The IL6/JAK/STAT3 pathway as a therapeutic target for histological subtypes of colorectal cancer

Kathryn AF Pennel

BSc (Hons), MSc

Submitted in fulfilment of the requirements for the degree of PhD

Institute of Cancer Sciences

College of Medical, Veterinary and Life Sciences

June 2021

The work presented in this thesis was performed by the author except where acknowledged. This thesis has not been submitted for a degree or diploma at this or any other institution.

Kathryn Pennel

June 2021

Acknowledgments

Firstly, thank you to my primary supervisor, Professor Joanne Edwards. I could not have asked for a more supportive, kind, warm and knowledgeable mentor. I have thoroughly enjoyed my time in her lab and the countless opportunities she has guided me through. I am incredibly grateful to have had Joanne as my supervisor.

Thank you also to my second supervisor, Dr Antonia Roseweir, for all her guidance and patience. I am hugely appreciative of her help throughout, particularly with teaching me histology and statistical analyses.

I really can't thank Dr Jean Quinn enough. I have learned so much from her and couldn't have got through my PhD without her mentorship. From coffee in the Wolfson Wohl corridor to Panda Express in New York, Jean has made my time during my PhD better.

Thank you to my reviewers Mr Campbell Roxburgh and Mr James Park for the support with clinical aspects of my PhD, and for introducing me to Mr Josh Smith to whom I'm grateful for visiting his lab to gain invaluable insight into organoid culture.

Thank you to my lab partner, Phungern Khongthong for her friendship, enthusiasm, and making me spicy lunches. I would also like to thank other members of Professor Edwards lab, past and present, for fostering a friendly, positive environment to work in. A special thank you to Dr Gerard Lynch and Dr Akhill Yedursi for their assistance with bioinformatics, and Sara-Al Badran for help with lab work and digital pathology. Thank you to Liz Morrow for giving me a caffeine addiction. Thank you to Chris Bigley for help with IT issues, and to Liam Hayman for keeping my ego in check.

Thank you to all the NHS GGC patients for providing tissue for the studies, and to the Glasgow Biorepository, and Clare Orange. Thank you to the Glasgow Tissue Research Facility and Dr Jennifer Hay for giving me an opportunity to use the TMA grandmaster. Thank you also to Colin Nixon from CRUK Beatson Institute for his assistance with histology.

Finally, many thanks to all my family and friends for helping me along the way. Thank you to Dr Hannah Jane Lawson for being 3 steps ahead of me in all walks of life and providing an endless source of humour and support. Thank you also to Paddy, you're a gem, you're a star, you raise the bar. Many thanks to my Swiss Nana Ursula, who is an inspiration at 92, having completed her PhD in 1956 with a year abroad spent at the University of Glasgow! Thanks to my brother, Johnny, I'll keep it short. Thank you to my mum for her unwavering support, guidance, phone calls, and lessons in cleanliness. Thank you to my dad for always being there, 'vlookup' tutorials, and liking my tweets.

Thank you to anyone else who has helped me along the way.

Summary

Colorectal cancer remains the 2nd most common cause of cancer-related mortality worldwide and 5-year survival is only ~60%. The current method of staging disease, TNM staging, was devised in the 1940s, and it is recognised this could be improved (1). Research has focused on identifying new ways to segregate disease for prognostic information, and to predict response to different types of therapy. This culminated in the development of the consensus molecular subtypes in 2015 from an international consortium (2). The CMS consist of 4 groups: immune, canonical, metabolic, and mesenchymal, based on transcriptomic and mutational profiles. CMS are devised from bulk sequencing data and are not solely focused on the tumour compartment. To combat this, the cancer cell intrinsic subtypes, CRIS, were developed in 2017, based specifically on the underlying biology of tumour cells, to classify patients into 1 of 5 groups, CRISA-E (3-5). Despite the advancement in our understanding of patient heterogeneity and prognostic information gained from CMS and CRIS, both methods require complex and costly laboratory processes not yet feasible for routine diagnostics.

A more clinically relevant tool for segregating patient disease, Glasgow Microenvironment Score (GMS), was devised by Park et al in 2015, which only requires a single H&E-stained resection (5-7). Tumours are assessed for level of inflammatory infiltrate via Klintrup-Mäkinen grading and stromal invasion by tumour stroma percentage, to form 3 independently prognostic groups: GMS0 immune, GMS1 intermediate, and GMS2 stromal (6, 7). Patients in the immune group observe the best prognosis in terms of cancer-specific survival (CSS), and those with stromal-rich tumours have the worst outcomes (6). In 2017 Roseweir et al built on the GMS through phenotypic subtyping, which splits GMS1 patients into 2 groups based on Ki67 proliferation index. This resulted in formation of immune, canonical, latent and stromal histological groups (8). These two subtyping methods are independently prognostic, could easily translate to clinical practise and therefore represent exciting strategies for segregating disease for prognosis and ultimately guiding treatment options. There is currently limited evidence as to how GMS or phenotypic subtypes could be utilised to predict response to specific therapies, and the underlying biology driving phenotypes is unknown.

This thesis aimed to address this firstly by validating histological subtyping in stage I-IV retrospective CRC cohorts, and exploring a potential mechanism driving phenotypic differences observed in GMS and phenotypic subtypes. It could be hypothesised that the observed histologic differences are driven by cancer-associated dysregulation of specific cellular signalling pathways. IL6/JAK/STAT3 represents one such signalling pathway implicated in CRC development and progression in the literature (9). To investigate this, a CRC tissue microarray (n=1030) with GMS/phenotypic subtyping data was stained via IHC for members of IL6/JAK/STAT3 signalling. Staining was scored semi-quantitatively and related to outcome, clinicopathological features and tumour phenotype. High

expression of key pathway members, IL6R, JAK2 and pSTAT3^{tyr705} were associated with significantly reduced CSS, particularly in GMS2 stromal subtyped patients.

