.

Given the limited scope of the work at this point to a small subset of immune cells, I chose to do a more in depth analysis of more immune cell types and their potential relation to structure in different cohorts. In order to do this I analysed images that had been obtained using state of the art multimarker IMC methods on samples from the BRITROC cohort. I found no significant correlation between collagen quantity and immune infiltration in this cohort. I also generated a new data set, applying the 30 marker collagen-immune panel from BRITROC to samples from the ICON7 cohort. This generated IMC data for a set of ICON7 patients and the region specific sampling process in the ICON7 trial means that this data set will be used for the examination of very specific immune populations at the tumour-stroma border and the relation of these infiltrates to structure in HGSOEC.

6.2 Conclusions

This data demonstrated that immune populations are continuously distributed on a log-normal scale and that quantitative measures of immune exclusion are too. The popular classification of the immune infiltrate in tumours as hot, cold (immune high or low) or excluded are merely arbitrary cutoffs for high, medium or differential immune cell infiltration where the underlying distribution is normal. This work highlights that investing time in research generating cut-points for these categories is not meaningful when the data is not naturally clustered. Due to this lack of clustering many studies use the median infiltration. Using the median value will group the majority of patients who are near the median, who have very similar immune infiltration, in different groups. Others classifications are guided by clinical data but Glaire *et al.* found similar proportions in the groups when using either of these two methods, indicating the effect is continuous as discussed by Goode *et al.*[185, 27]. As demonstrated in previous work, if cutoffs continue to be used they will vary between pathologists as well as between studies[186]. Any categorical scores of tumour inflammation

should only be utilised with the knowledge that the infiltration in the microenvironment is actually a continuum.

In Chapter 3 I also demonstrated that immune infiltration into the stroma and epithelium have different prognostic impacts. This implied that structure and localisation can play a significant role in the impact of this immune infiltration. A difference in survival impact that depended on macrophage location was observed, with stromal macrophages being positively prognostic and the ratio of epithelial macrophages to stromal macrophages was a negative prognostic indicator. This observation is in agreement with the plastic nature of macrophages and with the importance of spatial location in phenotyping[42, 45].

The derivation of structural metrics has shown that H&E can be utilised to assess tumour structure and that some tumour structure features are conserved and related to immune infiltration. The success of these methods may allow tumour architecture and grade, which was previously difficult to define and measure, to be assessed practically and these structural features to be further utilised[95, 93]. The relationship that was found between FOXP3+ and tissue structure may reflect properties of the tumour such as fluid or ascites build up[176, 187].

In accordance with Li *et al.*; despite finding relationships between tumour structure and FOXP3 and CD8 infiltration I found no link between collagen properties and infiltration[178]. This work supports the chemoattractant/repellant-model of tumour immune modulation.

This result, in combination with the analysis of structure, suggests that the spatial localisation of epithelial cells (at the edge or centre of a tumour) may affect their phenotype and immune cell recruitment. As shown, the structural metrics correlate strongly with collagen features and as such care must be taken when interpreting measures of collagen that these features aren't being conflated. This is particularly because, for example, the physical border hypothesis requires collagen deposition to be the driving feature of infiltration control, but the chemo-attractive/repellant hypothesis of epithelial infiltration can have structural dependence without the need for a collagen border.

6.3 Future Work

This project and its cohorts are well developed and allow for future work in many of the areas this thesis covered. The findings within spatial structure analysis and the derived metrics can be used to train classifiers to classify the architecture of tissues based on histopathologist scoring. A robust classification of tissue structure has been something attempted for a considerable amount of time[149, 188] and the methods discussed in this thesis now allow for an automated classification system to be built. Further structural analysis and structural metrics could also be investigated. Persistent homology, a subset of topological data analysis[189],

and related methods are gaining traction in the analysis of biological structures and have been found to be a prognostic classifier in breast cancer[190].

Future work upon these cohorts will also allow for in depth profiling of immune cells. Of most interest from this work are macrophages and analysing their phenotype in more detail with markers such as CD163 and CD56 may help to explain the results we observed which suggest plasticity based on location as in Yang *et al.*[42]. Once any immune subset has been identified, its distribution within the tumour can also be analysed. The phenotype of immune cells within the tumour and the stroma may vary and this can be investigated by expanding on the work in Chapter 5. This would help in further understanding why the localisation of CD68+ macrophages in the stroma was associated with improved survival and whether this effect is observed in other immune cell subsets.

Another question to consider within macrophages and the immune infiltration as a whole is whether there is significant clustering between immune cells and how this impacts survival. Macrophages have been shown to impede the function and dispersion of other immune cells[191] and given the requirement of these cells to interact this could be somewhat reflected in their localisation. Macrophages expressing PD1 have also been shown to be less effective at phagocytosis and PD1 expression increases with tumour stage[47]. Given the role of PD1 in immune suppression, the potential impacts of macrophage-lymphocyte interactions may also be investigated through the further analysis of cells expressing surface markers such as PD1. The data-sets in this project can be used in this way to identify cell subsets and the relationship between expression of PD1 and immune-immune distances. The k nearest neighbour metrics applied in this thesis to epithelial and stromal cells can also be applied to immune cells as has been done in breast cancer and melanoma[140, 134].

This work and its analysis also leads the way for an even more thorough investigation of the immune relationship with collagen structure, which can be obtained from both the IMC images and using SHG. SHG can be used alongside the dual IF protocol optimised for this project by the Core facilities at the CI. ICON7, being such a high dimensional data set with collagen staining, also provides a good opportunity for advanced structural analysis. Future work on the ICON7 data set will involve identifying very specific immune cell populations as has been carried out in an IMC analysis in breast cancer by Ali *et al*[192]. As the ICON7 tissue sampling is specifically targeted to three tissue regions per sample, a stromal, epithelial and tumour-stroma margin core, it is also ideal for comparing the infiltration in different tissue regions of the same patient, allowing us to compare infiltration at the centres of epithelial regions to infiltration at the margin.

Having demonstrated that the collagen density and coherency is not influencing infiltration by CD8+ and FOXP3+ cells, the chemo-attractant model could be explored further on the

cohorts from this analysis. Markers for chemokine receptor expression could be combined with immune and structural analysis on future IF, IHC or IMC panels to assess the spatial distribution within the epithelium and potentially explain the link between immune infiltration and structure[193, 194]. Mathematical models of immune infiltration patterns such as that by Li *et al.* could also be built and validated with these data sets[178]. Spatial modelling techniques could be carried out using diffusion models that contain repellant and attractant chemokine terms in similar methods to those by Makaryan *et al.*[195, 196]

Further work could also be done on improving the collagen metrics and the use of collagen alignment in these samples. Gole *et al.* use SHG imaging in triple negative breast cancer to classify collagen in each image into aggregated thick collagen (ATC) and dispersed thin collagen (DTC) categories in a similar method to the one in this thesis[197]. Improvements to my methods and metrics other than median values may both allow for more informative descriptions of the collagen and the assessment of it as a prognostic tool in ovarian cancer. Future work could also relate this tissue structural feature of the tumour to magnetic resonance imaging(MRI) texture and habitats within patient tumours. Work in mouse models related histology to MRI habitats[198] and the same could be done within human tissues to eventually obtain more information from non-invasive MRI scans. Furthermore, even if collagen is not the driver of immune infiltration into the epithelium, the collagen structure is strongly linked to metastasis in the literature[69] and these multiple cohorts with collagen imaging, structural and survival data would allow for a deeper analysis on the specific collagen structures that can lead to metastasis in patients.

References

- [1] D. Hanahan and R. A. Weinberg, "Hallmarks of Cancer: The Next Generation," *Cell*, vol. 144, pp. 646–674, Mar 2011.
- [2] "T Cells - Production of T Cells - Types of T Cells - TeachMePhysiology," Jan 2020. [Online; accessed 28. Jan. 2020].
- [3] L. Attisano and J. L. Wrana, "The return of Dr Jekyll in cancer metastasis," *EMBO J.*, vol. 31, p. 4486, Dec 2012.
- [4] E. Lengyel, "Ovarian Cancer Development and Metastasis," *Am. J. Pathol.*, vol. 177, pp. 1053–1064, Sep 2010.
- [5] P. Mehlen and A. Puisieux, "Metastasis: a question of life or death," *Nat. Rev. Cancer*, vol. 6, pp. 449–458, Jun 2006.
- [6] S. I. Labidi-Galy, E. Papp, D. Hallberg, N. Niknafs, V. Adleff, M. Noe, R. Bhattacharya, M. Novak, S. Jones, J. Phallen, *et al.*, "High grade serous ovarian carcinomas originate in the fallopian tube," *Nat. Commun.*, vol. 8, 2017.
- [7] D. Bell, A. Berchuck, M. Birrer, J. Chien, D. W. Cramer, F. Dao, R. Dhir, P. DiSaia, H. Gabra, P. Glenn, *et al.*, "Integrated genomic analyses of ovarian carcinoma," *Nature*, vol. 474, pp. 609–615, Jun 2011.
- [8] J. N. Buy and M. Ghossein, *Surface Epithelial-Stromal Tumors of the Ovary*. Springer, Berlin, Heidelberg, Jan 2013.
- [9] J. Prat, F. C. On Gynecologic Oncology, H. Belhadj, J. Berek, A. Bermudez, N. Bhatla, J. Cain, L. Denny, K. Fujiwara, N. Hacker, *et al.*, "Staging classification for cancer of the ovary, fallopian tube, and peritoneum," *Int. J. Gynaecol. Obstet.*, vol. 124, pp. 1–5, Jan 2014.
- [10] M. Stanske, S. Wienert, D. C. Castillo-Tong, C. Kreuzinger, I. Vergote, S. Lambrechts, H. Gabra, C. Gourley, R. N. Ganapathi, I. Kolaschinski, *et al.*, "Dynamics of the Intratumoral Immune Response during Progression of High-Grade Serous Ovarian Cancer," *Neoplasia*, vol. 20, pp. 280–288, Mar 2018.
- [11] E. Ricciardi, T. Baert, B. Ataseven, F. Heitz, S. Prader, M. Bommert, S. Schneider, A. du Bois, and P. Harter, "Low-grade Serous Ovarian Carcinoma," *Geburtshilfe Frauenheilkd.*, vol. 78, p. 972, Oct 2018.

- [12] J. Ferlay, I. Soerjomataram, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, D. M. Parkin, D. Forman, and F. Bray, "Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012," *Int. J. Cancer*, vol. 136, pp. E359–E386, Mar 2015.
- [13] M.-A. Lisio, L. Fu, A. Goyeneche, Z.-h. Gao, and C. Telleria, "High-Grade Serous Ovarian Cancer: Basic Sciences, Clinical and Therapeutic Standpoints," *Int. J. Mol. Sci.*, vol. 20, Feb 2019.
- [14] S. W. Johnson, R. F. Ozols, and T. C. Hamilton, "Mechanisms of drug resistance in ovarian cancer," *Cancer*, vol. 71, pp. 644–649, Jan 1993.
- [15] D. Hanahan and R. A. Weinberg, "The Hallmarks of Cancer," *Cell*, vol. 100, pp. 57–70, Jan 2000.
- [16] G. Macintyre, T. E. Goranova, D. De Silva, D. Ennis, A. M. Piskorz, M. Eldridge, D. Sie, L.-A. Lewsley, A. Hanif, C. Wilson, *et al.*, "Copy number signatures and mutational processes in ovarian carcinoma," *Nat. Genet.*, vol. 50, pp. 1262–1270, Aug 2018.
- [17] M. Olivier, M. Hollstein, P. Hainaut, M. Olivier, M. Hollstein, and P. Hainaut, "TP53 Mutations in Human Cancers: Origins, Consequences, and Clinical Use," *Cold Spring Harbor Perspect. Biol.*, vol. 2, p. a001008., Jan 2010.
- [18] B. Leroy, J. L. Fournier, C. Ishioka, P. Monti, A. Inga, G. Fronza, and T. Soussi, "The TP53 website: an integrative resource centre for the TP53 mutation database and TP53 mutant analysis," *Nucleic Acids Res.*, vol. 41, p. D962, Jan 2013.
- [19] A. K. Folkins, E. A. Jarboe, A. Saleemuddin, Y. Lee, M. J. Callahan, R. Drapkin, J. E. Garber, M. G. Muto, S. Tworoger, C. P. Crum, *et al.*, "A Candidate Precursor to Pelvic Serous Cancer (p53 Signature) and Its Prevalence in Ovaries and Fallopian Tubes From Women With BRCA Mutations," *Gynecol. Oncol.*, vol. 109, pp. 168–173, May 2008.
- [20] A. Saleemuddin, A. K. Folkins, L. Garrett, J. Garber, M. G. Muto, C. P. Crum, S. Tworoger, A. Saleemuddin, A. K. Folkins, L. Garrett, *et al.*, "Risk Factors for a Serous Cancer Precursor ("p53 Signature") in Women With Inherited BRCA Mutations," *Gynecol. Oncol.*, vol. 111, pp. 226–232, Nov 2008.
- [21] F. C. Martins, I. de Santiago, A. Trinh, J. Xian, A. Guo, K. Sayal, M. Jimenez-Linan, S. Deen, K. Driver, M. Mack, *et al.*, "Combined image and genomic analysis of high-grade serous ovarian cancer reveals PTEN loss as a common driver event and prognostic classifier," *Genome Biol.*, vol. 15, pp. 1–15, Dec 2014.
- [22] M. M. Kamieniak, I. Muñoz-Repeto, D. Rico, A. Osorio, M. Urioste, J. García-Donas, S. Hernando, L. Robles-Díaz, T. R. Y. Cajal, A. Cazorla, *et al.*, "DNA copy number profiling reveals extensive genomic loss in hereditary BRCA1 and BRCA2 ovarian carcinomas," *Br. J. Cancer*, vol. 108, pp. 1732–1742, Apr 2013.
- [23] J. A. Joyce and D. T. Fearon, "T cell exclusion, immune privilege, and the tumor microenvironment," *Science*, vol. 348, pp. 74–80, Apr 2015.

- [24] A. Jiménez-Sánchez, D. Memon, S. Pourpe, H. Veeraraghavan, Y. Li, H. A. Vargas, M. B. Gill, K. J. Park, O. Zivanovic, J. Konner, *et al.*, “Heterogeneous Tumor-Immune Microenvironments among Differentially Growing Metastases in an Ovarian Cancer Patient,” *Cell*, vol. 170, p. 927, Aug 2017.
- [25] B. Clarke, A. V. Tinker, C.-H. Lee, S. Subramanian, M. van de Rijn, D. Turbin, S. Kalloger, G. Han, K. Ceballos, M. G. Cadungog, *et al.*, “Intraepithelial T cells and prognosis in ovarian carcinoma: novel associations with stage, tumor type, and BRCA1 loss,” *Mod. Pathol.*, vol. 22, pp. 393–402, Dec 2008.
- [26] D. D. Bowtell, S. Böhm, A. A. Ahmed, P.-J. Aspuria, R. C. Bast, Jr., V. Beral, J. S. Berek, M. J. Birrer, S. Blagden, M. A. Bookman, *et al.*, “Rethinking ovarian cancer II: reducing mortality from high-grade serous ovarian cancer,” *Nat. Rev. Cancer*, vol. 15, pp. 668–679, Nov 2015.
- [27] Ovarian Tumor Tissue Analysis (OTTA) Consortium, E. L. Goode, M. S. Block, K. R. Kalli, R. A. Vierkant, W. Chen, Z. C. Fogarty, A. Gentry-Maharaj, A. Tołoczko, A. Hein, *et al.*, “Dose-Response Association of CD8+ Tumor-Infiltrating Lymphocytes and Survival Time in High-Grade Serous Ovarian Cancer,” *JAMA Oncol.*, vol. 3, p. e173290., Dec 2017.
- [28] E. Sato, S. H. Olson, J. Ahn, B. Bundy, H. Nishikawa, F. Qian, A. A. Jungbluth, D. Frosina, S. Gnjjatic, C. Ambrosone, *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.*, vol. 102, pp. 18538–18543, Dec 2005.
- [29] Y. Gao, L. Chen, G. Cai, X. Xiong, Y. Wu, D. Ma, S. C. Li, and Q. Gao, “Heterogeneity of immune microenvironment in ovarian cancer and its clinical significance: a retrospective study,” *OncImmunity*, vol. 9, p. 1760067, Jan 2020.
- [30] N. G. Nwani, L. E. Sima, W. Nieves-Neira, and D. Matei, “Targeting the Microenvironment in High Grade Serous Ovarian Cancer,” *Cancers*, vol. 10, Aug 2018.
- [31] T. J. Curiel, G. Coukos, L. Zou, X. Alvarez, P. Cheng, P. Mottram, M. Evdemon-Hogan, J. R. Conejo-Garcia, L. Zhang, M. Burow, *et al.*, “Specific recruitment of regulatory T cells in ovarian carcinoma fosters immune privilege and predicts reduced survival,” *Nat. Med.*, vol. 10, pp. 942–949, Sep 2004.
- [32] E. Sato, S. H. Olson, J. Ahn, B. Bundy, H. Nishikawa, F. Qian, A. A. Jungbluth, D. Frosina, S. Gnjjatic, C. Ambrosone, *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.*, vol. 102, pp. 18538–18543, Dec 2005.
- [33] J. F. A. P. Miller, G. F. Mitchell, and N. S. Weiss, “Cellular Basis of the Immunological Defects in Thymectomized Mice,” *Nature*, vol. 214, pp. 992–997, Jun 1967.
- [34] W. K. Decker, R. F. da Silva, M. H. Sanabria, L. S. Angelo, F. Guimarães, B. M. Burt, F. Kheradmand, and S. Paust, “Cancer Immunotherapy: Historical Perspective of a Clinical Revolution and Emerging Preclinical Animal Models,” *Front. Immunol.*, vol. 8, 2017.

- [35] J. Lu, Y. Xu, Y. Wu, X.-y. Huang, J.-w. Xie, J.-b. Wang, J.-x. Lin, P. Li, C.-h. Zheng, A.-m. Huang, and C.-m. Huang, "Tumor-infiltrating CD8+ T cells combined with tumor-associated CD68+ macrophages predict postoperative prognosis and adjuvant chemotherapy benefit in resected gastric cancer," *BMC Cancer*, vol. 19, pp. 1–10, Dec 2019.
- [36] C. G. Ioannides, R. S. Freedman, C. D. Platsoucas, S. Rashed, and Y. P. Kim, "Cytotoxic T cell clones isolated from ovarian tumor-infiltrating lymphocytes recognize multiple antigenic epitopes on autologous tumor cells," *J. Immunol.*, vol. 146, pp. 1700–1707, Mar 1991.
- [37] W.-T. Hwang, S. F. Adams, E. Tahirovic, I. S. Hagemann, and G. Coukos, "Prognostic Significance of Tumor-infiltrating T-cells in Ovarian Cancer: a Meta-analysis," *Gynecologic Oncology*, vol. 124, p. 192, Feb 2012.
- [38] J. Galon, A. Costes, F. Sanchez-Cabo, A. Kirilovsky, B. Mlecnik, C. Lagorce-Pagès, M. Tosolini, M. Camus, A. Berger, P. Wind, *et al.*, "Type, density, and location of immune cells within human colorectal tumors predict clinical outcome," *Science*, vol. 313, pp. 1960–1964, Sep 2006.
- [39] J. Galon, B. Mlecnik, G. Bindea, H. K. Angell, A. Berger, C. Lagorce, A. Lugli, I. Zlobec, A. Hartmann, C. Bifulco, *et al.*, "Towards the introduction of the 'Immunoscore' in the classification of malignant tumours," *J. Pathol.*, vol. 232, p. 199, Jan 2014.
- [40] A. Feichtenbeiner, M. Haas, M. Büttner, G. G. Grabenbauer, R. Fietkau, and L. V. Distel, "Critical role of spatial interaction between χ n-CD8-wz0a and χ n-Foxp3-qd4b cells in human gastric cancer: the distance matters," *Cancer Immunol. Immunother.*, vol. 63, pp. 111–119, Feb 2014.
- [41] S. Barua, P. Fang, A. Sharma, J. Fujimoto, I. Wistuba, A. U. K. Rao, and S. H. Lin, "Spatial Interaction of Tumor Cells and Regulatory T cells Correlates with Survival in Non-Small Cell Lung Cancer," *Lung cancer (Amsterdam, Netherlands)*, vol. 117, p. 73, Mar 2018.
- [42] M. Yang, D. McKay, J. W. Pollard, and C. E. Lewis, "Diverse Functions of Macrophages in Different Tumor Microenvironments," *Cancer Res.*, vol. 78, pp. 5492–5503, Oct 2018.
- [43] Z. Yafei, G. Jun, and G. Guolan, "Correlation between macrophage infiltration and prognosis of ovarian cancer-a preliminary study.," *Biomed. Res.*, vol. 27, no. 2, 2016.
- [44] Y. Tanaka, H. Kobayashi, M. Suzuki, N. Kanayama, M. Suzuki, and T. Terao, "Upregulation of bikunin in tumor-infiltrating macrophages as a factor of favorable prognosis in ovarian cancer," *Gynecol. Oncol.*, vol. 94, pp. 725–734, Sep 2004.
- [45] Y.-K. Huang, M. Wang, Y. Sun, N. Di Costanzo, C. Mitchell, A. Achuthan, J. A. Hamilton, R. A. Busuttil, and A. Boussioutas, "Macrophage spatial heterogeneity in gastric cancer defined by multiplex immunohistochemistry," *Nat. Commun.*, vol. 10, 2019.

- [46] A. Kythreotou, A. Siddique, F. A. Mauri, M. Bower, and D. J. Pinato, "PD-L1," *J. Clin. Pathol.*, vol. 71, pp. 189–194, Mar 2018.
- [47] S. R. Gordon, R. L. Maute, B. W. Dulken, G. Hutter, B. M. George, M. N. McCracken, R. Gupta, J. M. Tsai, R. Sinha, D. Corey, *et al.*, "PD-1 expression by tumor-associated macrophages inhibits phagocytosis and tumor immunity," *Nature*, vol. 545, p. 495, May 2017.
- [48] M. L. Drakes, S. Mehrotra, M. Aldulescu, R. K. Potkul, Y. Liu, A. Grisoli, C. Joyce, T. E. O'Brien, M. S. Stack, and P. J. Stiff, "Stratification of ovarian tumor pathology by expression of programmed cell death-1 (PD-1) and PD-ligand- 1 (PD-L1) in ovarian cancer," *J. Ovarian Res.*, vol. 11, p. 43., May 2018.
- [49] J. Ma, B. Zheng, S. Goswami, L. Meng, D. Zhang, C. Cao, T. Li, F. Zhu, L. Ma, Z. Zhang, *et al.*, "PD1 Hi CD8 + T cells correlate with exhausted signature and poor clinical outcome in hepatocellular carcinoma," *J. ImmunoTher. Cancer*, vol. 7, pp. 1–15, Dec 2019.
- [50] C. Bonnans, J. Chou, and Z. Werb, "Remodelling the extracellular matrix in development and disease," *Nat. Rev. Mol. Cell Biol.*, vol. 15, pp. 786–801, Dec 2014.
- [51] P. Lu, V. M. Weaver, and Z. Werb, "The extracellular matrix: a dynamic niche in cancer progression," *J. Cell Biol.*, vol. 196, pp. 395–406, Feb 2012.
- [52] C. A. Sherman-Baust, A. T. Weeraratna, L. B. A. Rangel, E. S. Pizer, K. R. Cho, D. R. Schwartz, T. Shock, and P. J. Morin, "Remodeling of the extracellular matrix through overexpression of collagen VI contributes to cisplatin resistance in ovarian cancer cells," *Cancer Cell*, vol. 3, pp. 377–386, Apr 2003.
- [53] F. Kai, H. Laklai, and V. M. Weaver, "Force Matters: Biomechanical Regulation of Cell Invasion and Migration in Disease," *Trends Cell Biol.*, vol. 26, pp. 486–497, Jul 2016.
- [54] L. G. Vincent, Y. S. Choi, B. Alonso-Latorre, J. C. del Álamo, and A. J. Engler, "Mesenchymal stem cell durotaxis depends on substrate stiffness gradient strength," *Biotechnol. J.*, vol. 8, pp. 472–484, Apr 2013.
- [55] A. Pathak and S. Kumar, "Independent regulation of tumor cell migration by matrix stiffness and confinement," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 109, p. 10334, Jun 2012.
- [56] K. R. Levental, H. Yu, L. Kass, J. N. Lakins, M. Egeblad, J. T. Erler, S. F. T. Fong, K. Csiszar, A. Giaccia, W. Weninger, *et al.*, "Matrix Crosslinking Forces Tumor Progression by Enhancing Integrin signaling," *Cell*, vol. 139, p. 891, Nov 2009.
- [57] A.-M. Baker, T. R. Cox, D. Bird, G. Lang, G. I. Murray, X.-F. Sun, S. M. Southall, J. R. Wilson, and J. T. Erler, "The role of lysyl oxidase in SRC-dependent proliferation and metastasis of colorectal cancer," *J. Natl. Cancer Inst.*, vol. 103, pp. 407–424, Mar 2011.

- [58] M. W. Pickup, H. Laklai, I. Acerbi, P. Owens, A. E. Gorska, A. Chytil, M. Aakre, V. M. Weaver, and H. L. Moses, “Stromally Derived Lysyl Oxidase Promotes Metastasis of Transforming Growth Factor- β Deficient Mouse Mammary Carcinomas,” *Cancer research*, vol. 73, p. 5336, Sep 2013.
- [59] B. L. Wen, M. A. Brewer, O. Nadiarnykh, J. D. Hocker, V. Singh, T. R. Mackie, and P. J. Campagnola, “Texture analysis applied to second harmonic generation image data for ovarian cancer classification,” *J. Biomed. Opt.*, vol. 19, p. 096007, Sep 2014.
- [60] M. De Donato, M. Petrillo, E. Martinelli, F. Filippetti, G. F. Zannoni, G. Scambia, and D. Gallo, “Uncovering the role of nuclear Lysyl oxidase (LOX) in advanced high grade serous ovarian cancer,” *Gynecol. Oncol.*, vol. 146, pp. 170–178, Jul 2017.
- [61] D.-J. Cheon, Y. Tong, M.-S. Sim, J. Dering, D. Berel, X. Cui, J. Lester, J. A. Beach, M. Tighiouart, A. E. Walts, B. Y. Karlan, and S. Orsulic, “A collagen-remodeling gene signature regulated by TGF- β signaling is associated with metastasis and poor survival in serous ovarian cancer,” *Clin. Cancer Res.*, vol. 20, pp. 711–723, Feb 2014.
- [62] H. A. Kenny, S. Kaur, L. M. Coussens, and E. Lengyel, “The initial steps of ovarian cancer cell metastasis are mediated by MMP-2 cleavage of vitronectin and fibronectin,” *J. Clin. Invest.*, vol. 118, pp. 1367–1379, Apr 2008.
- [63] S. W. Hayward, Y. Wang, M. Cao, Y. K. Hom, B. Zhang, G. D. Grossfeld, D. Sudilovsky, and G. R. Cunha, “Malignant transformation in a nontumorigenic human prostatic epithelial cell line,” *Cancer Res.*, vol. 61, pp. 8135–8142, Nov 2001.
- [64] B. Erdogan and D. J. Webb, “Cancer-associated fibroblasts modulate growth factor signaling and extracellular matrix remodeling to regulate tumor metastasis,” *Biochem. Soc. Trans.*, vol. 45, pp. 229–236, Feb 2017.
- [65] C. J. Whatcott, C. H. Diep, P. Jiang, A. Watanabe, J. LoBello, C. Sima, G. Hostetter, H. M. Shepard, D. D. Von Hoff, and H. Han, “Desmoplasia in Primary Tumors and Metastatic Lesions of Pancreatic Cancer,” *Clin. Cancer Res.*, vol. 21, pp. 3561–3568, Aug 2015.
- [66] R. Januchowski, M. Świerczewska, K. Sterzyńska, K. Wojtowicz, M. Nowicki, and M. Zabel, “Increased Expression of Several Collagen Genes is Associated with Drug Resistance in Ovarian Cancer Cell Lines,” *J. Cancer*, vol. 7, pp. 1295–1310, Jun 2016.
- [67] S. Xu, H. Xu, W. Wang, S. Li, H. Li, T. Li, W. Zhang, X. Yu, and L. Liu, “The role of collagen in cancer: from bench to bedside,” *J. Transl. Med.*, vol. 17, pp. 1–22, Dec 2019.
- [68] B. Jakubzig, F. Baltes, S. Henze, M. Schlesinger, and G. Bendas, “Mechanisms of Matrix-Induced Chemoresistance of Breast Cancer Cells—Deciphering Novel Potential Targets for a Cell Sensitization,” *Cancers*, vol. 10, Dec 2018.
- [69] K. Shield, C. Riley, M. A. Quinn, G. E. Rice, M. L. Ackland, and N. Ahmed, “ $\alpha 2\beta 1$ integrin affects metastatic potential of ovarian carcinoma spheroids by supporting disaggregation and proteolysis,” *J. Carcinog.*, vol. 6, p. 11, 2007.

- [70] K. Sawada, A. K. Mitra, A. R. Radjabi, V. Bhaskar, E. O. Kistner, M. Tretiakova, S. Jagadeeswaran, A. Montag, A. Becker, H. A. Kenny, M. E. Peter, V. Ramakrishnan, S. D. Yamada, and E. Lengyel, "Loss of E-cadherin promotes ovarian cancer metastasis via alpha 5-integrin, which is a therapeutic target," *Cancer Res.*, vol. 68, pp. 2329–2339, Apr 2008.
- [71] S. Natarajan, K. M. Foreman, M. I. Soriano, N. S. Rossen, H. Shehade, D. R. Fregoso, J. T. Eggold, V. Krishnan, O. Dorigo, A. J. Krieg, *et al.*, "Collagen remodeling in the hypoxic tumor-mesothelial niche promotes ovarian cancer metastasis," *Cancer Res.*, Jan 2019.
- [72] Y. Zhao, J. Cao, A. Melamed, M. Worley, A. Gockley, D. Jones, H. T. Nia, Y. Zhang, T. Stylianopoulos, A. S. Kumar, *et al.*, "Losartan treatment enhances chemotherapy efficacy and reduces ascites in ovarian cancer models by normalizing the tumor stroma," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 116, pp. 2210–2219, Feb 2019.
- [73] H. Bougherara, A. Mansuet-Lupo, M. Alifano, C. Ngô, D. Damotte, M.-A. Le Frère-Belda, E. Donnadieu, and E. Peranzoni, "Real-Time Imaging of Resident T Cells in Human Lung and Ovarian Carcinomas Reveals How Different Tumor Microenvironments Control T Lymphocyte Migration," *Front. Immunol.*, vol. 6, p. 500, Oct 2015.
- [74] H. C. Pruitt, D. Lewis, M. Ciccaglione, S. Connor, Q. Smith, J. W. Hickey, J. P. Schneck, and S. Gerecht, "Collagen fiber structure guides 3D motility of cytotoxic T lymphocytes," *Matrix Biol.*, Feb 2019.
- [75] R. R. da Cunha Colombo Bonadio, R. N. Fogace, V. C. Miranda, and M. del Pilar Estevez Diz, "Homologous recombination deficiency in ovarian cancer: a review of its epidemiology and management," *Clinics*, vol. 73, no. Suppl 1, 2018.
- [76] Q. Chen, L. Sun, and Z. J. Chen, "Regulation and function of the cGAS–STING pathway of cytosolic DNA sensing," *Nat. Immunol.*, vol. 17, pp. 1142–1149, Sep 2016.
- [77] E. E. Parkes, S. M. Walker, L. E. Taggart, N. McCabe, L. A. Knight, R. Wilkinson, K. D. McCloskey, N. E. Buckley, K. I. Savage, M. Salto-Tellez, *et al.*, "Activation of STING-Dependent Innate Immune Signaling By S-Phase-Specific DNA Damage in Breast Cancer," *J. Natl. Cancer Inst.*, vol. 109, p. djw199., Oct 2016.
- [78] L. Ding, H.-J. Kim, Q. Wang, M. Kearns, T. Jiang, C. E. Ohlson, B. B. Li, S. Xie, J. F. Liu, E. H. Stover, *et al.*, "PARP Inhibition Elicits STING-Dependent Antitumor Immunity in Brca1-Deficient Ovarian Cancer," *Cell Rep.*, vol. 25, pp. 2972–29805, Dec 2018.
- [79] E. Brown, T. McKee, E. DiTomaso, A. Pluen, B. Seed, Y. Boucher, and R. K. Jain, "Dynamic imaging of collagen and its modulation in tumors in vivo using second-harmonic generation," *Nat. Med.*, vol. 9, pp. 796–800, Jun 2003.
- [80] R. A. Soslow, G. Han, K. J. Park, K. Garg, N. Olvera, D. R. Spriggs, N. D. Kauff, and D. A. Levine, "Morphologic patterns associated with BRCA1 and BRCA2 genotype in ovarian carcinoma," *Mod. Pathol.*, vol. 25, pp. 625–636, Dec 2011.