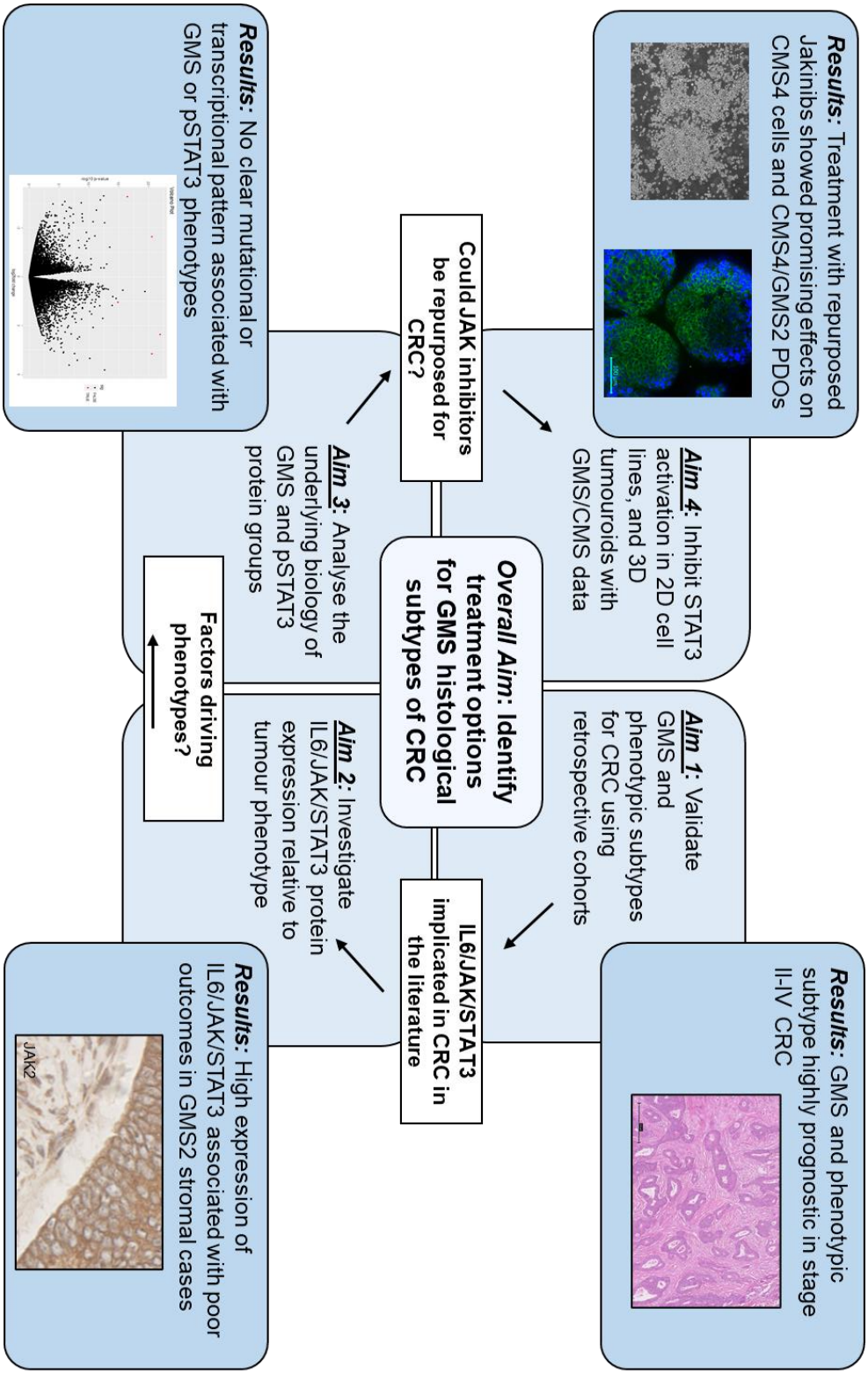
Next, mutational (n=285), and transcriptional profiling (n=100) were performed on a subset of these patients to investigate the underlying biology driving GMS phenotypes and activated STAT3 at the protein level. There were no clear mutational patterns driving GMS, however patients with high pSTAT3^{tyr705} protein observed a significantly higher frequency of MAP2K mutations. In terms of transcriptomics there were no differentially expressed genes associated with pSTAT3^{tyr705} phenotypes, however 5 differentially expressed genes were observed in GMS1 compared to GMS0/2 patients. Gene set enrichment analysis identified a strong inflammatory component to these genes that were downregulated in GMS1 patients. These data represent a step towards understanding the mechanisms driving tumour phenotypes, however the study was underpowered and larger cohorts should be utilised to validate findings and uncover other genomic/transcriptomic patterns.

Data from the IHC studies highlighted IL6/JAK/STAT3 as a promising therapeutic target, specifically in patients with stromal-rich tumours (GMS2). There are already a number of drugs which target intermediate pathway members (JAK1/2/3) approved for use in inflammatory disorders such as rheumatoid arthritis (10). The next aim of this thesis was to perform preliminary drug screening experiments using 2 repurposed Jakinibs, Ruxolitinib and Tofacitinib, in 7 distinct 2D CRC cell lines. HCT116, SW620 and SW480 lines showed the highest reduction in cell viability following treatment, which was of interest as all 3 are classified as stromal-rich CMS4 via their transcriptomic profiles (11).

It is now widely accepted that 2D cell lines are not representative of patient disease, therefore 3D patient-derived tumouroids were established and used to screen JAK inhibitors. H&E and Ki67 IHC was performed on the diagnostic specimen from which organoids were devised and utilised to establish GMS and phenotypic subtypes to assess subtype relationship to drug response. The organoid lines observed to have the greatest reduction in cell viability following JAK inhibition were from organoids derived from GMS2/Stromal tumours. This observation requires further validation in a larger independent cohort of organoids, and future work should include developing co-culture models to further recapitulate human disease.

In summary, data from this thesis have validated histological scoring of CRC and identified a role for IL6/JAK/STAT3 signal transduction in stromal dense tumours. Preliminary data from preclinical cell line and organoid work suggested that inhibiting STAT3 activation may be a promising, novel therapeutic approach for GMS2/stromal subtyped patients. These GMS2 patients observe the worst clinical outcomes and therefore require new treatment options. Ruxolitinib and Tofacitinib are already FDA approved for inflammatory disorders, and therefore they may be easily translated into clinical trials for CRC should validation work prove successful.

Graphical Abstract



Publications and presentations

Publications relating to this thesis

Pennel, K.A., Park, J.H., McMillan, D.C., Roseweir, A.K. and Edwards, J., 2019. Signal interaction between the tumour and inflammatory cells in patients with gastrointestinal cancer: Implications for treatment. Cellular signalling, 54, pp.81-90.

Alexander, P.G., Roseweir, A.K., **Pennel, K.A**., van Wyk, H.C., Powell, A.G., McMillan, D.C., Horgan, P.G., Kelly, C., Hay, J., Sansom, O. and Harkin, A., (2020). The Glasgow Microenvironment Score associates with prognosis and adjuvant chemotherapy response in colorectal cancer. British Journal of Cancer, pp.1-11

Al-Badran, S.S., Grant, L., Campo, M.V., Inthagard, J., **Pennel, K**., Quinn, J., Konanahalli, P., Hayman, L., Horgan, P.G., McMillan, D.C. and Roxburgh, C.S., 2020. Relationship between immune checkpoint proteins, tumour microenvironment characteristics, and prognosis in primary operable colorectal cancer. The Journal of Pathology: Clinical Research

Knight, J.R., Alexandrou, C., Skalka, G.L., Vlahov, N., **Pennel, K**., Officer, L., Teodosio, A., Kanellos, G., Gay, D.M., May-Wilson, S. and Smith, E.M., 2020. MNK inhibition sensitizes KRAS-mutant colorectal cancer to mTORC1 inhibition by reducing eIF4E phosphorylation and c-MYC expression. Cancer Discovery.

Submitted to British Journal of Cancer June 2021- **Kathryn AF Pennel***, Colin S Wood*, Xabier Cortes-Lavaud, Jean A Quinn, Susanna Cooke, Antonia K Roseweir, James H Park, Campbell SD Roxburgh, Donald C McMillan, Andrew V Biankin, Owen J Sansom, Nigel B Jamieson, Paul G Horgan, Colin W Steele, Joanne Edwards. Histological scores predict time to recurrence and overall survival in patients undergoing curative-intent surgical treatment of synchronous colorectal liver metastases.

Submitted to European Journal of Cancer June 2021- P.G. Alexander*, A.A.M. Matly*, N. Jirapongwattana, **K.A.F. Pennel**, H.C. van Wyk, D.C. McMillan, P.G. Horgan, C.S.D. Roxburgh, C. Thuwajit, A.K. Roseweir, J. Quinn, J.H. Park, J. Edwards. The relationship between the Glasgow Microenvironment Score and Markers of Epithelial-to-Mesenchymal Transition in TNM II-III Colorectal Cancer

In preparation- Patel M, **Pennel KAF**, Quinn JA, Hood H, Cooke S, Biankin A, Horgan PG, McMillan DC, Edwards J. Spatial expression of IKK-alpha is associated with a differential mutational landscape and survival in primary colorectal cancer

In preparation- **Pennel KAF**, Al-Badran S, Quinn JA, Hay J, Nixon C, Anderson D, Cooke S, Biankin A, Horgan PG, McMillan DC, Park JH, Roxburgh CSD, Sansom O, Roseweir AK, Edwards

J. IL6/JAK/STAT3 signalling as a novel, repurposable therapeutic target for histological subtypes of colorectal cancer.

Oral presentations

Phenotype dictates outcome following synchronous resection of primary colorectal cancer and liver metastases. **Kathryn Pennel**, Colin Steele, Antonia Roseweir, Rene Jackstadt, Colin Nixon, Xabier Cortes-Lavaud, Campbell Roxburgh, Hester van Wyk, Donald McMillan, Paul Horgan, Owen Sansom, Joanne Edwards. National Cancer Research Institute conference 2019 – Glasgow, UK

Phenotypic subtyping of matched primary colonic tumours and liver metastases. **Kathryn Pennel**, Arfon Powell, Donald McMillan, Paul Horgan, Antonia Roseweir, Joanne Edwards. National Cancer Research Institute conference 2018 – Glasgow, UK

Poster presentations

Effect of phenotype on outcome in synchronously resected primary colorectal cancer and matched liver metastases. **Kathryn AF Pennel***, Colin W Steele*, Jean A Quinn, Antonia K Roseweir, Rene Jackstadt, Colin Nixon, Xabier Cortes-Lavaud, Matthew Neilson, Hester van Wyk, Campbell SD Roxburgh, Donald C McMillan, Paul G Horgan, Owen J Sansom and Joanne Edwards. ASCO Gastrointestinal Symposium 2020 – San Francisco, CA, USA

Phenotype dictates outcome following synchronous resection of primary colorectal cancer and liver metastases **Kathryn Pennel**, Colin W Steele, Hester van Wyk, Antonia Roseweir, Rene Jackstadt, Colin Nixon, Xabier Cortes-Lavaud, Campbell SD Roxburgh1, Donald C McMillan, Paul G Horgan, Owen J Sansom, Joanne Edwards. British Association of Cancer Research Tumour Microenvironment conference 2019 – Nottingham, UK

Phenotypic subtyping of primary colonic tumours and matched liver metastases. **Kathryn AF Pennel**, Donald C McMillan, Arfon G Powel, Antonia K Roseweir and Joanne Edwards, British Association of Cancer Research Student Conference 2018 – London, UK

Contents

List of Figures

List of Tables

Abbreviations

ADAM10 A disintegrin and metalloproteinases 10

ADAM17 A disintegrin and metalloproteinases 17

ADDMEM Advanced

AKT Protein kinase B

APC Adenomatous polyposis coli

α-SMA Alpha smooth muscle actin

BCL2 B-cell lymphoma 2

BME Basement membrane extract

BRAF v-raf murine sarcoma viral oncogene homolog B1

CAFs cancer-associated fibroblasts

CAPOX oxaliplatin & capecitabine

CASP7 Caspase 7

CI Confidence interval

CIMP CpG Island methylation

CIN Chromosomal instability

CMS Consensus molecular subtype

CRC Colorectal cancer

CRIS Cancer cell intrinsic subtype

CRUK Cancer research UK

CSS Cancer specific survival

cRT chem-radiotherapy

CTLA4 Cytotoxic T-Lymphocyte Associated Protein 4

CXCR3 C-X-C Motif Chemokine Receptor 3

DAB 3,3'-diaminobenzidine

DFS Disease-free survival

DMSO Dimethyl sulfoxide

DNA Deoxyribonucleic acid

DO Deterector oligonucleotides

DSS Dextran sodium sulphate

EDTA Ethylenediaminetetraacetic acid

EGFR Epidermal growth factor receptor

ELISA Enzyme-linked immunosorbent assay

ERK Extracellular signal-regulated kinase

FAK Focal adhesion kinase

FAP Familial adenomatous polyposis

FBS Fetal bovine serum

FDA Food and drug administration

FFPE Formalin fixed paraffin embedded

FOBT Faecal occult blood test

FOLFIRI Folinic acid, fluorouracil, and irinotecan

FOLFOX Fluorouracil, Folinic Acid (modified de Gramont) and. Oxaliplatin

NHS GGC National health service grater Glasgow and Clyde

GMS Glasgow Microenvironment Score

GPOL Glasgow precision oncology laboratory

GTRF Glasgow tissue research facility

HCl Hydrogen chloride

H&E Haematoxylin and Eosin

HeLa Henrietta Lacks

HNPCC Hereditary nonpolyposis cancer

H_2O_2 Hydrogen peroxide

HR Hazard ratio

HRP Horse radish peroxidase

ICCC Intra class correlation coefficient

IF Immunofluorescence

IGF2 Insulin Like Growth Factor 2

IHC Immunohistochemistry

IL6 Interleukin 6

IL8 Interleukin 8

IL6R Interleukin 6 receptor

JAK1 Janus Kinase 1

JAK2 Janus Kinase 2

JAK3 Janus Kinase 3

KM Klintrup-Mäkinen

KRAS Ki-ras2 Kirsten rat sarcoma viral oncogene homolog

MAP2K Mitogen activated protein kinase

MCL1 Myeloid cell leukemia sequence 1

mCRC metastatic colorectal cancer

mGPS modified Glasgow prognostic score

MMP9 Matrix Metallopeptidase 9

MMR Mismatch repair

MSCs Mesenchymal stromal cell

MSI Microsatellite instability

MSS Microsatellite stability

mTORC1 mechanistic target of rapamycin complex 1

MYC Master Regulator of Cell Cycle Entry and Proliferative Metabolism

NaH Sodium hydroxide

NFκB Nuclear Factor kappa B

NLR Neutrophil to lymphocyte ratio

Padj Adjusted p value

PBS Phosphate buffered saline

PCR Polymerase chain reaction

PD1 Programmed death protein 1

PDL1 Programmed death ligand 1

PDO Patient derived organoid

PFA Paraformaldehyde

PIK3CA Phosphatidylinositol-4,5-bisphosphate 3-kinase

PKR protein kinase R

PVDF polyvinylidene difluoride

RA Rheumatoid arthritis

RAF Rapidly Accelerated Fibrosarcoma

RAS Rat sarcoma

RNA Ribonucleic acid

ROC Receiver operator curve

ROS Reactive oxygen species

RPM Revolutions per minute

RT Radiotherapy

SNV Single-nucleotide variants

STAT3 Signal transducer and activator of transcription 3

TAMs Tumour associated macrophage

TANs Tumour associated Neutrophil

TBST Tris buffer saline solution plus tween

TGFβ Transforming growth factor beta

TMA Tissue microarray

TNM Tumour node metastases

TP53 tumour protein (EC :2.7.1.37)

TRIS Tris buffered saline

TSP Tumour stroma percentage

TYK2 Tyrosine kinase 2

VEGFR Vascular Endothelial Growth Factor Receptor

WNT Wingless-related integration site

WST-1 Water-soluble tetrazolium salt 1

5FU fluorouracil

Chapter 1: Introduction

1.1 Colorectal cancer incidence

Colorectal cancer (CRC) is the 2nd most common cause of cancer-related death worldwide and incidences are still rising amongst low- and middle-income countries (12). In the UK, overall incidence of CRC has remained relatively stable between 2002 and 2014, at around 8.37 per 100000 in 2014 (13). In Scotland, between the years of 1997-2017, there were 77,262 CRC diagnoses (14). The introduction of the Scottish Bowel Screening programme in 2007 resulted in adults aged between 50-74 being offered a faecal occult blood test (FOBT) once every two years to improve earlier detection (15). Uptake of FOBT is approximately 55% of individuals offered (16). Since screening was introduced, overall incidence has fallen however rates amongst young people (<50 years old) have increased from 5.3 per 100,000 in 2000 to 6.8 per 100,000 in 2017 (14). The overall 5-year survival rate for patients with CRC is ~60%, however this varies with respect to stage at time of detection (17).