- [81] J. N. McAlpine, H. Porter, M. Köbel, B. H. Nelson, L. M. Prentice, S. E. Kalloger, J. Senz, K. Milne, J. Ding, S. P. Shah, *et al.*, “BRCA1 and BRCA2 mutations correlate with TP53 abnormalities and presence of immune cell infiltrates in ovarian high-grade serous carcinoma,” *Mod. Pathol.*, vol. 25, pp. 740–750, Jan 2012.
- [82] J. N. McAlpine, H. Porter, M. Köbel, B. H. Nelson, L. M. Prentice, S. E. Kalloger, J. Senz, K. Milne, J. Ding, S. P. Shah, *et al.*, “BRCA1 and BRCA2 mutations correlate with TP53 abnormalities and presence of immune cell infiltrates in ovarian high-grade serous carcinoma,” *Mod. Pathol.*, vol. 25, pp. 740–750, Jan 2012.
- [83] H. Failmezger, S. Muralidhar, A. Rullan, C. E. de Andrea, E. Sahai, and Y. Yuan, “Topological Tumor Graphs: a graph-based spatial model to infer stromal recruitment for immunosuppression in melanoma histology,” *Cancer Res.*, Jan 2019.
- [84] L. Sorokin, “The impact of the extracellular matrix on inflammation,” *Nat. Rev. Immunol.*, vol. 10, pp. 712–723, Oct 2010.
- [85] S. Lyford-Pike, S. Peng, G. D. Young, J. M. Taube, W. H. Westra, B. Akpeng, T. C. Bruno, J. D. Richmon, H. Wang, J. A. Bishop, *et al.*, “Evidence for a role of the PD-1:PD-L1 pathway in immune resistance of HPV-associated head and neck squamous cell carcinoma,” *Cancer Res.*, vol. 73, pp. 1733–1741, Mar 2013.
- [86] T. C. G. A. R. Network, “Integrated Genomic Analyses of Ovarian Carcinoma,” *Nature*, vol. 474, p. 609, Jun 2011.
- [87] F. C. Martins, I. de Santiago, A. Trinh, J. Xian, A. Guo, K. Sayal, M. Jimenez-Linan, S. Deen, K. Driver, M. Mack, *et al.*, “Combined image and genomic analysis of high-grade serous ovarian cancer reveals PTEN loss as a common driver event and prognostic classifier,” *Genome Biol.*, vol. 15, pp. 1–15, Dec 2014.
- [88] S. Al-Khalidi, *Genomic aberrations as determinants of immune infiltrates in high grade serous ovarian carcinoma*. PhD thesis, University of Cambridge, 2020.
- [89] Y. K. Wang, A. Bashashati, M. S. Anglesio, D. R. Cochrane, D. S. Grewal, G. Ha, A. McPherson, H. M. Horlings, J. Senz, L. M. Prentice, *et al.*, “Genomic consequences of aberrant DNA repair mechanisms stratify ovarian cancer histotypes,” *Nat. Genet.*, vol. 49, pp. 856–865, Apr 2017.
- [90] K. Yoshihara, T. Tsunoda, D. Shigemizu, H. Fujiwara, M. Hatae, H. Fujiwara, H. Masuzaki, H. Katabuchi, Y. Kawakami, A. Okamoto, *et al.*, “High-risk ovarian cancer based on 126-gene expression signature is uniquely characterized by downregulation of antigen presentation pathway,” *Clin. Cancer Res.*, vol. 18, pp. 1374–1385, Mar 2012.
- [91] R. W. Tothill, A. V. Tinker, J. George, R. Brown, S. B. Fox, S. Lade, D. S. Johnson, M. K. Trivett, D. Etemadmoghadam, B. Locandro, *et al.*, “Novel molecular subtypes of serous and endometrioid ovarian cancer linked to clinical outcome,” *Clin. Cancer Res.*, vol. 14, pp. 5198–5208, Aug 2008.

- [92] S. Kommoss, B. Winterhoff, A. L. Oberg, G. E. Konecny, C. Wang, S. M. Riska, J.-B. Fan, M. J. Maurer, C. April, V. Shridhar, *et al.*, “Bevacizumab May Differentially Improve Ovarian Cancer Outcome in Patients with Proliferative and Mesenchymal Molecular Subtypes,” *Clin. Cancer Res.*, vol. 23, pp. 3794–3801, Jul 2017.
- [93] Y. Shimizu, S. Kamoi, S. Amada, K. Hasumi, F. Akiyama, and S. G. Silverberg, “Toward the Development of a Universal Grading System for Ovarian Epithelial Carcinoma: I. Prognostic Significance of Histopathologic Features—Problems Involved in the Architectural Grading System,” *Gynecol. Oncol.*, vol. 70, pp. 2–12, Jul 1998.
- [94] E. G. Silva and D. M. Gershenson, “Standardized Histologic Grading of Epithelial Ovarian Cancer: Elusive after All These Years,” *Gynecol. Oncol.*, vol. 70, p. 1, Jul 1998.
- [95] Y. R. Hussein, J. A. Ducie, A. G. Arnold, N. D. Kauff, H. A. Vargas-Alvarez, E. Sala, D. A. Levine, and R. A. Soslow, “Invasion Patterns of Metastatic Extrauterine High Grade Serous Carcinoma with BRCA Germline Mutation and Correlation with Clinical Outcomes,” *Am. J. Surg. Pathol.*, vol. 40, p. 404, Mar 2016.
- [96] A. Ayhan, R. J. Kurman, A. Yemelyanova, R. Vang, S. Logani, J. D. Seidman, and I.-M. Shih, “Defining the cut point between low-grade and high-grade ovarian serous carcinomas: a clinicopathologic and molecular genetic analysis,” *Am. J. Surg. Pathol.*, vol. 33, pp. 1220–1224, Aug 2009.
- [97] R. Murakami, N. Matsumura, M. Mandai, K. Yoshihara, H. Tanabe, H. Nakai, K. Yamanoi, K. Abiko, Y. Yoshioka, J. Hamanishi, *et al.*, “Establishment of a Novel Histopathological Classification of High-Grade Serous Ovarian Carcinoma Correlated with Prognostically Distinct Gene Expression Subtypes,” *Am. J. Pathol.*, vol. 186, pp. 1103–1113, May 2016.
- [98] M. Schwede, L. Waldron, S. C. Mok, W. Wei, A. Basunia, M. A. Merritt, C. S. Mitsiades, G. Parmigiani, D. P. Harrington, J. Quackenbush, *et al.*, “The Impact of Stroma Admixture on Molecular Subtypes and Prognostic Gene Signatures in Serous Ovarian Cancer,” *Cancer Epidemiol. Biomarkers Prev.*, Dec 2019.
- [99] D. S. Chen and I. Mellman, “Elements of cancer immunity and the cancer-immune set point,” *Nature*, vol. 541, pp. 321–330, Jan 2017.
- [100] T. Mandelkow, N. C. Blessin, E. Lueerss, L. Pott, R. Simon, W. Li, B. Wellge, N. F. Debatin, D. Höflmayer, J. R. Izbicki, *et al.*, “Immune Exclusion Is Frequent in Small-Cell Carcinoma of the Bladder,” *Dis. Markers*, vol. 2019, p. 2532518, May 2019.
- [101] A. W. Zhang, A. McPherson, K. Milne, D. R. Kroeger, P. T. Hamilton, A. Miranda, T. Funnell, N. Little, C. P. E. de Souza, S. Laan, *et al.*, “Interfaces of Malignant and Immunologic Clonal Dynamics in Ovarian Cancer,” *Cell*, vol. 173, pp. 1755–1769.e22, Jun 2018.
- [102] J. N. Kather, M. Suarez-Carmona, P. Charoentong, C.-A. Weis, D. Hirsch, P. Bankhead, M. Horning, D. Ferber, I. Kel, E. Herpel, S. Schott, I. Zörnig, J. Utikal, A. Marx, T. Gaiser, H. Brenner, J. Chang-Claude, M. Hoffmeister, D. Jäger, and N. Halama, “Topography of cancer-associated immune cells in human solid tumors,” *eLife*, vol. 7, 2018.

- [103] A. Jiménez-Sánchez, D. Memon, S. Pourpe, H. Veeraraghavan, Y. Li, H. A. Vargas, M. B. Gill, K. J. Park, O. Zivanovic, J. Konner, *et al.*, “Heterogeneous Tumor-Immune Microenvironments among Differentially Growing Metastases in an Ovarian Cancer Patient,” *Cell*, vol. 170, pp. 927–93820, Aug 2017.
- [104] M. Tuefferd, A. De Bondt, I. Van Den Wyngaert, W. Talloen, T. Verbeke, B. Carvalho, D.-A. Clevert, M. Alifano, N. Raghavan, D. Amaratunga, *et al.*, “Genome-wide copy number alterations detection in fresh frozen and matched FFPE samples using SNP 6.0 arrays,” *Genes Chromosom. Cancer*, vol. 47, pp. 957–964, Nov 2008.
- [105] N. C. M. Visser, A. A. M. van der Wurff, J. M. A. Pijnenborg, L. F. A. G. Massuger, J. Bulten, and I. D. Nagtegaal, “Tissue microarray is suitable for scientific biomarkers studies in endometrial cancer,” *Virchows Arch.*, vol. 472, pp. 407–413, Mar 2018.
- [106] F. P. Fonseca, B. A. B. de Andrade, A. L. C. A. Rangel, R. D. Coletta, M. A. Lopes, O. P. de Almeida, and P. A. Vargas, “Tissue microarray is a reliable method for immunohistochemical analysis of pleomorphic adenoma,” *Oral Surg. Oral Med. Oral Pathol. Oral Radiol.*, vol. 117, pp. 81–88, Jan 2014.
- [107] M. H. Khouja, M. Baekelandt, A. Sarab, J. M. Nesland, and R. Holm, “Limitations of tissue microarrays compared with whole tissue sections in survival analysis,” *Oncology Letters*, vol. 1, p. 827, Sep 2010.
- [108] J. Permeth-Wey, D. Boulware, N. Valkov, S. Livingston, S. Nicosia, J.-H. Lee, R. Sutphen, J. Schildkraut, S. Narod, A. Parker, *et al.*, “Sampling Strategies for Tissue Microarrays to evaluate biomarkers in Ovarian Cancer,” *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, vol. 18, p. 28, Jan 2009.
- [109] A. H. Fischer, K. A. Jacobson, J. Rose, and R. Zeller, “Hematoxylin and eosin staining of tissue and cell sections,” *CSH Protoc.*, vol. 2008, p. pdb.prot4986, May 2008.
- [110] O. P. Kallioniemi, U. Wagner, J. Kononen, and G. Sauter, “Tissue microarray technology for high-throughput molecular profiling of cancer,” *Hum. Mol. Genet.*, vol. 10, pp. 657–662, Apr 2001.
- [111] M. Rakaee, T. K. Kilvaer, S. M. Dalen, E. Richardsen, E.-E. Paulsen, S. M. Hald, S. Al-Saad, S. Andersen, T. Donnem, R. M. Bremnes, *et al.*, “Evaluation of tumor-infiltrating lymphocytes using routine H&E slides predicts patient survival in resected non-small cell lung cancer,” *Hum. Pathol.*, vol. 79, pp. 188–198, Sep 2018.
- [112] A. Yemelyanova, R. Vang, M. Kshirsagar, D. Lu, M. A. Marks, I. M. Shih, and R. J. Kurman, “Immunohistochemical staining patterns of p53 can serve as a surrogate marker for TP53 mutations in ovarian carcinoma: an immunohistochemical and nucleotide sequencing analysis,” *Mod. Pathol.*, vol. 24, pp. 1248–1253, May 2011.
- [113] R. Y. Tsien and A. Waggoner, *Fluorophores for Confocal Microscopy*. Springer, Boston, MA, 1995.

- [114] C. Giesen, H. A. O. Wang, D. Schapiro, N. Zivanovic, A. Jacobs, B. Hattendorf, P. J. Schüffler, D. Grolimund, J. M. Buhmann, S. Brandt, *et al.*, “Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry,” *Nat. Methods*, vol. 11, pp. 417–422, Mar 2014.
- [115] R. Remark, T. Merghoub, N. Grabe, G. Litjens, D. Damotte, J. D. Wolchok, M. Merad, and S. Gnjatic, “In-depth tissue profiling using multiplexed immunohistochemical consecutive staining on single slide,” *Sci. Immunol.*, vol. 1, p. aaf6925., Jul 2016.
- [116] E. R. Parra, A. Francisco-Cruz, and I. I. Wistuba, “State-of-the-Art of Profiling Immune Contexture in the Era of Multiplexed Staining and Digital Analysis to Study Paraffin Tumor Tissues,” *Cancers*, vol. 11, p. 247, Feb 2019.
- [117] R. Hellwarth and P. Christensen, “Nonlinear optical microscopic examination of structure in polycrystalline ZnSe,” *Opt. Commun.*, vol. 12, pp. 318–322, Nov 1974.
- [118] A. D. Hofemeier, H. Hachmeister, C. Pilger, M. Schürmann, J. F. W. Greiner, L. Nolte, H. Sudhoff, C. Kaltschmidt, T. Huser, and B. Kaltschmidt, “Label-free nonlinear optical microscopy detects early markers for osteogenic differentiation of human stem cells,” *Sci. Rep.*, vol. 6, pp. 1–10, May 2016.
- [119] F. Fereidouni, A. Todd, Y. Li, C.-W. Chang, K. Luong, A. Rosenberg, Y.-J. Lee, J. W. Chan, A. Borowsky, K. Matsukuma, *et al.*, “Dual-mode emission and transmission microscopy for virtual histochemistry using hematoxylin- and eosin-stained tissue sections,” *Biomed. Opt. Express*, vol. 10, p. 6516, Dec 2019.
- [120] N. Vuillemin, P. Mahou, D. Débarre, T. Gacoin, P.-L. Tharoux, M.-C. Schanne-Klein, W. Supatto, and E. Beaurepaire, “Efficient second-harmonic imaging of collagen in histological slides using Bessel beam excitation,” *Sci. Rep.*, vol. 6, pp. 1–13, Jul 2016.
- [121] K. R. Campbell, R. Chaudhary, J. M. Handel, M. S. Patankar, and P. J. Campagnola, “Polarization-resolved second harmonic generation imaging of human ovarian cancer,” *J. Biomed. Opt.*, vol. 23, pp. 1–8, Jun 2018.
- [122] K. R. Campbell and P. J. Campagnola, “Assessing local stromal alterations in human ovarian cancer subtypes via second harmonic generation microscopy and analysis,” *J. Biomed. Opt.*, vol. 22, pp. 1–7, Nov 2017.
- [123] K. R. Campbell, B. Wen, E. M. Shelton, R. Swader, B. L. Cox, K. Eliceiri, and P. J. Campagnola, “3D second harmonic generation imaging tomography by multi-view excitation,” *Optica*, vol. 4, pp. 1171–1179, Oct 2017.
- [124] O. Nadiarnykh, R. B. LaComb, M. A. Brewer, and P. J. Campagnola, “Alterations of the extracellular matrix in ovarian cancer studied by Second Harmonic Generation imaging microscopy,” *BMC Cancer*, vol. 10, pp. 1–14, Dec 2010.
- [125] T. Vicar, J. Balvan, J. Jaros, F. Jug, R. Kolar, M. Masarik, and J. Gumulec, “Cell segmentation methods for label-free contrast microscopy: review and comprehensive comparison,” *BMC Bioinf.*, vol. 20, pp. 1–25, Dec 2019.

- [126] M. Veta, P. J. van Diest, R. Kornegoor, A. Huisman, M. A. Viergever, and J. P. W. Pluim, "Automatic Nuclei Segmentation in H&E Stained Breast Cancer Histopathology Images," *PLoS One*, vol. 8, p. e70221, Jul 2013.
- [127] M. Veta, J. P. W. Pluim, P. J. van Diest, and M. A. Viergever, "Breast cancer histopathology image analysis: a review," *IEEE Trans. Biomed. Eng.*, vol. 61, pp. 1400–1411, May 2014.
- [128] M. Macenko, M. Niethammer, J. S. Marron, D. Borland, J. T. Woosley, X. Guan, C. Schmitt, and N. E. Thomas, "A method for normalizing histology slides for quantitative analysis," *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 1107–1110, Jun 2009.
- [129] P. Kleczek, J. Jaworek-Korjakowska, and M. Gorgon, "A novel method for tissue segmentation in high-resolution H&E-stained histopathological whole-slide images.," *Comput. Med. Imaging Graph.*, vol. 79, p. 101686, Nov 2019.
- [130] F. S. Nielsen, M. J. Pedersen, M. V. Olsen, M. S. Larsen, R. Røge, and A. S. Jørgensen, "Automatic Bone Marrow Cellularity Estimation in H&E Stained Whole Slide Images," *Cytometry*, vol. 95, pp. 1066–1074, Oct 2019.
- [131] H. D. Couture, L. A. Williams, J. Geradts, S. J. Nyante, E. N. Butler, J. S. Marron, C. M. Perou, M. A. Troester, and M. Niethammer, "Image analysis with deep learning to predict breast cancer grade, ER status, histologic subtype, and intrinsic subtype," *npj Breast Cancer*, vol. 4, pp. 1–8, Sep 2018.
- [132] P. Bankhead, M. B. Loughrey, J. A. Fernández, Y. Dombrowski, D. G. McArt, P. D. Dunne, S. McQuaid, R. T. Gray, L. J. Murray, H. G. Coleman, *et al.*, "QuPath: Open source software for digital pathology image analysis," *Sci. Rep.*, vol. 7, pp. 1–7, Dec 2017.
- [133] J. L. Carstens, P. Correa de Sampaio, D. Yang, S. Barua, H. Wang, A. Rao, J. P. Allison, V. S. LeBleu, and R. Kalluri, "Spatial computation of intratumoral T cells correlates with survival of patients with pancreatic cancer," *Nat. Commun.*, vol. 8, pp. 1–13, Apr 2017.
- [134] H. R. Ali, A. Dariush, J. Thomas, E. Provenzano, J. Dunn, L. Hiller, A.-L. Vallier, J. Abraham, T. Piper, J. M. S. Bartlett, *et al.*, "Lymphocyte density determined by computational pathology validated as a predictor of response to neoadjuvant chemotherapy in breast cancer: secondary analysis of the ARTemis trial," *Ann. Oncol.*, vol. 28, pp. 1832–1835, Aug 2017.
- [135] D. K. Wells, Y. Chuang, L. M. Knapp, D. Brockmann, W. L. Kath, and J. N. Leonard, "Spatial and Functional Heterogeneities Shape Collective Behavior of Tumor-Immune Networks," *PLoS Comput. Biol.*, vol. 11, Apr 2015.
- [136] S. Nawaz, A. Heindl, K. Koelble, and Y. Yuan, "Beyond immune density: critical role of spatial heterogeneity in estrogen receptor-negative breast cancer," *Mod. Pathol.*, vol. 28, pp. 766–777, Jun 2015.

- [137] Y. Yuan, H. Failmezger, O. M. Rueda, H. R. Ali, S. Gräf, S.-F. Chin, R. F. Schwarz, C. Curtis, M. J. Dunning, H. Bardwell, *et al.*, “Quantitative Image Analysis of Cellular Heterogeneity in Breast Tumors Complements Genomic Profiling,” *Sci. Transl. Med.*, vol. 4, p. 157ra143, Oct 2012.
- [138] C. Lan, A. Heindl, X. Huang, S. Xi, S. Banerjee, J. Liu, and Y. Yuan, “Quantitative histology analysis of the ovarian tumour microenvironment,” *Sci. Rep.*, vol. 5, pp. 1–12, Nov 2015.
- [139] A. Heindl, C. Lan, D. N. Rodrigues, K. Koelble, and Y. Yuan, “Similarity and diversity of the tumor microenvironment in multiple metastases: critical implications for overall and progression-free survival of high-grade serous ovarian cancer,” *Oncotarget*, vol. 7, p. 71123, Nov 2016.
- [140] D. Massi, E. Rulli, M. Cossa, B. Valeri, M. Rodolfo, B. Merelli, F. De Logu, R. Nassini, M. Del Vecchio, L. Di Guardo, *et al.*, “The density and spatial tissue distribution of CD8 + and CD163 + immune cells predict response and outcome in melanoma patients receiving MAPK inhibitors,” *J. ImmunoTher. Cancer*, vol. 7, pp. 1–13, Dec 2019.
- [141] Anglian Breast Cancer Study Group, “Prevalence and penetrance of BRCA1 and BRCA2 mutations in a population-based series of breast cancer cases,” *Br. J. Cancer*, vol. 83, p. 1301, Nov 2000.
- [142] T. Goranova, D. Ennis, A. M. Piskorz, G. Macintyre, L. A. Lewsley, J. Stobo, C. Wilson, D. Kay, R. M. Glasspool, M. Lockley, *et al.*, “Safety and utility of image-guided research biopsies in relapsed high-grade serous ovarian carcinoma—experience of the BriTROC consortium,” *Br. J. Cancer*, vol. 116, pp. 1294–1301, Mar 2017.
- [143] C. A. Parkinson, D. Gale, A. M. Piskorz, H. Biggs, C. Hodgkin, H. Addley, S. Freeman, P. Moyle, E. Sala, K. Sayal, *et al.*, “Exploratory Analysis of TP53 Mutations in Circulating Tumour DNA as Biomarkers of Treatment Response for Patients with Relapsed High-Grade Serous Ovarian Carcinoma: A Retrospective Study,” *PLoS Med.*, vol. 13, p. e1002198, Dec 2016.
- [144] “ISRCTN - ISRCTN91273375: A randomised, two-arm, multicentre Gynaecologic Cancer InterGroup trial of adding bevacizumab to standard chemotherapy (carboplatin and paclitaxel) in patients with epithelial ovarian cancer,” Aug 2020. [<http://www.isrctn.com/ISRCTN91273375>; accessed 2. Aug. 2020].
- [145] A. M. Oza, A. D. Cook, J. Pfisterer, A. Embleton, J. A. Ledermann, E. Pujade-Lauraine, G. Kristensen, M. S. Carey, P. Beale, A. Cervantes, *et al.*, “Standard chemotherapy with or without bevacizumab for women with newly diagnosed ovarian cancer (ICON7): overall survival results of a phase 3 randomised trial,” *Lancet Oncol.*, vol. 16, p. 928, Aug 2015.
- [146] E. Schubert, J. Sander, M. Ester, H. P. Kriegel, and X. Xu, “DBSCAN Revisited, Revisited: Why and How You Should (Still) Use DBSCAN,” *ACM Trans. Database Syst.*, vol. 42, pp. 1–21, Jul 2017.
- [147] M. Batty, “Spatial Entropy,” *Geog. Anal.*, vol. 6, pp. 1–31, Jan 1974.

- [148] . Groundai, “SpatEntropy: Spatial Entropy Measures in R,” *GroundAI*, Apr 2018.
- [149] S.-C. Lam, “Texture feature extraction using gray level gradient based co-occurrence matrices,” vol. 1, pp. 267–271, IEEE, Oct 1996.
- [150] R. M. Haralick, K. Shanmugam, and I. Dinstein, “Textural Features for Image Classification,” *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-3, pp. 610–621, Nov 1973.
- [151] J. Ho, S. M. Ahlers, C. Stratman, O. Aridor, L. Pantanowitz, J. L. Fine, J. A. Kuzmishin, M. C. Montalto, and A. V. Parwani, “Can digital pathology result in cost savings? A financial projection for digital pathology implementation at a large integrated health care organization,” *J. Pathol. Inform.*, vol. 5, p. 33., Aug 2014.
- [152] H. Song, S. J. Ramus, L. Quaye, R. A. DiCioccio, J. Tyrer, E. Lomas, D. Shadforth, E. Hogdall, C. Hogdall, V. McGuire, *et al.*, “Common variants in mismatch repair genes and risk of invasive ovarian cancer,” *Carcinogenesis*, vol. 27, pp. 2235–2242, Nov 2006.
- [153] F. Zeppernick and I. Meinhold-Heerlein, “The new FIGO staging system for ovarian, fallopian tube, and primary peritoneal cancer,” *Arch. Gynecol. Obstet.*, vol. 290, pp. 839–842, Nov 2014.
- [154] S. Walters, C. Maringe, J. Butler, J. D. Brierley, B. Rachet, and M. P. Coleman, “Comparability of stage data in cancer registries in six countries: lessons from the International Cancer Benchmarking Partnership,” *Int. J. Cancer*, vol. 132, pp. 676–685, Feb 2013.
- [155] A. M. Piskorz, D. Ennis, G. Macintyre, T. E. Goranova, M. Eldridge, N. Segui-Gracia, M. Valganon, A. Hoyle, C. Orange, L. Moore, *et al.*, “Methanol-based fixation is superior to buffered formalin for next-generation sequencing of DNA from clinical cancer samples,” *Ann. Oncol.*, vol. 27, pp. 532–539, Mar 2016.
- [156] M. Köbel, A. M. Piskorz, S. Lee, S. Lui, C. LePage, F. Marass, N. Rosenfeld, A.-M. M. Masson, and J. D. Brenton, “Optimized p53 immunohistochemistry is an accurate predictor of TP53 mutation in ovarian carcinoma,” *J. Pathol. Clin. Res.*, vol. 2, pp. 247–258, Jul 2016.
- [157] H. Song, M. S. Cicek, E. Dicks, P. Harrington, S. J. Ramus, J. M. Cunningham, B. L. Fridley, J. P. Tyrer, J. Alsop, M. Jimenez-Linan, *et al.*, “The contribution of deleterious germline mutations in BRCA1, BRCA2 and the mismatch repair genes to ovarian cancer in the population,” *Hum. Mol. Genet.*, vol. 23, pp. 4703–4709, Sep 2014.
- [158] Y. Dong, J.-A. Richards, R. Gupta, P. P. Aung, A. Emley, Y. Kluger, S. K. Dogra, M. Mahalingam, and N. Wajapeyee, “PTEN functions as a melanoma tumor suppressor by promoting host immune response,” *Oncogene*, vol. 33, pp. 4632–4642, Sep 2014.
- [159] Y. Dai, C. Sun, Y. Feng, Q. Jia, and B. Zhu, “Potent immunogenicity in BRCA1-mutated patients with high-grade serous ovarian carcinoma,” *J. Cell. Mol. Med.*, vol. 22, pp. 3979–3986, May 2018.

- [160] M. Cordani, R. Pacchiana, G. Butera, G. D’Orazi, A. Scarpa, and M. Donadelli, “Mutant p53 proteins alter cancer cell secretome and tumour microenvironment: Involvement in cancer invasion and metastasis,” *Cancer Lett.*, vol. 376, pp. 303–309, Jul 2016.
- [161] W. Peng, J. Q. Chen, C. Liu, S. Malu, C. Creasy, M. T. Tetzlaff, C. Xu, J. A. McKenzie, C. Zhang, X. Liang, *et al.*, “Loss of PTEN Promotes Resistance to T Cell-Mediated Immunotherapy,” *Cancer Discov.*, vol. 6, pp. 202–216, Feb 2016.
- [162] M. Zhang, Y. He, X. Sun, Q. Li, W. Wang, A. Zhao, and W. Di, “A high M1/M2 ratio of tumor-associated macrophages is associated with extended survival in ovarian cancer patients,” *J. Ovarian Res.*, vol. 7, p. 19., Feb 2014.
- [163] Z. Li, D. Maeda, M. Yoshida, M. Umakoshi, H. Nanjo, K. Shiraishi, M. Saito, T. Kohno, H. Konno, H. Saito, *et al.*, “The intratumoral distribution influences the prognostic impact of CD68- and CD204-positive macrophages in non-small cell lung cancer,” *Lung Cancer*, vol. 123, pp. 127–135, Sep 2018.
- [164] Q.-w. Zhang, L. Liu, C.-y. Gong, H.-s. Shi, Y.-h. Zeng, X.-z. Wang, Y.-w. Zhao, and Y.-q. Wei, “Prognostic Significance of Tumor-Associated Macrophages in Solid Tumor: A Meta-Analysis of the Literature,” *PLoS One*, vol. 7, no. 12, 2012.
- [165] Y. Zhu, H. Fu, Z. Liu, J. Zhang, and D. Ye, “Immune-desert, immune-excluded and inflamed phenotypes predict survival and adjuvant chemotherapy response in patients with MIBC,” *European Urology Supplements*, vol. 17, pp. e128–e130, Mar 2018.
- [166] G. Kayser, J. Görtler, C. A. Weis, S. Borkenfeld, and K. Kayser, “The application of structural entropy in tissue based diagnosis,” *Diagn. Pathol.*, vol. 3, Aug 2017.
- [167] K. Milne, M. Köbel, S. E. Kalloger, R. O. Barnes, D. Gao, C. B. Gilks, P. H. Watson, and B. H. Nelson, “Systematic analysis of immune infiltrates in high-grade serous ovarian cancer reveals CD20, FoxP3 and TIA-1 as positive prognostic factors,” *PLoS One*, vol. 4, p. e6412., Jul 2009.
- [168] R. F. Schwarz, C. K. Y. Ng, S. L. Cooke, S. Newman, J. Temple, A. M. Piskorz, D. Gale, K. Sayal, M. Murtaza, P. J. Baldwin, N. Rosenfeld, H. M. Earl, E. Sala, M. Jimenez-Linan, C. A. Parkinson, F. Markowitz, and J. D. Brenton, “Spatial and Temporal Heterogeneity in High-Grade Serous Ovarian Cancer: A Phylogenetic Analysis,” *PLoS Med.*, vol. 12, p. e1001789, Feb 2015.
- [169] R. Natrajan, H. Sailem, F. K. Mardakheh, M. A. Garcia, C. J. Tape, M. Dowsett, C. Bakal, and Y. Yuan, “Microenvironmental Heterogeneity Parallels Breast Cancer Progression: A Histology–Genomic Integration Analysis,” *PLoS Med.*, vol. 13, p. e1001961, Feb 2016.
- [170] N. Rahmah and I. S. Sitanggang, “Determination of optimal epsilon (eps) value on db-scan algorithm to clustering data on peatland hotspots in sumatra,” in *IOP Conference Series: Earth and Environmental Science*, vol. 31, p. 012012, IOP Publishing, 2016.
- [171] Y. Ou, M. J. Cannon, and M. Nakagawa, “Regulatory T Cells in Gynecologic Cancer,” *MOJ immunology*, vol. 6, no. 2, p. 34, 2018.