1.2 CRC development and pathogenesis

Several genetic and environmental factors cause an individual to be more at risk of developing CRC. Modifiable risk factors for CRC include high body mass index, high consumption of red meats, low physical activity, age, smoking status, and poor diet (17, 18). Non-modifiable risk factors include comorbidities such as inflammatory bowel disease and genetic predisposition, namely hereditary CRC. There are two main types of hereditary CRC: familial adenomatous polyposis and hereditary non-polyposis colorectal cancer (19).

1.2.1 Hereditary CRC

1.2.1.1 Familial Adenomatous Polyposis

Familial Adenomatous Polyposis (FAP) is autosomal dominant and occurs as a result of inherited germline truncating mutations in the adenomatous polyposis coli (APC) gene on chromosome 5q21 (19). Patients with FAP have a 100% lifetime risk of developing CRC and make up less than 1% of total CRC cases (19).

1.2.1.2 Hereditary non-polyposis colorectal cancer

Hereditary non-polyposis colorectal cancer (HNPCC) or lynch syndrome is an autosomal dominant condition that arises as a result of germline mutation in at least one of 4 mismatch repair genes (MMR) genes; MLH1, MSH6, PMS2, MSH2 (20). HNPCC accounts for 1-3% of total CRC cases worldwide (20).

1.2.2 Sporadic CRC

Colorectal cancers occur sporadically in over 95% of cases. CRC is heterogeneous and the mechanisms of pathogenesis are complex. Development and progression are characterised by 3 main pathways of genetic instability, including the chromosomal instability pathway (CIN), CpG island methylation and microsatellite instability.

1.2.3 Signet ring CRC

Signet ring carcinoma is an extremely rare form of CRC that accounts for <1% of cases (21). It is characterised by distinct ring-like pathology whereby mucinous deposits push the tumour cell nuclei outwards (21). Signet ring carcinoma confers very poor prognosis, in part due to diagnoses at late-stage disease because of late-onset symptoms (21).

1.2.4 Chromosomal Instability Pathway

Approximately 65-75% of all CRC cases are thought to occur via the CIN pathway (22). The multi-step process was devised by Vogelstein in 1990 and the first stage of the pathway involves inactivation of the APC gene (Figure 1.1). APC loss of function is associated with formation of early adenoma (23) (Figure 1.1). This is followed by activating mutations in the KRAS, which leads to larger adenoma formation (23) (Figure 1.1). Accumulation of alterations in TGFß, PIK3CA and TP53 signalling pathways ultimately cause development of carcinoma (23) (Figure 1.1).

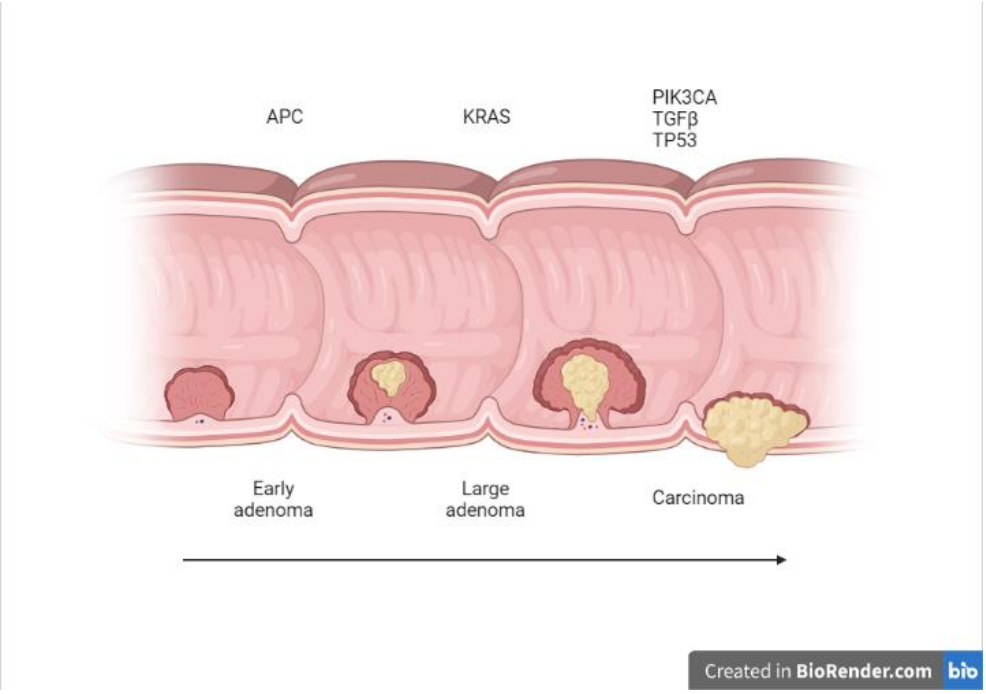


Figure 1.1 Chromosomal instability pathway. *Diagram showing sequence of mutation accumulation resulting in formation of adenoma to carcinoma in the Vogelstein model of CRC development. Loss-of function mutation in the APC gene leads to formation of an early adenoma. Activating mutation in the KRAS gene then subsequently leads to development of a large adenoma. Progression to carcinoma occurs after accumulation of PIK3CA, TGFβ and TP53 mutations.*

1.2.5 Microsatellite Instability

The normal process for repair of DNA replication errors is called mismatch repair (MMR). Microsatellite instability (MSI) pathway refers to alterations or deletions in DNA repeats caused by mutations in ≥1 MMR gene (MSH2, MSH6, MLH1, PMS1 and/or PMS2). The most common reason for MSI in CRC is epigenetic silencing of the MLH1 promoter. Diagnoses of MSI is performed via polymerase chain reaction using the Bethesda panel of genes: BAT-25, BAT-26, D2S123, D5S346 and D17S250 (24). Cancer with instability at ≥2 of these genes is considered MSI-h (24). IHC is used to detect MMR deficiency at the protein level by staining for MSH2, MLH1, MSH6, PMS1 and PMS2(25). Patients with MMR deficiency will show complete loss of protein expression of at least one of the four markers within the tumour cell nests (25). There is a high degree of concordance between MSI/dMMR cases identified by each technique (25). MSI-h/MMR deficiency leads to tumour characterised by mucinous, poorly differentiated phenotypes with high levels of tumour-infiltrating lymphocytes. This increase in immunogenicity is coupled with improved clinical outcome and improved response to immunotherapies (26). MSI accounts for 10-15% of sporadic CRC cases (27).

1.2.6 CpG island methylation

CpG (cytosine preceding guanine) islands are found in the promoter regions of ~50% of all genes (27). CpG dinucleotides tend to be methylated in normal cells and in cancer become unmethylated or hypermethylated (27). This results in promotion of tumourigenic genes and patients who exhibit these epigenetic changes are classified as CpG-island methylate phenotype (CIMP) (27). Approximately 30-40% of proximal colon cancer patients observe CpG methylation, and 3-12% of distal colon and rectal cancer patients display CIMP (28). Patients with this phenotype tend to have worse clinical outcomes when coupled with microsatellite stable (MSS) disease (29).

1.3 Colorectal cancer pathology

1.3.1 Anatomy

The bowel is a large tubular organ, which consists of connective tissue and muscular walls with an inner mucosal layer (30). Its main functions surround absorption of water and nutrients, and formation of stool. The most proximal portion of the colon is the cecum which is joined to the ascending colon (30) (Figure 1.2). At the hepatic flexure the ascending colon turns ~90 degrees to become the transverse colon (30) (Figure 1.2). This is generally the largest part of the bowel in length and terminates at the splenic flexure where at a 90 degree turn it becomes the descending colon (30) (Figure 1.2). The sigmoid section of the colon joins the descending colon and the rectum (30). The rectum is the terminal portion of the large intestine which is responsible for temporary storage of stool and connects to the anus (30) (Figure 1.2).

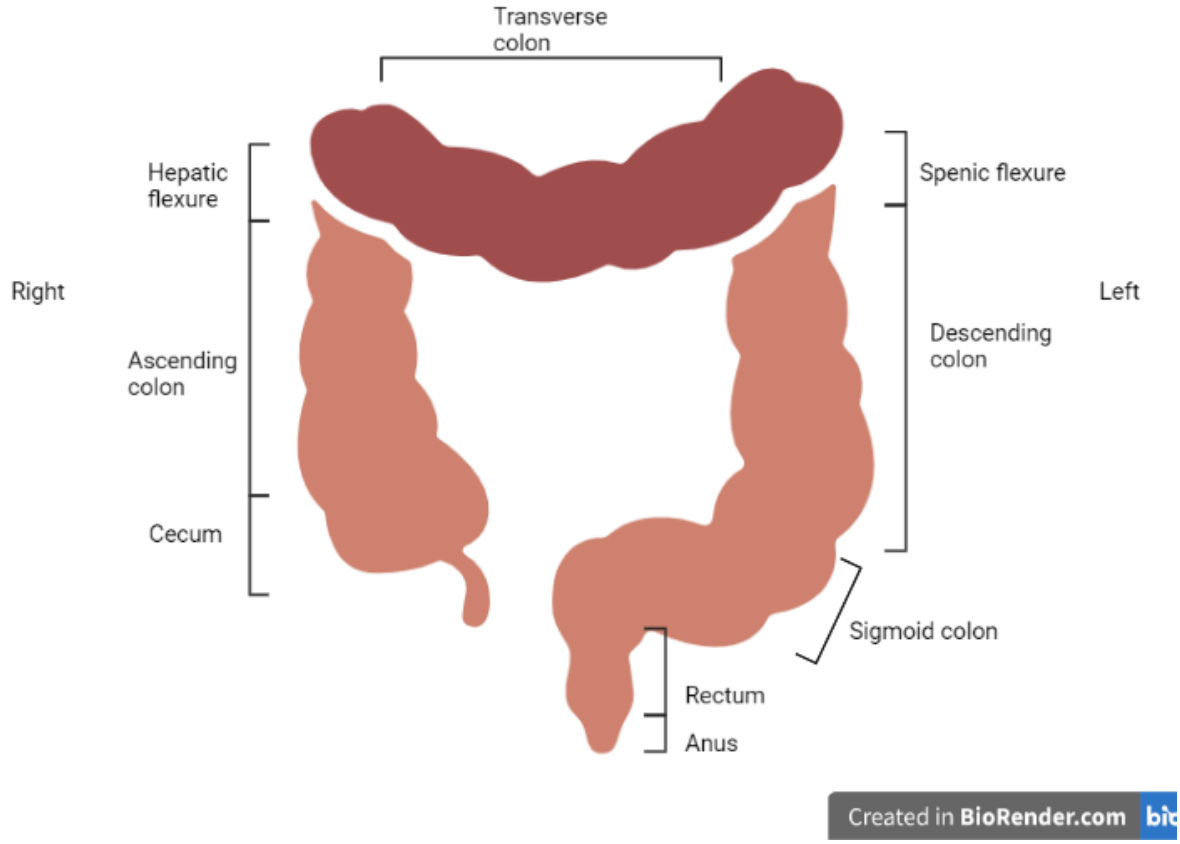


Figure 1.2 Anatomy of the bowel. *Illustrative diagram showing the different anatomical areas of the large intestine. The proximal part of the colon is the cecum which leads to the ascending colon and turns at the hepatic flexure to form the transverse colon. The splenic flexure joins the transverse colon to the descending colon, which ultimately leads to the sigmoid colon, rectum and anus.*

1.3.2 Sidedness

The anatomy of the bowel is important in terms of CRC, as there are notable differences in disease depending on the precise location of tumour formation. Broadly, CRC tumour location can be divided into 3 subclasses; right-sided, left-sided and rectal tumours (Figure 1.2). Right-sided tumours form in the ascending colon and proximal 2/3rds of the transverse colon. Left-sided tumours originate in the descending and sigmoid colon (31). Rectal tumours originate in the most distal part of the bowel, the rectum. Right-sided tumours tend to have the worst prognosis and are associated with unfavourable clinical characteristics such as increased systemic inflammatory response (32, 33).

1.3.3 TNM Staging

Tumour Node Metastases (TNM) staging is the current method used clinically for segregating disease, indicating how advanced the cancer is, predicting prognosis and guiding treatment regimens (34). T stage is determined by the level of tumour spread through the bowel wall, as outlined in Table 1.

Patients with no tumour detected are T0, when the tumour has grown into the submucosa tumours are classified as T1 and if the tumour has grown through the muscularis propria T2 is assigned. T3 occurs classification is when the tumour has grown into the subserosa or peritoneum and T4 is assigned if the tumour has invaded into different organs from the primary site, or perforation into the visceral peritoneum is present (Table 1.1). N stage is based on the presence of cancer within lymph nodes. If there are no lymph node metastases this is classified at N0, if there between 1 and 3 lymph nodes with metastases this is N1 and patients with ≥4 affected nodes are classified as N2 (Table 1.2). M stage is based on the absence (M0) or presence (M1) of any distant metastases (Table 1.3). The most common site of metastases for colorectal cancer is the liver. Overall TMM stage ranges from I-IV (Table 1.4). TNM stage I patients are T1N0M0 or T2N0M0 and TNM stage II are either T3N0M0 or T4N0M0 (Table 1.4). TNM stage III patients are classified as any T stage, N1 M0 or N2M0 (Table 1.4). Patients are categorised at TNM IV if there is presence of any metastases- Any T, Any N M1 (Table 1.4). Patients at TNM stage I have the best prognosis and stage IV patients observe the worst outcomes. TNM staging was first devised in the 1940s and is now considered to be outdated as there is no account for tumour heterogeneity and its inability to effectively guide therapeutics (22). More advanced prognostic measures and ways to predict optimal therapy regimes for patients needs to be translated into routine clinical practise.

T Stage	Clinical Description
T0	No primary tumour detected
T1	Tumour grown into submucosa
T2	Tumour grown into muscularis propria
T3	Tumour grown into subserosa or into peritoneum
T4	Tumour invaded into other organs or perforated through the visceral peritoneum

Table 1.1 T stage. *Table outlining the clinical description for each colorectal cancer T stage with reference to T0, T1, T2, T3 and T4 describing various stages of tumour growth through the bowel wall.*

N Stage	Clinical Description
N0	No lymph node metastases
N1	1-3 lymph nodes with metastases
N2	≥4 lymph nodes with metastases

Table 1.2 N stage. *Table outlining the clinical description for each colorectal cancer N stage. This is a classification system based on the number of lymph nodes with detectable metastases.*

M Stage	Clinical Description
0	No distant metastases present
1	Distant metastases present

Table 1.3 M stage. *Table outlining the clinical description for each colorectal cancer M stage, with M stage 0 patients showing no presence of distant metastases, and M stage 1 patients with distant metastases present.*

TNM Stage	Classification
I	T1 N0 M0, T2 N0 M0
II	T3 N0 M0, T4 N0 M0
III	Any T N1 M0, Any T N2 M0
IV	Any T Any N M1

Table 1.4 TNM stage. *Table outlining how TNM stages are determined based on the combination of T stage, N stage and M stage to form 4 prognostic groups; TNM stages I-IV.*

1.3.4 Dukes staging

Dukes staging is a similar method to TNM staging devised in 1932 specifically for rectal cancer, but is now known to also be prognostic in colon cancers (35). Dukes' stages include 4 patient groups: A, B, C and D (36). Dukes stage A is assigned if the patient's tumour is confined to the inner lining of the bowel and stage B occurs if the cancer has grown into the muscle wall. Dukes stage C is characterised by spread to lymph nodes and stage D encompasses patients with metastatic spread of the primary colorectal tumour(36). TNM staging is currently the preferred method, over Dukes, for grading disease in the clinic.