- [172] A. Toker, L. T. Nguyen, S. C. Stone, S. Y. C. Yang, S. R. Katz, P. A. Shaw, B. A. Clarke, D. Ghazarian, A. Al-Habeeb, A. Easson, *et al.*, “Regulatory T Cells in Ovarian Cancer Are Characterized by a Highly Activated Phenotype Distinct from that in Melanoma,” *Clin. Cancer Res.*, vol. 24, pp. 5685–5696, Nov 2018.
- [173] M. Takenaka, N. Seki, U. Toh, S. Hattori, A. Kawahara, T. Yamaguchi, K. Koura, R. Takahashi, H. Otsuka, H. Takahashi, *et al.*, “FOXP3 expression in tumor cells and tumor-infiltrating lymphocytes is associated with breast cancer prognosis,” *Molecular and Clinical Oncology*, vol. 1, pp. 625–632, Jul 2013.
- [174] C. Hermans, D. Anz, J. Engel, T. Kirchner, S. Endres, and D. Mayr, “Analysis of FoxP3+ T-Regulatory Cells and CD8+T-Cells in Ovarian Carcinoma: Location and Tumor Infiltration Patterns Are Key Prognostic Markers,” *PLoS One*, vol. 9, p. e111757, Nov 2014.
- [175] J. C. Barnett, S. M. Bean, R. S. Whitaker, E. Kondoh, T. Baba, S. Fujii, J. R. Marks, H. K. Dressman, S. K. Murphy, and A. Berchuck, “Ovarian cancer tumor infiltrating T-regulatory (Treg) cells are associated with a metastatic phenotype,” *Gynecol. Oncol.*, vol. 116, pp. 556–562, Mar 2010.
- [176] T. Feigenberg, B. Clarke, C. Virtanen, A. Plotkin, M. Letarte, B. Rosen, M. Q. Bernardini, A. Kollara, T. J. Brown, and K. J. Murphy, “Molecular Profiling and Clinical Outcome of High-Grade Serous Ovarian Cancer Presenting with Low- versus High-Volume Ascites,” *Biomed Res. Int.*, vol. 2014, 2014.
- [177] Z. Yang, X. Yang, S. Xu, P. Jin, X. Li, X. Wei, D. Liu, K. Huang, S. Long, Y. Wang, C. Sun, G. Chen, J. Hu, L. Meng, D. Ma, and Q. Gao, “Reprogramming of stromal fibroblasts by SNAI2 contributes to tumor desmoplasia and ovarian cancer progression,” *Mol. Cancer*, vol. 16, pp. 1–15, Dec 2017.
- [178] X. Li, T. Gruosso, D. Zuo, A. Omeroglu, S. Meterissian, M.-C. Guiot, A. Salazar, M. Park, and H. Levine, “Infiltration of CD8+ T cells into tumor cell clusters in triple-negative breast cancer,” *Proc. Natl. Acad. Sci. U.S.A.*, vol. 116, p. 3678, Feb 2019.
- [179] T. J. Perren, A. M. Swart, J. Pfisterer, J. A. Ledermann, E. Pujade-Lauraine, G. Kristensen, M. S. Carey, P. Beale, A. Cervantes, C. Kurzeder, A. d. Bois, J. Sehouli, R. Kimmig, A. Stähle, F. Collinson, S. Essapen, C. Gourley, A. Lortholary, F. Selle, M. R. Mirza, A. Leminen, M. Plante, D. Stark, W. Qian, M. K. B. Parmar, and A. M. Oza, “A Phase 3 Trial of Bevacizumab in Ovarian Cancer,” *N. Engl. J. Med.*, vol. 365, pp. 2484–2496, Dec 2011.
- [180] “ICON7 | MRC Clinical Trials Unit at UCL,” Jan 2020. [Online; accessed 14. Jan. 2020].
- [181] K. Minami, K. Hiwatashi, S. Ueno, M. Sakoda, S. Iino, H. Okumura, M. Hashiguchi, Y. Kawasaki, H. Kurahara, Y. Mataka, *et al.*, “Prognostic significance of CD68, CD163 and Folate receptor- β positive macrophages in hepatocellular carcinoma,” *Experimental and Therapeutic Medicine*, vol. 15, p. 4465, May 2018.

- [182] L. A. Elliott, G. A. Doherty, K. Sheahan, and E. J. Ryan, “Human Tumor-Infiltrating Myeloid Cells: Phenotypic and Functional Diversity,” *Front. Immunol.*, vol. 8, p. 86., Feb 2017.
- [183] J. A. Harris, S. Jain, Q. Ren, A. Zarineh, C. Liu, and S. Ibrahim, “CD163 versus CD68 in tumor associated macrophages of classical hodgkin lymphoma,” *Diagn. Pathol.*, vol. 7, p. 12, 2012.
- [184] A. Mantovani, S. Sozzani, M. Locati, P. Allavena, and A. Sica, “Macrophage polarization: tumor-associated macrophages as a paradigm for polarized M2 mononuclear phagocytes,” *Trends Immunol.*, vol. 23, pp. 549–555, Nov 2002.
- [185] M. A. Glaire, E. Domingo, A. Sveen, J. Bruun, A. Nesbakken, G. Nicholson, M. Novelli, K. Lawson, D. Oukrif, W. Kildal, *et al.*, “Tumour-infiltrating CD8+ lymphocytes and colorectal cancer recurrence by tumour and nodal stage,” *Br. J. Cancer*, vol. 121, pp. 474–482, Aug 2019.
- [186] Z. Kos, E. Roblin, R. S. Kim, S. Michiels, B. D. Gallas, W. Chen, K. K. van de Vijver, S. Goel, S. Adams, S. Demaria, *et al.*, “Pitfalls in assessing stromal tumor infiltrating lymphocytes (sTILs) in breast cancer,” *npj Breast Cancer*, vol. 6, pp. 1–16, May 2020.
- [187] R. Posselt, K. Erlenbach-Wünsch, M. Haas, J. JeΔberger, M. Büttner-Herold, M. Haderlein, M. Hecht, A. Hartmann, R. Fietkau, and L. V. Distel, “Spatial distribution of FoxP3+ and CD8+ tumour infiltrating T cells reflects their functional activity,” *Oncotarget*, vol. 7, p. 60383, Sep 2016.
- [188] M. Ester, H.-P. Kriegel, J. Sander, X. Xu, *et al.*, “A density-based algorithm for discovering clusters in large spatial databases with noise.” in *Kdd*, vol. 96, pp. 226–231, 1996.
- [189] E. J. Amézquita, M. Y. Quigley, T. Ophelders, E. Munch, and D. H. Chitwood, “The shape of things to come: Topological data analysis and biology, from molecules to organisms,” *Dev. Dyn.*, vol. 249, pp. 816–833, Apr 2020.
- [190] M. Nicolau, A. J. Levine, and G. Carlsson, “Topology based data analysis identifies a subgroup of breast cancers with a unique mutational profile and excellent survival,” *Proc. Natl. Acad. Sci. U.S.A.*, vol. 108, pp. 7265–7270, Apr 2011.
- [191] E. Peranzoni, J. Lemoine, L. Vimeux, V. Feuillet, S. Barrin, C. Kantari-Mimoun, N. Bercovici, M. Guérin, J. Biton, H. Ouakrim, *et al.*, “Macrophages impede CD8 T cells from reaching tumor cells and limit the efficacy of anti-PD-1 treatment,” *Proc. Natl. Acad. Sci. U.S.A.*, vol. 115, pp. E4041–E4050, Apr 2018.
- [192] H. R. Ali, H. W. Jackson, V. R. T. Zanutelli, E. Danenberg, J. R. Fischer, H. Bardwell, E. Provenzano, H. R. Ali, M. Al Sa’d, S. Alon, *et al.*, “Imaging mass cytometry and multiplatform genomics define the phenogenomic landscape of breast cancer,” *Nat. Cancer*, vol. 1, pp. 163–175, Feb 2020.
- [193] P. Marques, S. Barry, E. Carlsen, D. Collier, A. Ronaldson, S. Awad, N. Dorward, J. Grieve, N. Mendoza, S. Muquit, *et al.*, “Chemokines modulate the tumour microenvironment in pituitary neuroendocrine tumours,” *Acta Neuropathol. Commun.*, vol. 7, pp. 1–21, Dec 2019.

- [194] M. Joshi, M. Oltean, P. B. Patil, D. Hallberg, M. Kleman, J. Holgersson, M. Olausson, and S. Sumitran-Holgersson, “Chemokine-Mediated Robust Augmentation of Liver Engraftment: A Novel Approach,” *Stem Cells Transl. Med.*, vol. 4, pp. 21–30, Jan 2015.
- [195] S. Z. Makaryan, C. G. Cess, and S. D. Finley, “Modeling immune cell behavior across scales in cancer,” *WIREs Syst. Biol. Med.*, vol. 12, p. e1484, Jul 2020.
- [196] G. Bocharov, V. Volpert, B. Ludewig, and A. Meyerhans, “Editorial: Mathematical Modeling of the Immune System in Homeostasis, Infection and Disease,” *Front. Immunol.*, vol. 10, Jan 2020.
- [197] L. Gole, J. Yeong, J. C. T. Lim, K. H. Ong, H. Han, A. A. Thike, Y. C. Poh, S. Yee, J. Iqbal, W. Hong, *et al.*, “Quantitative stain-free imaging and digital profiling of collagen structure reveal diverse survival of triple negative breast cancer patients,” *Breast Cancer Res.*, vol. 22, pp. 1–13, May 2020.
- [198] B. V. Jardim-Perassi, S. Huang, W. Dominguez-Viqueira, J. Poleszczuk, M. M. Budzevich, M. A. Abdalah, S. R. Pillai, E. Ruiz, M. M. Bui, D. A. P. C. Zuccari, *et al.*, “Multiparametric MRI and Coregistered Histology Identify Tumor Habitats in Breast Cancer Mouse Models,” *Cancer Res.*, vol. 79, pp. 3952–3964, Aug 2019.

Appendix A

Probability Density

The probability density function is the function that gives the probability of a particular value occurring. $P(x) = f(x)$

And the equivalent form that all probabilities sum to 1 is the following equation:

$$\int_{-\infty}^{\infty} f(x)dx = 1 \quad (\text{A.1})$$

The expected or mean value of x is

$$\bar{x} = \int_{-\infty}^{\infty} xf(x)dx \quad (\text{A.2})$$

In 2D when we analyse a point distribution with a constant probability of a point occurring, the point density k , the probability density function is the following;

$$f(r) = k \quad (\text{A.3})$$

When we integrate this in radial coordinates over a circular area with radius R ;

$$\int_0^{\infty} k.2\pi r.dr = \int_0^R k.2\pi r.dr = 1 \quad (\text{A.4})$$

$$k.\pi R^2 = 1 \quad (\text{A.5})$$

$$k = \frac{1}{\pi R^2} \quad (\text{A.6})$$

The probability k of finding a point at radius, r , is $\frac{1}{\pi R^2}$. The probability of finding N points is $\frac{N}{\pi R^2}$

Therefore the probability of finding a point between r and $r + dr$ is

$$\frac{2\pi}{\pi R^2} dr \quad (\text{A.7})$$

Appendix B

Published Work

- B.1 Combining measures of immune infiltration shows additive effect on survival prediction in high-grade serous ovarian carcinoma**



ARTICLE

Molecular Diagnostics

Combining measures of immune infiltration shows additive effect on survival prediction in high-grade serous ovarian carcinoma

Anne Montfort¹, Stephanie Owen^{2,3}, Anna M. Piskorz², Anna Supernat⁴, Luiza Moore^{2,3}, Sarwah Al-Khalidi², Steffen Böhm¹, Paul Pharoah^{2,5}, Jacqueline McDermott^{1,6}, Frances R. Balkwill¹ and James D. Brenton^{2,3,7}

BACKGROUND: In colorectal and breast cancer, the density and localisation of immune infiltrates provides strong prognostic information. We asked whether similar automated quantitation and combined analysis of immune infiltrates could refine prognostic information in high-grade serous ovarian carcinoma (HGSOC) and tested associations between patterns of immune response and genomic driver alterations.

METHODS: Epithelium and stroma were semi-automatically segmented and the infiltration of CD45RO⁺, CD8⁺ and CD68⁺ cells was automatically quantified from images of 332 HGSOC patient tissue microarray cores.

RESULTS: Epithelial CD8 [$p = 0.027$, hazard ratio (HR) = 0.83], stromal CD68 ($p = 3 \times 10^{-4}$, HR = 0.44) and stromal CD45RO ($p = 7 \times 10^{-4}$, HR = 0.76) were positively associated with survival and remained so when averaged across the tumour and stromal compartments. Using principal component analysis, we identified optimised multiparameter survival models combining information from all immune markers ($p = 0.016$, HR = 0.88). There was no significant association between PTEN expression, type of TP53 mutation or presence of BRCA1/BRCA2 mutations and immune infiltrate densities or principal components.

CONCLUSIONS: Combining measures of immune infiltration provided improved survival modelling and evidence for the multiple effects of different immune factors on survival. The presence of stromal CD68⁺ and CD45RO⁺ populations was associated with survival, underscoring the benefits evaluating stromal immune populations may bring for prognostic immunoscores in HGSOC.

British Journal of Cancer (2020) 122:1803–1810; <https://doi.org/10.1038/s41416-020-0822-x>

INTRODUCTION

There is a heterogeneous immune response in the tumour microenvironment of HGSOC, but the presence of intra-epithelial CD8⁺ T cells is consistently associated with improved survival.^{1,2} Prolonged survival is also associated with the presence of B cells and other immune cells, including CD45RO⁺ memory cells.^{3–6}

Prognostic scoring has predominantly focused on the presence, absence or semi-quantitative analysis of immune cells in tumour epithelium.⁷ Moreover, cell densities of lymphocytes, macrophages and dendritic cells have also been shown to be prognostic in breast, ovarian and lung cancer without restricting analyses to malignant epithelial cell areas.^{8–10} Approaches that combine analyses of multiple immune infiltrates, such as the CD3/CD8 immunoscore in colorectal cancer,¹¹ have not yet been developed for ovarian cancer prognosis.

The notion that anti-tumoural immune responses may be altered by tumour cell-intrinsic factors is supported by several observations. Detailed temporal and spatial histological and genomic studies in a single patient with high-grade serous

ovarian carcinoma (HGSOC) showed that Wnt signalling was upregulated in a progressing tumour nodule, a phenomenon correlated with focal immunosuppression.¹² Mutations in the TP53 and BRCA1/BRCA2 genes as well as loss of PTEN expression are driver events in HGSOC development.¹³ HGSOC cases with BRCA1 mutations have increased CD8 and CD20 intra-epithelial infiltrates,^{14,15} suggesting that loss of homologous recombination and DNA damage may prime immune responses. Whether different classes of TP53 mutations alter immune infiltration in HGSOC remains unknown, but non-synonymous mutations in TP53 may have gain-of-function or other cellular effects distinct from loss-of-function mutations.¹⁶ Mutant p53 protein may drive B cell responses and auto-antibody production.¹⁷ Loss of PTEN expression in melanoma was associated with both reduced T cell infiltration and resistance to immune checkpoint inhibitors,¹⁸ but these associations have not been investigated in HGSOC.

Despite the strong association of CD8⁺ infiltrate with prognosis, routine immunoscore for ovarian cancer is not performed in the clinic and development of automatic cellular recognition tools

¹Barts Cancer Institute, Queen Mary University of London, London, UK; ²Cancer Research UK Cambridge Institute, Cambridge, UK; ³Department of Oncology, University of Cambridge, Cambridge, UK; ⁴Laboratory of Translational Oncology, Intercollegiate Faculty of Biotechnology, Medical University of Gdańsk, 80-211 Gdańsk, Poland; ⁵Department of Public Health and Primary Care, Centre for Cancer Genetic Epidemiology, University of Cambridge, Cambridge, UK; ⁶Department of Pathology, University College London Hospital, London, UK and ⁷Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK

Correspondence: Frances R. Balkwill (f.balkwill@qmul.ac.uk) or James D. Brenton (James.Brenton@cruk.cam.ac.uk)

These authors contributed equally: Anne Montfort, Stephanie Owen

Received: 19 September 2019 Revised: 8 February 2020 Accepted: 11 March 2020

Published online: 6 April 2020

could be beneficial for high-throughput pathology workflows. To test the hypothesis that a more integrated analysis of cytotoxic, mature and organised immune responses in the tumour micro-environment might have greater prognostic value in HGSOc, we developed new image analysis methods and workflows to test the single and integrated analysis of CD8⁺, CD45RO⁺ and CD68⁺ in the malignant epithelium and adjacent stroma. The secondary aim of this study was to test for association between quantitative measures of CD8⁺, CD45RO⁺ and CD68⁺ cells and driver genomic alterations in *BRCA1*, *BRCA2*, *PTEN* and different classes of *TP53* mutation.

MATERIALS AND METHODS

Patients

Samples from 570 patients from the prospective SEARCH ovarian cancer population-based study were used to construct tissue microarrays (TMAs).¹⁹ Ethical approval was granted by the Eastern Multicenter Research Ethics Committee. Among the samples from 570 patients with primary epithelial ovarian tumours, 332 were high-grade serous ovarian cancer patients. All cases underwent detailed histopathological review by a gynaecological pathologist (J.Mc.D.). Patients were staged as having localised, regional or distant disease (L/R/D).²⁰

Immunohistochemistry

Microarray slides composed of formalin-fixed paraffin-embedded ovarian tumour cores were dewaxed and rehydrated prior to heat-induced epitope retrieval using a pressure cooker and a citrate-based antigen unmasking solution (Vector Laboratory). Detection of CD8⁺ T cells, CD45RO⁺ memory lymphocytes and CD68⁺ macrophages was performed using the mouse anti-human CD8 (clone C8/144B, Dako), mouse anti-human CD45RO (clone UCLH, Dako) and mouse anti-human CD68 (clone M0876, Dako) antibodies, using ultrasensitive Polymer-HRP IHC Detection system (Biogenex). Immunohistochemical protocols and slide hybridisations were carried out manually. Sections were counterstained with haematoxylin and mounted with DPX mounting medium (Sigma). Previously published PTEN immunostaining data was used where high PTEN expression was considered to be positive staining and low expression to be weak, heterogeneous or negative staining, respectively.²¹

Mutation analysis

The coding regions of *TP53* were sequenced by tagged-amplicon next-generation sequencing as previously described²² and confirmed by immunohistochemical analysis using a 4-tier core system.²³ Sequencing of germline mutations in the *BRCA1* and *BRCA2* genes was performed as previously described.²⁴

Immune cell quantification

Stained slides were scanned using the Panoramic Slash Scanner (3D Histech). The number of CD8⁺ and CD45RO⁺ cells per mm² of epithelial and stroma areas, as well as the percentage of epithelial and stromal areas covered by CD68 staining, were digitally determined using the Tissue Studio software (Definiens™). Definiens image analysis algorithms for detection of epithelial and stromal areas were trained and the segmentation for each core was manually refined by two researchers, including a consultant gynaecological-histopathologist (J.Mc.D.). Supplementary Fig. 1 shows examples of classifications of tissue regions and cell detection and the entire data set, including these epithelial and stromal assignments, can be downloaded from the repository.

Statistical analyses

R (version 3.5.1) was used for statistical analysis and an R markdown document containing the entire data set allowing for performing all analyses is available (<https://bitbucket.org/>

[jamesdbrenton/search-montfort/src/master/](https://bitbucket.org/jamesdbrenton/search-montfort/src/master/)). Quality checking for spatial bias across TMAs and effects of varying tissue area was carried out upon all cores and across TMAs using heatmaps and Shapiro–Wilk tests. All count data were transformed to log base 10 after adding a small offset to zero values. Wilcoxon's signed-rank test was used to compare the mean infiltrate between groups. Continuous data were presented as median and interquartile range (IQR) and groups were compared by the Kruskal–Wallis and pairwise Kruskal–Wallis tests. Discrete data were presented as count and percentage.

Cox proportional hazard regression analysis was applied to assess the effect of each infiltrate on overall survival. The functional form of each of the immune variables was assessed using comparison with cubic splines. The best approximations to the functional forms were carried forward for the Cox models. The clinical variables of age at diagnosis, menopause status and stage were available for the cohort and were included in the analysis. Univariable Cox regressions were used to identify best-fitting variables for the final multivariable Cox regression model. The refined model was compared with a combined multivariable Cox regression model including all immune infiltrates. Hazard ratios (HRs) refer to a single unit increase in continuous variables. The proportional hazards assumption was tested and satisfied in all cases using Schoenfeld residuals. The Kaplan–Meier analysis (with log-rank test) was applied to illustrate survival differences graphically. Two-sided *p* values <0.05 were used to indicate statistical significance. Principal component analysis (PCA) using the R package *prcomp* was used to extract the independent components of variance between patients. The package *prcomp* uses singular value decomposition and the variables were scaled to have unit variance before creating composite linear independent variables. These were then passed forward to the survival modelling. The Akaike information criterion (AIC) was used to compare the performance of survival models, which includes a penalty on the number of terms to reduce overfitting. Bonferroni *p* value corrections were carried out for all multiple testing. *P* < 0.05 was considered significant for all analyses.

RESULTS

Patient characteristics

Supplementary Fig. 2 shows the REMARK diagram for this study and Supplementary Table 1 shows the clinical characteristics of the 332 HGSOc patients from the study cohort. Immunohistochemical analyses on TMAs were performed to detect CD8⁺, CD45RO⁺ and CD68⁺ cells in tissue cores from primary ovarian specimens. One hundred and fifty-two HGSOc cases were available for analysis after quality assurance, data cleaning and the reduction of the data set to only cases with complete results for CD8, CD45RO and CD68 staining in both epithelium and stroma, as well as survival data.

Tagged-amplicon sequencing was performed on 248 cases and *TP53* mutation was detected in 231 samples (93%) (Supplementary Table 1). Previously published data for germline *BRCA1* and *BRCA2* mutation and PTEN expression were available for 297 and 155 cases, respectively.^{18,22}

Digital pathology analysis of tumour composition and immune cell densities

Image analysis software was used to determine the area of tumour epithelium and stroma in each core (Fig. 1 and Supplementary Fig 1). Of 964 images representing 332 HGSOc patients, 69 patients (20.8%) had images that contained malignant epithelium but no stroma; 250 patients (75.3%) had images that contained epithelium and >1% adjacent stroma and 13 patients (3.9%) had images containing no tumour epithelium (Fig. 1a). The median proportion of epithelium and stroma was 85.1% (IQR 51–100%) and 14.9% (IQR 0–49%), respectively. We expected the proportion

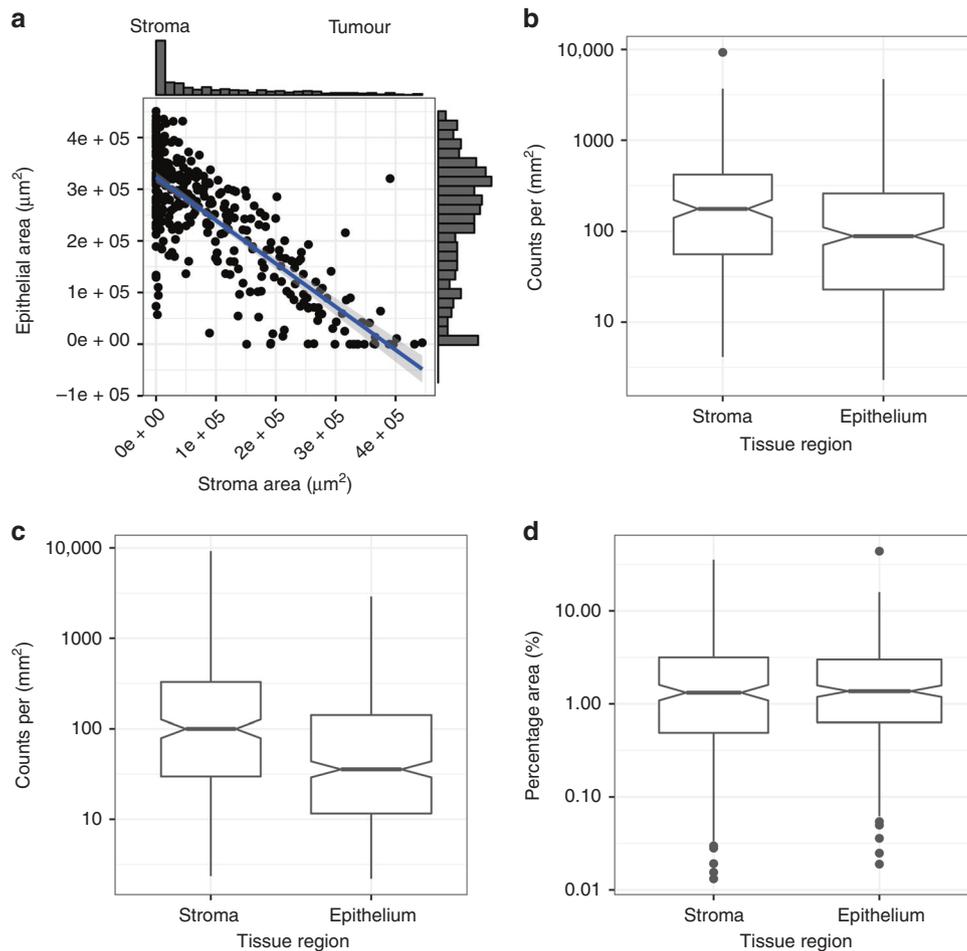


Fig. 1 Proportions of immune cells vary between epithelial and stroma areas of tumours. **a** Scatter plot of the average stroma and epithelial tumour areas for each patient. **b–d** show respectively the distribution of densities of CD8⁺, CD45RO⁺ and CD68⁺ cell in epithelium and stroma. CD8⁺ and CD45RO⁺ densities were defined as counts per mm² and CD68⁺ as the percentage of tissue stained for this marker. Notches on box plots extend $1.58 \times \text{IQR}/\sqrt{n}$ and approximate the 95% confidence interval for the median. Box plot whiskers extend to $1.5 \times \text{IQR}$.

of tumour in a sample to be correlated with p53 mutant allele fraction, a measure of sample purity, and found them to be positively correlated ($R^2 = 0.25$, $p = 0.0004$).

Quantitative densities of all immune populations were then generated automatically through image analysis, the quantitative assessment of CD8⁺ T cell, CD45RO⁺ memory lymphocyte and CD68⁺ macrophage densities in each compartment are shown in Supplementary Fig. 1. The relationship between the fraction of tumour in a core and the density of immune infiltration in the epithelium was examined. Intra-epithelial CD8⁺ and CD45RO⁺ densities were weakly correlated with the purity/tumour fraction of the sample ($R^2 = 0.17$, $p = 0.003$ and $R^2 = 0.16$, $p = 0.006$), but CD68⁺ epithelial density was not.

The three immune populations in our samples showed moderate to strong correlation between epithelium and stroma (Supplementary Fig. 3 and Supplementary Table 2). Samples with low density of stromal immune populations generally had low density of epithelial infiltrate and vice versa. The distribution of densities of immune populations within tumour epithelium and stromal areas were compared (Fig. 1b–d). The density of CD8⁺ and CD45RO⁺ cells were significantly higher in stroma than in tumour epithelium ($p = 0.005$ and $p = 0.004$, respectively; Welch's *t* test), but not significantly different for CD68⁺ cells.

In describing the patterns of immune infiltrate, the terms immune inflamed, immune desert and immune excluded have

been used to describe varying T cell infiltration based on histological and transcriptional analyses.^{12,25–27} Immune-inflamed and immune-desert patterns reflect high positive or negative correlations between all infiltrates, but T cell exclusion describes tumours where CD8⁺ cells are significantly absent from tumour epithelium while still being present in the surrounding stroma.^{28,29} Given the higher infiltration in stroma than epithelial compartments, we defined immune cell exclusion as a 10-fold difference between tumour epithelium and stromal infiltration as the standard deviation of the log₁₀-transformed counts was ~ 1 . CD8⁺ exclusion was present in 20 (10.6%) cases and 36 (20%) cases had CD45RO⁺ exclusion. No cases had significant exclusion of CD68⁺ infiltrate from tumour epithelium. Notably none of the cases had both CD8⁺ and CD45RO⁺ exclusion.

Stromal CD68⁺ and CD45RO⁺ densities are the strongest individual prognostic markers

Survival was modelled using Cox proportional hazards and the relationship between the immune variables and survival was found to be approximately log linear. The clinical variables accompanying the cohort were age at diagnosis, stage and menopause status and the relationship between age and survival was found to be approximately linear (see Methods, Supplementary Fig. 4 and Supplementary Table 3).

Table 1. Hazard ratios from the Cox proportional model for all infiltrates in all regions and averaged across the whole core in HGSOc, measured as log 10 (counts per mm²).

Functional form	Evaluable cases	Tissue compartment	Univariable		Multivariable ^a (adjusted for stage)	
			HR	p Value	HR	p Value
CD8⁺						
Log 10	301	Epithelium	0.89	0.15	0.83	0.027
Log 10	202	Stroma	0.97	0.74	0.93	0.40
Log 10	315	Average	0.79	0.010	0.72	0.0006
CD45RO⁺						
Log 10	290	Epithelium	0.86	0.033	0.85	0.022
Log 10	196	Stroma	0.76	0.001	0.76	0.0007
Log 10	306	Average	0.82	0.006	0.80	0.003
CD68⁺						
Linear	293	Epithelium	0.99	0.67	0.99	0.43
Log 10	226	Stroma	0.53	0.003	0.44	0.0003
Log 10	308	Average	0.67	0.042	0.62	0.017
Stage						
-	312	Localised	1	0	1	0
		Regional	1.47	0.26	1.15	0.25
		Distant	3.96	<<0.001	5.58	<<0.001
		Unstaged	3.35	<0.001	3.34	<<0.001

^aMultivariable analysis was adjusted for stage.

Univariable analysis showed improved survival with increasing stromal density of CD45RO⁺ [HR 0.76, 95% confidence interval (CI): 0.65–0.90, *p* = 0.001] and CD68⁺ (HR 0.53, 95% CI: 0.34–0.81, *p* = 0.003) (Table 1). Modelling each immune variable with stage showed improved predictive value for epithelial CD8⁺ density (HR = 0.83, *p* value = 0.027) as well as stromal CD68⁺ and CD45RO⁺ density and epithelial CD45RO⁺ density (Table 1). Figure 2 and Supplementary Fig. 5 shows illustrative Kaplan–Meier survival curves for high and low stromal and epithelial CD68⁺, CD45RO⁺ and CD8⁺ densities.