1.4 Current therapeutic approaches in CRC

1.4.1 Surgery

The primary treatment option for CRC patients with no distant metastases is surgical resection of the affected area. This can involve local resection for smaller early-stage tumours or can involve removal of larger parts of the bowel in procedures such as right or left hemi-colectomies. For rectal cancer there is evidence that total mesorectal excision can be beneficial in terms of preventing recurrence and improving prognosis (37).

1.4.2 Chemotherapy

In current clinical practise, TNM staining is utilised to determine which patients are offered chemotherapy. Regimes can be neoadjuvant (before surgical procedure) or adjuvant (post-surgical procedure). Patients with rectal cancer generally receive neoadjuvant short course radiotherapy or neoadjuvant long course chemoradiotherapy (cRT) and may receive additional adjuvant chemotherapy. Colon cancer patients are less likely to receive any radiotherapy but may be offered adjuvant chemotherapy if they have high risk TNM stage II or stage III disease. There is a lack of biomarkers used to guide which chemotherapy patients should receive and more research needs to focus on identifying better disease segregation methods to predict the best course of action.

The main chemotherapeutic agents utilised for CRC (stage III) are fluoropyrimidine with oxaliplatin (38). These can be administered intravenously as FOLFOX (bolus and infused fluorouracil with oxaliplatin) or orally via the CAPOX regime (capecitabine and oxaliplatin) (38). Chemotherapy is generally given over a 6-month period, however results from the recent SCOT non-inferiority clinical trial have provided evidence that 3-month regimens may be equally as beneficial and result in fewer adverse side effects (38). FOLFIRI (Folinic acid, fluorouracil, and irinotecan) is also being increasingly used for CRC, in the adjuvant setting for colon cases and both pre- and post-operatively for rectal cancer (39).

1.4.3 Radiotherapy

Radiotherapy is only used rarely for downsizing advanced colon cancer cases preoperatively, however, is often used as neoadjuvant therapy in rectal cancer for downstaging purposes and to improve surgical margins (40). The improvement in magnetic resonance imaging (MRI) for rectal cancer diagnoses and staging has assisted with identifying patients at more at risk of local recurrence and thus likely to benefit from preoperative radiotherapy (40). Other than TNM staging there are no biomarker methods to identify which patients will benefit most from preoperative cRT. There are two main Neoadjuvant courses of RT used for rectal cancer, short and long course. Short course or hypofractionation generally involves administration of 5 fractions of RT over 1-2 weeks followed by up to 12 weeks of 5 cycles of chemotherapy. Long course RT or hyper fractionation involves administration of 28 fractions over a 5–6-week period followed by up to 12 weeks of 4 cycles of chemotherapy. There is now evidence that cRT may provide benefit to the patient not only as a standalone treatment but may also prime aspects of the TME to produce a better response to targeted therapies.

1.4.4 Targeted therapies and immunotherapy

The use of biologic therapy for CRC is restricted to stage IV metastatic disease (mCRC). Epidermal growth factor receptor (EGFR) neutralising antibodies such as cetuximab are approved for use in stage IV KRAS wild-type patients (41). In clinical trials cetuximab stabilised disease in 32-52% of patients (41). Anti-angiogenic biologic Bevacisumab which targets vascular endothelial growth factor receptor (VEGFR) was approved for mCRC in 2004 and produces a survival benefit in patients who have their primary tumour resected (42). More recently there has been a focus on immunomodulatory checkpoint inhibitors which has culminated in the approval of programmed death cell protein 1 (PD1) inhibitors Pembrolizumab and Nivolumab for refractory MMR deficient mCRC in 2017 (43). Recent phase II clinical trials investigating combined inhibition of programmed death ligand 1 (PDL1) and Cytotoxic T lymphocyte-associated protein (CTLA4) with Durvalumab and Tremelimumab,

respectively, has demonstrated survival benefit in both MSS and MSI disease (44). There are ongoing phase II clinical trials (PRIME-RT, NCT04621370 and DUREC, NCT04293419) investigating the immune priming effect of radiotherapy to enhance response to Durvalumab in rectal cancer patients. There are also promising results from the PANDORA (NCT04083365) phase II study which is ongoing but has reported limited toxicity and good rates of complete pathological responses thus far (45). Although the advent of targeted therapies has profoundly revolutionised treatment of CRC, further research is required to translate biological therapy to precision medicine approaches for earlier TNM stage patients.

1.5 CRC subtyping

1.5.1 Genomic and Transcriptomic CRC subtypes

In order to segregate CRC to better predict prognosis and response to specific treatment regimens there has been a vast body of research focused on developing better subtypes of CRC. This culminated in the identification of the Consensus Molecular Subtypes (CMS) by Guinney and colleagues in 2015 (2). CMS were devised from a consortium of six independent research groups with transcriptomic data on over 3000 CRC patients resulting in the identification of four distinct prognostic groups. The main characteristics of the CMS subgroups are shown in table 1.5. CMS1, the immune subtype, is characterised by a strong inflammatory component, MMR proficient disease, and hypermutation. CMS2 canonical patients' tumours are characterised by MYC and WNT dysregulation and the CMS3 metabolic group have tumours with altered metabolism. CMS4, the mesenchymal subtype, is characterised by angiogenesis, stromal invasion and TGFß signalling. There is evidence that CMS can predict response to current therapies, for example CMS1 patients are most likely to respond to immunotherapy (Table 1.5). CMS2 patients respond best to anti-EGFR such as cetuximab and FOLFOX (46). CMS3 patients also respond to FOLFOX and CMS4 patients have shown most promising results when treated with FOLFIRI (46). CMS1 patients with strongly immunogenic tumours observe the best prognosis and patients classified as stromal CMS4 are likely to have poor prognosis (Table 1.5) (46). The development of the CMS was a huge step forward for CRC precision medicine and understanding of disease heterogeneity. However, the techniques required to determine an individual's CMS are expensive, time-consuming, and not feasible for routine diagnostics.

	CMS1 Immune	CMS2 Canonical	CMS3 Metabolic	CMS4 Mesenchymal
Characteristics	MSI-H BRAF ^{V600E} Strong immune component	TP53 Wnt, Myc, Src	KRAS Dysregulated metabolism	EMT, TGF-β CAFS
Prognosis	Best	Intermediate	Intermediate	Worst
Therapeutics	Immunotherapy	EGFR FOLFOX	FOLFOX	FOLFIRI

Table 1.5 CMS classifications of disease. *Table outlining the characteristics of the consensus molecular subtypes of colorectal cancer and relative survival predication. CMS1, the immune subtype is characterised by MSI-H disease, BRAF mutation, influx of immune cells, good prognosis and good response to immunotherapy. The canonical CMS2 subtype consists of patients with TP53, Wnt, Myc and Src mutations, and these patients have intermediate prognosis but show a good response to EGFR inhibitors and the FOLFOX chemotherapy regime. CMS3 Metabolic patients generally have KRAs mutations, dysregulated cellular metabolism, intermediate prognosis and benefit most from FOLFOX chemotherapy. The mesenchymal/stromal group are characterised by epithelial to mesenchymal transition, TGF-β, presence of CAFs, worst prognosis and respond preferentially to FOLFIRI chemotherapy.* Modified from Buikhuisen et al, 2020 (46)