In cores with <1% stroma, epithelial CD8⁺ malignant epithelial infiltrate remained an independent prognostic factor, but epithelial CD45RO⁺ density was not significant (Supplementary Table 4).

In clinical reporting, quantifying immune populations in exclusively tumour epithelium is technically challenging and time consuming. We tested the effects of using the average density of each marker averaged across both tumour and stromal regions from each core (Table 1). Averaging the tissue density of CD8⁺ increased the strength of the associated HR and significance of the model (HR = 0.79, *p* value = 0.010) indicating increased prognostic value over quantitation of individual epithelial and stromal infiltrates. Supplementary Fig. 5 shows illustrative Kaplan–Meier survival curves for high and low CD68⁺, CD45RO⁺ and CD8⁺ densities over combined epithelium and stroma compartments.

Multivariable Cox regression analysis including all infiltrates and stage was carried out on patients with complete data for all infiltrates (*n* = 152) (Table 2). In this model, only stage and CD68⁺ stromal infiltrate were significant predictors of survival.

We then refined the model by removing the least significant elements (defined as those with *p* > 0.1) (Table 2, *n* = 152). Interestingly, we found that the *p* values for CD68⁺ and CD45RO⁺ stromal infiltrates in the refined model become less significant and the hazard ratios are attenuated in comparison to both the univariable regression and the full model. This is likely due to the inability of Cox regression to distinguish with confidence whether stromal CD45RO⁺ or CD68⁺ density is the most significant

predictor when there is a strong correlation between all the immune variables.

Principal components of the combined immunospace describe biologically interpretable effects

As the three types of immune infiltrate vary continuously across epithelium and stroma these variables can be regarded as six dimensions of an ‘immunospace’ (three infiltrates, two localisations). Given the strong correlations between infiltrates (Supplementary Fig. 3), these immune variables are not independent. We used PCA to determine the independent patterns across these dimensions, using the 152 patients for whom complete data were available. PCA transformed the six correlated infiltrate variables into six independent axes with the first component containing the largest proportion of variance (60%) in the data set (Supplementary Table 5).

In principal component 1 (PC1), the weightings of all immune infiltrates are positive and similar in magnitude (Supplementary Table 5). This indicates that as one infiltrate increases so do all the others and represents the degree of coordinated immune response. The remaining PCs characterise additional patterns across immune infiltrates independent of PC1. The additive contribution of PC2 characterises negative correlation between CD8⁺ infiltrates and CD68⁺ macrophages and CD45RO⁺ memory cells. PC3 characterises additional variation where epithelial and stromal infiltrates are negatively correlated, the most positive values of PC3 correspond to high infiltration in tumour epithelium compared to stroma and the most negative values of PC3 correspond to the opposite, the aforementioned immune exclusion.

Supplementary Fig. 6 shows representative images with the largest magnitudes of PC1, PC2 and PC3 to visually illustrate the patterns described above. The variance in the remaining PCs (4–6) is smaller and less informative. Variance in PC4 is predominantly from CD45RO⁺ stromal density, PC5 is from CD45RO⁺ epithelial density and PC6 is from CD68⁺ epithelial density. Patients are plotted by their PC1 and PC2 values in Supplementary Fig. 7. Cox

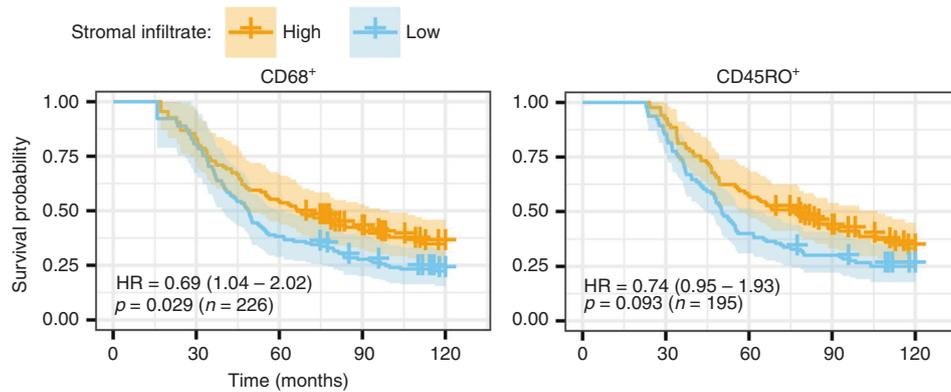


Fig. 2 Survival analysis of HGSOC patients relative to the density of CD68⁺ macrophages and CD45RO⁺ cells in their stroma. Kaplan–Meier survival curves using cut point of median density of stromal CD68⁺ macrophages and CD45RO⁺ memory T cells using left truncation and right censoring. Median entry to the study for all patients after diagnosis was 26.4 months. Median follow-up time from diagnosis to exit or death was 105.1 months.

Table 2. Multivariable Cox regression hazard ratios and associated *p* values for a model with all infiltrates and stage and for a reduced model with most significant variables only (*n* = 152).

	Multivariable (all combined)		Refined model	
	HR	<i>p</i> Value	HR	<i>p</i> Value
CD8⁺				
Epithelium	0.96	0.81	–	–
Stroma	1.07	0.63	–	–
CD45RO⁺				
Epithelium	1.12	0.37	–	–
Stroma	0.83	0.09	0.68	0.11
CD68⁺				
Epithelium	1.16	0.63	–	–
Stroma	0.53	0.038	0.88	0.17
Stage				
Localised	1	0	1	0
Regional	2.00	0.16	2.03	0.14
Distant	4.82	≪0.001	4.70	0.0001
Unstaged	8.15	≪0.001	8.25	0.0001

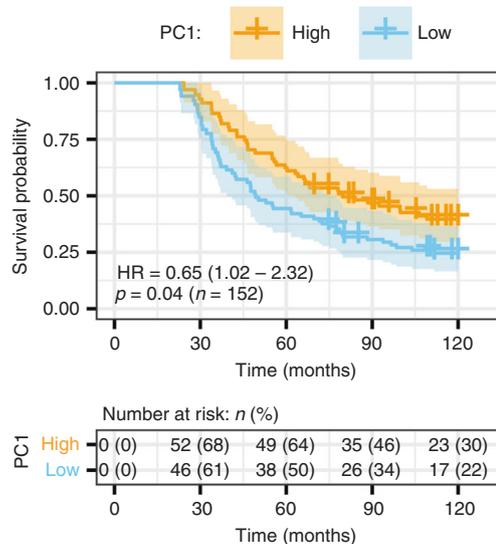


Fig. 3 Survival analysis by principal component 1 (PC1). Kaplan–Meier analysis using median cut point for PC1 with left truncation for study entry variables and right censored at 120 months. Median entry to the study for all patients after diagnosis was 26.4 months. Median follow-up time from diagnosis to exit or death was 105.1 months.

proportional hazard regression was used to assess whether these PCs were predictive of survival. Only PC1 was an independent predictor of survival in our cohort and was associated with improved survival (univariate: HR = 0.89, *p* value = 0.024; PC1 + stage: HR: 0.88, *p* value = 0.019) (Supplementary Table 6), reflecting the good prognosis of a strong coordinated immune response. The association of this PC with survival is also illustrated graphically using Kaplan–Meier curves in Fig. 3.

Cox regression survival models were also calculated on all combinations of PCs and stage. The AIC was used to compare the performance of these survival models and includes a penalty on the number of terms to reduce overfitting (Supplementary Fig. 8). The model combining stage, PC1 and PC5, had the best performance for predicting overall survival. The improvement with the addition of PC5 shows that the addition of this PC has a suppressor effect in the model, increasing the significance of other variables when included. This demonstrates that survival is predominantly determined by the coordinated immune response and further variation in survival from this trend can be encoded by the quantity of epithelial CD45RO⁺ infiltrate. Interestingly, the models that contained stage and either stromal CD45RO⁺ or

CD68⁺ infiltrate contained a similar amount of information about patient survival as the one that contained stage and PCs 1 and 5. In our cohort, the density of CD68⁺ and CD45RO⁺ stromal infiltrates are therefore the best single infiltrates for survival modelling.

Are genetic defects associated with HGSOC driving individual infiltrates or the coordinated immune response in tumours? The presence of germline *BRCA2* mutations was significantly associated with lower CD8⁺ cell density than patients with a *BRCA1* after multiple testing correction (Table 3). There was no significant association between the quantity of CD45RO⁺ or CD68⁺ infiltrate and the mutational status of either *BRCA1* or *BRCA2* genes (Table 3). No significant association was detected between *TP53* gain-of-function (GOF) and loss-of-function (LOF) mutations or PTEN expression and the densities of CD8⁺, CD45RO⁺ and CD68⁺ cells in the epithelium or stroma (Table 3). Similarly, changes in the PCs were not significantly associated with PTEN expression, *TP53* GOF or LOF mutation or germline *BRCA1/BRCA2* mutation status (Table 3).

Table 3. Association between immune infiltrate and genetic alterations in HGSOC tumours.

	<i>BRCA1/BRCA2</i>	p53 GOF/LOF	PTEN
CD8⁺			
Epithelium	0.04	0.54	0.67
Stroma	0.17	0.29	0.07
Average	0.17	0.53	0.84
CD45RO⁺			
Epithelium	0.20	0.42	0.92
Stroma	0.46	0.84	0.82
Average	0.10	0.36	0.87
CD68⁺			
Epithelium	0.43	0.72	0.79
Stroma	0.78	0.85	0.65
Average	0.68	0.66	0.41
Principal component			
1	0.81	0.80	0.55
2	0.10	0.26	0.17
3	0.27	0.58	0.51
4	0.58	0.99	0.17
5	0.76	0.88	0.79
6	0.97	0.44	0.18

P values were associated with Kruskal–Wallis test for detecting differences in mean ranks of immune infiltrate in patients grouped by mutation in *BRCA1*, *BRCA2* or not-detected, p53 GOF or LOF and PTEN high or low in HGSOC.

DISCUSSION

Infiltration of immune cells in HGSOC tumours has previously been correlated to prognosis for patients.^{1,30} However, translation of these tools to the clinic has been impeded by the lack of standardised and reliable quantitation methods. In this work, we used a semi-automated approach to quantify the CD8⁺, CD45RO⁺ and CD68⁺ infiltrates in both stromal and epithelial areas of primary HGSOC tumours. We found that immune infiltration in the tumour microenvironment was continuously distributed across a wide range and CD8⁺, CD45RO⁺ and CD68⁺ infiltrates were strongly correlated. T cell exclusion from epithelial regions has been previously reported.²⁹ We only observed epithelial exclusion of CD8⁺ or CD45RO⁺ in 10% and 20% of patients, respectively. These groups were mutually exclusive, suggesting that immune exclusion from epithelial or stromal regions, is a tumour-specific effect that is distinct from coordinated multi-infiltrate immune responses.

Our results for the positive prognostic effects of epithelial CD8⁺ T cells are consistent with the large study by Goode et al.,² who described a near log-linear relationship between the density of epithelial CD8⁺ T cells and survival. However, our observations also show that averaging CD8 density across the total core (epithelial and stromal areas) improves survival prediction. The alternative possibility that epithelial and stroma compartments have been mis-assigned is highly unlikely as these assignments were all reviewed by a gynaecological pathologist. The combination of both stroma and tumour epithelium potentially provides a better representation of the dynamics of epithelial infiltration by including cells that may be poised to infiltrate.

CD45RO is a memory marker for T and B cells. In the tumour microenvironment, we and others^{5,31,32} have shown that CD4⁺ and CD8⁺ T cells, myeloid cells and B cells are mainly located in the stroma of tumours, whereas CD8⁺ T cells can be found in both areas. Meta-analyses investigating the prevalence of CD45RO in

solid cancers reported that high CD45RO infiltrate was associated with better prognosis for patients.³³ We confirmed these results in HGSOC and demonstrated that the quantity of stromal and average densities of CD45RO⁺ cells over the whole core are prognostic. We also demonstrated that epithelial CD45RO⁺ density was a significant predictor of survival in cores with >1% stroma, but not in those with <1% stroma. This result emphasises the need to contextualise epithelial infiltrate and the confounding nature of tumour composition on survival impact.

The link between tumour infiltration with macrophages and patient survival is more complex. Previous studies demonstrated the positive predictive value of classically activated (M1) over tumour-promoting (M2) macrophages in ovarian cancer.^{9,34} However, no significant association was found between the level of intra-epithelial CD68 infiltrate and patient survival.⁴ Our study confirms this latter result and also demonstrates a significant association between a high coverage of stromal areas with CD68⁺ macrophages and improved overall survival. This result is in line with a recent study showing a high infiltration of CD68⁺ macrophages at the invasive front of tumour sections from colorectal cancer patients is correlated to better response to chemotherapy.³⁵ Whereas tumour-associated macrophages (TAM) promote cancer progression via stimulation of multiple processes including angiogenesis, inflammation and immune escape, it now appears evident that the phenotype of TAM changes in different areas of the tumour.^{36,37} Taking into account the seemingly contradictory results obtained on this subject, additional studies aiming at deciphering the precise role distinct subpopulations of macrophages infiltrating different areas of tumours play in cancer progression is required.

All three infiltrates were also found to be predictive in multivariate analyses adjusted for tumour stage. We previously demonstrated that optimal response to neoadjuvant chemotherapy was significantly associated with a decrease in the density of FoxP3⁺ regulatory T cells in the stroma of human HGSOC tumours.³⁸ As we did not have access to treatment information for patients included in this cohort, we cannot rule it out as a potential confounding factor.

Importantly, we show that combining our six correlated types of immune infiltrates using PCA allowed us to transform inter-patient variations of the tumour immune landscape into independent, biologically interpretable PCs. We find that the main source of variation between patients is reflective of the quantity of concerted immune response. The second PC demonstrates that there is a pattern of significant variations in the CD8⁺ infiltrate of some patients that is independent and acting in opposition to the concerted immune response. The third PC measures the extent of negatively correlated epithelial and stromal infiltrates and is weighted differently by infiltrates. This supports our observation that immune exclusion is common, variable by infiltrate and shows that this is an additional effect to that of the concerted immune response.

The main limitation of our analysis of the ‘immunospace’ and its PCs is that we could only apply it to images that contain both epithelium and stroma. This reduced the size of the cohort we could analyse. With foresight, sampling could be designed to include both regions in TMA and sectioning and imputation could be carried out for missing data. It is worth noting that being able to combine multiple correlated infiltrates and elucidate patterns and sources of variation will become even more important as many more (30+) immune markers are combined on single sections using imaging methods such as Hyperion. This method is particularly useful in that it also allows us to measure the strength of immune patterns occurring across multiple infiltrates. Analysing PCs also avoids some statistical issues that are associated with frequent co-correlation between different immune populations. Including correlated variables in a typical multivariable Cox regression survival model not only reduces

interpretability but can also cause model results to be variable over bootstrapping as seen in the refined model of our cohort.

We found that the first PC was independently associated with survival. This result is consistent with a previous study showing that the simultaneous infiltration of tumours by different subsets of leucocytes (e.g. T cells, plasma cells, B cells), likely reflecting the establishment of a concerted immune response, gives a survival advantage to ovarian cancer patients.⁵ We used the PCs to model survival and found that a model including PC1 and PC5 was as predictive as individually modelled stromal CD45RO and CD68 infiltrates. Therefore, CD68 or CD45RO markers analysed in epithelium-adjacent stroma constitute the best single prognostic markers in our cohort.

In human lung adenocarcinoma, Mansuet-Lupo et al.³⁹ showed that oncogenic mutations in the *TP53* gene were positively correlated with CD8⁺ infiltrate. In their study, intra-epithelial CD8⁺ T cell numbers and *TP53* status were both linked to prognosis with patients harbouring tumours with low CD8⁺ infiltrate and non-disruptive *TP53* mutation (associated with GOF) being linked to poorer survival. In glioblastoma, *TP53* mutations were also associated with increased immune infiltrate.⁴⁰ In the present work, we integrated histological and genomic features to evaluate whether genetic alterations could be linked to different quantities of immune infiltration. We found no correlation between the nature of *TP53* mutations and the amount of CD8⁺ T cells, CD45RO⁺ memory lymphocytes and CD68⁺ macrophages in the stromal and malignant areas of 197 HGSOc tumours or the values of the PCs we derived. Overall, the exact nature of the relationship between subtypes of *TP53* mutations and immune infiltrate is still unclear and likely to vary across different cancers, as evidenced by two studies showing *TP53* GOF mutations to be associated with pro-tumour effects related to inflammation in glioblastoma⁴⁰ and immunosuppression in lung adenocarcinoma.⁴¹

PTEN, another essential tumour suppressor, regulates the production of immunosuppressive cytokines by melanoma cells.⁴² In human melanoma, *PTEN* deletion was correlated with a decrease in infiltrating CD8⁺ T cells.¹⁸ In HGSOc, however, we observed no correlation between the expression level of *PTEN* and the number and/or localisation of CD8⁺, CD45RO⁺ and CD68⁺ leucocytes in the tumour microenvironment or the PCs we derived. The discrepancy between the results obtained with HGSOc and melanoma tumours might lie in the fact that in our study loss of *PTEN* expression was only partial with tumours divided into expressing high or low levels of *PTEN*.

In contradiction to a previous work,⁴³ we did not observe a significant correlation between the number of intra-epithelial CD8⁺ T cells and mutations in the *BRCA1* gene. As there were only 18 patients with a *BRCA1* mutation in this cohort, this is most likely related to a broad confidence interval as compared with the null population. We did however find a significant difference in intra-epithelial CD8⁺ T cell infiltration between patients with *BRCA1* and *BRCA2* mutations, with *BRCA1* patients having significantly higher infiltration than *BRCA2* cases. The discrepancy could also be linked to the automated and continuous quantitation method used in our study to quantify the immune infiltrate of tumours, as scoring in the previous study was done manually and in a stratified manner.⁴³ Furthermore, we did not assess the methylation in the gene coding region for *BRCA1* and restricted our analysis to the identification of germline mutations in the *BRCA1* and *BRCA2* genes.

Our work shows that averaging immune infiltration over the whole tissue core could be as useful as current methods and potentially provides more prognostic information. Methods that average counts of immune cells across tissues are also simpler to implement than epithelial:stromal segmentation methods. Nonetheless, careful sampling of the tumour microenvironment with inclusion of both epithelium and stroma remains very important as both areas have independent prognostic significance.

Our results showing strong positive prognostic significance of stromal CD68⁺ infiltrate in HGSOc tumours warrants further investigation into the role and properties of CD68⁺ macrophages in HGSOc and also may caution against the use of new potential macrophage-depleting therapies.

ACKNOWLEDGEMENTS

We are grateful to the patients who consented to donate tumour samples and to Dr. Filipe Correia Martins for helpful discussions about the PTEN analysis.

AUTHOR CONTRIBUTIONS

Experiments and data analysis: A.M., S.O., S.B., J.McD., P.P., J.D.B., F.R.B., A.M.P., A.S., L.M., S.A.-K.; TMA cohort: P.P.; writing paper: J.D.B., F.R.B., A.M., S.O.; designed study and acquired funding: F.R.B. and J.D.B.

ADDITIONAL INFORMATION

Ethics approval and consent to participate Ethical approval for the collection of tumour biopsies from ovarian cancer patients was granted by the Eastern Multicenter Research Ethics Committee. All patients enrolled in this study provided informed consent for the use of their tissue samples for research purposes. This study was performed in accordance with the Declaration of Helsinki.

Data availability The R markdown document containing the entire data set and allowing for reproducing all analyses performed in this manuscript is available online (<https://bitbucket.org/jamesdbrenton/search-montfort/src/master/>).

Competing interests P.P. is a member of the British Journal Cancer Editorial Board. All other authors declare no competing interests.

Funding information We acknowledge funding and support from Cancer Research UK (grant numbers A16354, A25714, A15973, A15601, A22905, A17197), ERC Advanced Grant ERC322566, the Queen Mary University of London, the National Institute for Health Research Cambridge, the National Cancer Research Network and the Cancer Research UK Experimental Cancer Medicine Centres and Hutchison Whampoa Limited. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Supplementary information is available for this paper at <https://doi.org/10.1038/s41416-020-0822-x>.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

- Zhang, L., Conejo-Garcia, J. R., Katsaros, D., Gimotty, P. A., Massobrio, M., Regnani, G. et al. Intratumoral T cells, recurrence, and survival in epithelial ovarian cancer. *N. Engl. J. Med.* **348**, 203–213 (2003).
- Goode, E. L., Block, M. S., Kalli, K. R., Vierkant, R. A., Chen, W., Fogarty, Z. C. et al. Dose-response association of CD8⁺ tumor-infiltrating lymphocytes and survival time in high-grade serous ovarian cancer. *JAMA Oncol.* **3**, e173290 (2017).
- Zhang, Z., Huang, J., Zhang, C., Yang, H., Qiu, H., Li, J. et al. Infiltration of dendritic cells and T lymphocytes predicts favorable outcome in epithelial ovarian cancer. *Cancer Gene Ther.* **22**, 198–206 (2015).
- Milne, K., Köbel, M., Kaloger, S. E., Barnes, R. O., Gao, D., Gilks, C. B. et al. Systematic analysis of immune infiltrates in high-grade serous ovarian cancer reveals CD20, FoxP3 and TIA-1 as positive prognostic factors. *PLoS ONE* **4**, e6412 (2009).
- Kroeger, D. R., Milne, K. & Nelson, B. H. Tumor-infiltrating plasma cells are associated with tertiary lymphoid structures, cytolytic T-cell responses, and superior prognosis in ovarian cancer. *Clin. Cancer Res.* **22**, 3005–3015 (2016).
- Leffers, N., Gooden, M. J., de Jong, R. A., Hoogbeem, B. N., ten Hoor, K. A., Hollema, H. et al. Prognostic significance of tumor-infiltrating T-lymphocytes in primary and metastatic lesions of advanced stage ovarian cancer. *Cancer Immunol. Immunother.* **58**, 449–459 (2009).
- Fridman, W. H., Zitvogel, L., Sautès-Fridman, C. & Kroemer, G. The immune contexture in cancer prognosis and treatment. *Nat. Rev. Clin. Oncol.* **14**, 717–734 (2017).

8. Ali, H. R., Dariush, A., Thomas, J., Provenzano, E., Dunn, J., Hiller, L. et al. Lymphocyte density determined by computational pathology validated as a predictor of response to neoadjuvant chemotherapy in breast cancer: secondary analysis of the ARTemis trial. *Ann. Oncol.* **28**, 1832–1835 (2017).
9. Zhang, M., He, Y., Sun, X., Li, Q., Wang, W., Zhao, A. et al. A high M1/M2 ratio of tumor-associated macrophages is associated with extended survival in ovarian cancer patients. *J. Ovarian Res.* **7**, 19 (2014).
10. Dieu-Nosjean, M. C., Antoine, M., Danel, C., Heudes, D., Wislez, M., Poulot, V. et al. Long-term survival for patients with non-small-cell lung cancer with intratumoral lymphoid structures. *J. Clin. Oncol.* **26**, 4410–4417 (2008).
11. 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. *J. Clin. Oncol.* **27**, 5944–5951 (2009).
12. Jiménez-Sánchez, A., Memon, D., Pourpe, S., Veeraghavan, H., Li, Y., Vargas, H. A. et al. Heterogeneous Tumor-Immune Microenvironments among Differentially Growing Metastases in an Ovarian Cancer Patient. *Cell* **170**, 927–938.e920 (2017).
13. Bowtell, D. D., Böhm, S., Ahmed, A. A., Aspuria, P. J., Bast, R. C., Beral, V. et al. Rethinking ovarian cancer II: reducing mortality from high-grade serous ovarian cancer. *Nat. Rev. Cancer* **15**, 668–679 (2015).
14. Dai, Y., Sun, C., Feng, Y., Jia, Q. & Zhu, B. Potent immunogenicity in BRCA1-mutated patients with high-grade serous ovarian carcinoma. *J. Cell Mol. Med.* <https://doi.org/10.1111/jcmm.13678> (2018).
15. McAlpine, J. N., Porter, H., Köbel, M., Nelson, B. H., Prentice, L. M., Kallinger, S. E. et al. BRCA1 and BRCA2 mutations correlate with TP53 abnormalities and presence of immune cell infiltrates in ovarian high-grade serous carcinoma. *Mod. Pathol.* **25**, 740–750 (2012).
16. Zhou, X., Hao, Q. & Lu, H. Mutant p53 in cancer therapy—the barrier or the path. *J. Mol. Cell. Biol.* **11**, 293–305 (2019).
17. Kumar, S., Mohan, A. & Guleria, R. Prognostic implications of circulating anti-p53 antibodies in lung cancer—a review. *Eur. J. Cancer Care (Engl)* **18**, 248–254 (2009).
18. Peng, W., Chen, J. Q., Liu, C., Malu, S., Creasy, C., Tetzlaff, M. T. et al. Loss of PTEN promotes resistance to T cell-mediated immunotherapy. *Cancer Discov.* **6**, 202–216 (2016).
19. Song, H., Ramus, S. J., Quaye, L., DiCioccio, R. A., Tyrer, J., Lomas, E. et al. Common variants in mismatch repair genes and risk of invasive ovarian cancer. *Carcinogenesis* **27**, 2235–2242 (2006).
20. Walters, S., Maringe, C., Butler, J., Brierley, J. D., Rachet, B. & Coleman, M. P. Comparability of stage data in cancer registries in six countries: lessons from the International Cancer Benchmarking Partnership. *Int. J. Cancer* **132**, 676–685 (2013).
21. Martins, F. C., Santiago, I., Trinh, A., Xian, J., Guo, A., Sayal, K. et al. Combined image and genomic analysis of high-grade serous ovarian cancer reveals PTEN loss as a common driver event and prognostic classifier. *Genome Biol.* **15**, 526 (2014).
22. Piskorz, A. M., Ennis, D., Macintyre, G., Goranova, T. E., Eldridge, M., Segui-Gracia, N. et al. Methanol-based fixation is superior to buffered formalin for next-generation sequencing of DNA from clinical cancer samples. *Ann. Oncol.* **27**, 532–539 (2016).
23. Köbel, M., Piskorz, A. M., Lee, S., Lui, S., LePage, C., Marass, F. et al. Optimized p53 immunohistochemistry is an accurate predictor of. *J. Pathol. Clin. Res.* **2**, 247–258 (2016).
24. Song, H., Cicek, M. S., Dicks, E., Harrington, P., Ramus, S. J., Cunningham, J. M. et al. The contribution of deleterious germline mutations in BRCA1, BRCA2 and the mismatch repair genes to ovarian cancer in the population. *Hum. Mol. Genet.* **23**, 4703–4709 (2014).
25. Ros, X. R. & Vermeulen, L. Turning cold tumors hot by blocking TGF- β . *Trends Cancer* **4**, 335–337 (2018).
26. Haanen, J. B. A. G. Converting cold into hot tumors by combining immunotherapies. *Cell* **170**, 1055–1056 (2017).
27. Chen, D. S. & Mellman, I. Elements of cancer immunity and the cancer-immune set point. *Nature* **541**, 321–330 (2017).
28. Peranzoni, E., Lemoine, J., Vimeux, L., Feuillet, V., Barrin, S., Kantari-Mimoun, C. et al. Macrophages impede CD8 T cells from reaching tumor cells and limit the efficacy of anti-PD-1 treatment. *Proc. Natl Acad. Sci. USA* **115**, E4041–E4050 (2018).
29. Joyce, J. A. & Fearon, D. T. T cell exclusion, immune privilege, and the tumor microenvironment. *Science* **348**, 74–80 (2015).
30. Sato, E., Olson, S. H., 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. USA* **102**, 18538–18543 (2005).
31. Montfort, A., Pearce, O., Maniati, E., Vincent, B. G., Bixby, L., Böhm, S. et al. A strong B-cell response is part of the immune landscape in human high-grade serous ovarian metastases. *Clin. Cancer Res.* **23**, 250–262 (2017).
32. Bindea, G., Mlecnik, B., Tosolini, M., Kirilovsky, A., Waldner, M., Obenauf, A. C. et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity* **39**, 782–795 (2013).
33. Hu, G. & Wang, S. Tumor-infiltrating CD45RO. *Sci. Rep.* **7**, 10376 (2017).
34. Zhang, Q. W., Liu, L., Gong, C. Y., Shi, H. S., Zeng, Y. H., Wang, X. Z. et al. Prognostic significance of tumor-associated macrophages in solid tumor: a meta-analysis of the literature. *PLoS ONE* **7**, e50946 (2012).
35. Malesci, A., Bianchi, P., Celesti, G., Basso, G., Marchesi, F., Grizzi, F. et al. Tumor-associated macrophages and response to 5-fluorouracil adjuvant therapy in stage III colorectal cancer. *Oncoimmunology* **6**, e1342918 (2017).
36. Pinto, M. L., Rios, E., Duraes, C., Ribeiro, R., Machado, J. C., Mantovani, A. et al. The two faces of tumor-associated macrophages and their clinical significance in colorectal cancer. *Front Immunol.* **10**, 1875 (2019).
37. Yang, M., McKay, D., Pollard, J. W. & Lewis, C. E. Diverse functions of macrophages in different tumor microenvironments. *Cancer Res.* **78**, 5492–5503 (2018).
38. Bohm, S., Montfort, A., Pearce, O. M., Topping, J., Chakravarty, P., Everitt, G. L. et al. Neoadjuvant chemotherapy modulates the immune microenvironment in metastases of tubo-ovarian high-grade serous carcinoma. *Clin. Cancer Res.* **22**, 3025–3036 (2016).
39. Mansuet-Lupo, A., Alifano, M., Pécuchet, N., Biton, J., Becht, E., Goc, J. et al. Intratumoral immune cell densities are associated with lung adenocarcinoma gene alterations. *Am. J. Respir. Crit. Care Med.* **194**, 1403–1412 (2016).
40. Ham, S. W., Jeon, H. Y., Jin, X., Kim, E. J., Kim, J. K., Shin, Y. J. et al. TP53 gain-of-function mutation promotes inflammation in glioblastoma. *Cell Death Differ.* **26**, 409–425 (2019).
41. Biton, J., Mansuet-Lupo, A., Pecuchet, N., Alifano, M., Ouakrim, H., Arrondeau, J. et al. TP53, STK11, and EGFR mutations predict tumor immune profile and the response to anti-PD-1 in lung adenocarcinoma. *Clin. Cancer Res.* **24**, 5710–5723 (2018).
42. Dong, Y., Richards, J. A., Gupta, R., Aung, P. P., Emley, A., Kluger, Y. et al. PTEN functions as a melanoma tumor suppressor by promoting host immune response. *Oncogene* **33**, 4632–4642 (2014).
43. Clarke, B., Tinker, A. V., Lee, C. H., Subramanian, S., van de Rijn, M., Turbin, D. et al. Intraepithelial T cells and prognosis in ovarian carcinoma: novel associations with stage, tumor type, and BRCA1 loss. *Mod. Pathol.* **22**, 393–402 (2009).



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