Subsequently in 2017 the cancer-cell intrinsic subtypes (CRIS) were developed in patient-derived xenografts (PDX) and consist of 5 patient groups: CRIS-A-CRIS-E, as shown in Table 1.6 (3). PDXs are defined as models which involve transplanting human tissue (tumours) into mice. The authors argue that by utilising PDX models they were able to exclude stromal cell signatures and focus solely on tumour cell transcriptomics/genomics, by removing the murine stromal compartment from analyses. CRIS-A consists of tumours that are mainly either BRAF mutated and MSI, or KRAS mutated, and MSS. CRIS-B consists of tumours with dysregulated TGFß signalling, specific to the tumour compartment and therefore differs from CMS4 bulk data. The main characteristic of CRIS-C tumours is responsiveness to anti-EGFR therapy. CRIS-D tumours exhibit IGF2 overexpression and CRIS-E patients generally have KRAS mutated and Paneth cell-like phenotype tumours. Patients classified as CRIS-D observe the best clinical outcomes, and patients in the CRIS-B group have the worst survival outcomes. Like CMS, CRIS subtyping has some disadvantages, as the transcriptomic and genomic techniques are not yet used in routine diagnostics. Additionally, use of PDX models may have introduced bias into the signatures. However, CRIS subtyping is more likely than CMS to predict concordant results when multiple tumour samples are analysed from the same patient (4). CRIS focuses specifically on tumour epithelium areas and therefore combats intra-tumour heterogeneity better than CMS (4).

	CRIS-A	CRIS-B	CRIS-C	CRIS-D	CRIS-E
Characteristics	MSI-H or BRAF	TGFβ dysregulation in tumour	TP53	IGF2 overexpression	KRAS
Prognosis	Intermediate	Worst	Intermediate	Best	Intermediate
Therapeutics	-	-	Anti-EGFR	-	-

Table 1.6 CRIS classifications of disease. *Table outlining the characteristics of the cancer cell intrinsic subtypes of colorectal cancer and relative survival predication. CRIS-A classified patients are generally MSI-H or BRAF mutated with intermediate prognosis. CRIS-B patients observe dysregulated TGF-β signalling in the TME and have the worst outcomes. CRIS-C patients have mutations in TP53, intermediate prognosis and respond to EGFR therapy. CRIS-D categorised patients have overexpression of IGF2 and the best clinical outcomes. CRIS-E patients generally have KRAS mutations and intermediate prognosis.* Modified from Buikhuisen et al, 2020 (46)

1.5.2 Histological CRC subtypes

To identify more clinically translatable segregation of CRC disease, Park et al devised Glasgow Microenvironment scoring (GMS) in 2015 (6). Using a simple Haematoxylin and Eosin-stained section, patients can be assigned to one of three independently prognostic GMS groups based on inflammatory infiltrate and stromal invasion. GMS groups are outlined in Table 1.5 and include GMS0, immune, GMS1, intermediate and GMS2, stromal (Table 1.7). The inflammatory infiltrate is assessed by Klintrup Mäkinen grading (KM) which is an established pathological measure of the level of immune cell presence at the invasive margin (7). Stromal invasion is measured using tumour-stroma percentage (TSP) to assess the level of intra-tumour stroma (5). GMS had been validated in a number of cohorts, and more recently was found to associate with response to chemotherapeutic regime in the TransSCOT clinical trial cohort (47). Patients in the GMS0 groups had a significantly better clinical response to FOLFOX chemotherapy over CAPOX (47). The authors hypothesise that the dense inflammatory infiltrate of GMS0 tumours interferes with capecitabine metabolism, which reduces its cytotoxicity and thus harbours a reduced clinical effect (47). Further research is needed to validate these findings, understand the mechanisms of action, and identify other favourable treatment options for each GMS group.

KM Grade (Immune Infiltrate)	TSP (Stromal Invasion)	GMS	Prognosis
High (2-3)	Any	0	Good
Low (0-1)	<50%	1	Intermediate
Low (0-1)	>50%	2	Worst

Table 1.7 Glasgow Microenvironment Scoring. *Table outlining the components of Glasgow Microenvironment Scoring and relative prognosis of each subgroup. GMS0 is the high immune group, GMS2 is characterised by high intra-tumour stromal invasion and GMS1 is an intermediate group with low inflammation and low stromal invasion.*

In 2017, Roseweir et al, preformed a review of the CMS classification to assess how to histologically define the four groups. Interestingly, they identified similar phenotypic features to the GMS, including stromal and immune infiltration. However, they also identified growth rate as a key feature and therefore further built on GMS by segregating GMS1 patients into two intermediate groups based on Ki67 proliferation index (48). This gave rise to 4 phenotypic subtypes: Immune, Canonical, Latent and Stromal as shown in table 1.8 The immune subtype confers the best prognosis, followed by canonical, latent and stromal groups. It could be argued that GMS1 patients represent a diverse array of CRC disease phenotypes, and phenotypic subtyping is able to stratify this GMS1 such that the canonical and latent subtypes provide more accurate segregation of patient disease. Phenotypic subtypes are independently prognostic, predict risk of recurrence and could easily be translated into routine diagnostics (49). Furthermore, the immune subtype showed the same response to FOLFOX chemotherapy as GMS0, therefore, identifying targeted therapies for each of the other three phenotypic subtype groups represents a promising therapeutic strategy for CRC precision medicine going forward.

Phenotypic Subtype	KM Grade	TSP	Ki67	Prognosis
Immune	High	Any	Any	Best
Canonical	Low	Low	High	Intermediate
Latent	Low	Low	Low	Intermediate
Stromal	Low	High	Any	Worst

Table 1.8 Phenotypic subtyping. *Table outlining the components of phenotypic subtypes of colorectal cancer and relative prognosis of each subgroup. The immune subtype is characterised by high inflammatory infiltrate at the invasive edge and patients have the best prognosis. Patients in the canonical subtype have intermediate prognosis and tumours are immunologically cold, with low stroma but high proliferation. Latent patients are low for all 3 histological measures and have intermediate prognosis. The stromal subtype patients have a high stromal invasion and exhibit the worst clinical outcomes.*

Another histopathological score which has been extensively validated in the literature is the Immunoscore developed by Galon et al in 2012 (50). In contrast to GMS and phenotypic subtype, Immunscore relies on measurement of specific populations of immune cells via immunohistochemical staining. Tumours are graded based on presence of lymphocyte populations CD3+/CD45RO+, CD3+/CD8+ or CD45RO+/CD8+ cells within the tumour centre and at the invasive margin (50). The Immunoscore ranges from I0 to I4 based on density of cells per mm^2 and is highly prognostic for disease-free survival in CRC (50). Immunoscore was investigated in relation to disease-free survival in the IDEA clinical trial investigating stage II CRC patients receiving 3 or 6 months standard of care chemotherapy, CAPOX or FOLFOX regimens (51). Patients classified as I2-I4 (high T cell infiltrates) showed increased DFS with administration of FOLFOX over 6 months compared to 3 months treatment(51).

1.5.3 Inflammatory subsets in CRC

In addition to the two lymphocyte populations utilised in the Immunoscore there are numerous other inflammatory cell populations known to associate with prognosis (52). The immune system is extremely complex, and expression of markers used to define specific populations can be transient depending on the surrounding milieu. Broadly, immune cell types can be linked to pro-tumour actions and poor prognosis or anti-tumour and better prognosis. Pro-tumour populations promote an immunosuppressive microenvironment whereby tumour cells can evade the immune response (52).

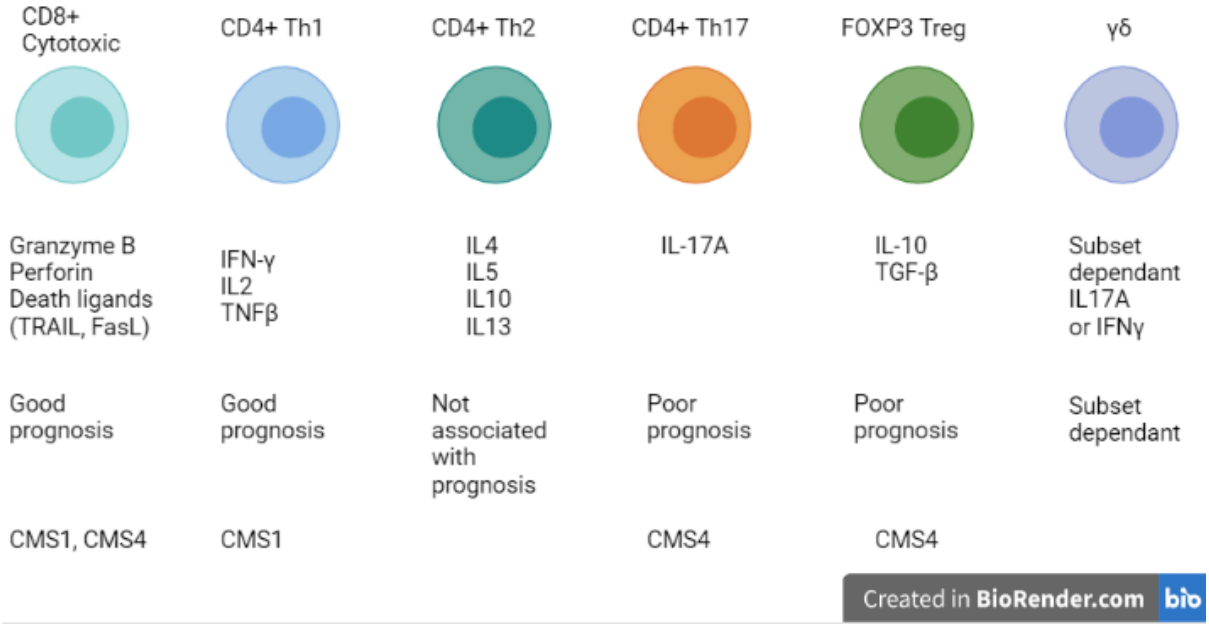


Figure 1.3. T-lymphocyte subsets in colorectal cancer. *Diagram showing the main subsets of T lymphocytes studied with regard to colorectal cancer, production of signalling molecules, effect on CRC prognosis and links with consensus molecular subtypes. The most well-studied subsets include CD8+ cytotoxic T cells, CD4+ Th1, CD4+ Th2, CD4+ Th17, FOXP3+ regulatory T cells and more recently investigated γδ T cells.*

The general role of most subsets of T lymphocytes in the tumour microenvironment are detailed in figure 1.3). As encompassed by Immunoscore, influx of CD3+ T cells to the TME is associated with good prognosis, and more specifically CD3+CD8+ cytotoxic T cells are strong drivers of anti-tumour responses (53). CD8+ T cells exert anti-tumour effects through production of granzyme B, perforin, and expression of death receptor ligands such as FasL and TRAIL. High levels of CD8+ T cells within the invasive margin and tumour centre (and overall CD3+ T cell infiltrates) are classified as Immunoscore 4 or I4. T helper cells with a Th1 phenotype are associated with good prognosis, and function through production of IFN-γ, IL2 and TNF-β. Cytotoxic T cells and Th1 cells are associated with influx to CMS1 classified tumours. Th2 cells are characterised by production of IL4, IL5, IL10, and IL13, and to date there has been no strong association with prognosis or CMS in CRC identified. In contrast, Th17 cells have shown a strong association with poor outcome in CRC and stromal-rich CMS4, functioning through production of IL-17A. FOXP3+ regulatory T cells are also generally associated with poor prognosis and CMS4. Tregs are regarded as a mainly immunosuppressive population dampening the cytotoxic nature of anti-tumour cells through production of IL10 and TGF-β (54). A significant increase in FOXP3+ expression amongst systemic CD4+ T cells has been noted in CRC patients compared to age-matched healthy controls (55). A less well-studied population of T cells, Gamma delta (γδ) T cells, role in the TME can be pro or anti-tumourigenic depending on the

specific subset. The presence of those which produce IL-17A are generally associated with worse clinicopathological characteristics and those which produce IFN-γ are anti-tumourigenic (56).

The innate arm of the immune system is also implicated in CRC progression and prognosis. Tumour-associated macrophages (TAMs) are generally polarized to a pro-tumour phenotype, characterised by high expression of CD163. Presence of high numbers of CD163+ TAMs at the invasive margin is associated with unfavourable clinicopathological characteristics and reduced recurrence-free survival (57). There are a variety of mechanisms through which TAMs promote tumour progression including expression of PDL1/CTLA4 leading to T cell suppression, secretion of anti-inflammatory cytokines TGF-β and IL10 and chemokines CCL5, CCL20 and CCL22, which recruit regulatory T cells to the TME and production of reactive oxygen species(58, 59). Tumour-associated Neutrophils are also generally polarised to a pro-tumour phenotype functioning through production of IL8, CCL2 and TNFα. Interestingly, recent studies in gastric cancer have identified crosstalk between MSCs and TANs within the TME, which results in promotion of angiogenesis and migration. MSCs produce IL6, which activates STAT3 and ERK1/2 in the TANs, which causes promotion of MSCs differentiation into pro-tumourigenic CAFs(60).

Natural killer cells (CD56+) are mainly regarded as functioning in as an anti-tumour cell type via their production of IFNγ, and cytolytic nature. However, infiltration to the tumour bed is relatively sparse with CD56+ cells only detectable (>4 cells/section) in about 30% of cases. It is hypothesised that their main function in the TME is working collaboratively with CD8+ T cells, and it has been shown that patients with high levels of infiltrating CD8+ T cells and NK cells observe good prognosis in CRC(61).

There is increasing evidence that the spatial orientation of immune populations relative to each other is of paramount importance in determining outcome. Development of multiplex IHC-based tissue studies has allowed for assessment of a vast number of different cell types on a single tumour section, such as the co-detection by indexing (CODEX) system. In a recent study it was shown that the presence of PD1+ CD4+ T cells was only associated with poor outcomes in cases where granulocytes were in close proximity(62). Another study looking at the spatial orientation of different subsets of immune cells found that the distance between myeloid populations (CD11b CD14, and CD11b CD15) and cytotoxic CD8s T cells was higher amongst MSI patients (63).

Single populations of inflammatory cells influence prognosis, but there is also strong evidence that in fact the density of the infiltrate is more predictive than its composition (7). The recent advances in the development of immunotherapies, highlight the importance of the immune system in CRC and the potential to harness inflammation to improve outcome (52).

1.5.4 Stromal cells and prognosis in CRC

In addition to the important role of the inflammatory infiltrate, stromal cell populations are also involved in shaping the tumour microenvironment. Stromal populations associated with CRC include cancer-associated fibroblasts (CAFs), endothelial cells, myofibroblasts and mesenchymal stem cells (MSCs) (64). CAFs can be derived from numerous stromal precursor cells and are defined as activated fibroblasts within or around the tumour. CAFs promote progression of CRC through a number of mechanisms including promotion of invasion and angiogenesis, modulating tumour cell metabolism and driving metastases (65). MSCs are phenotypically similar to CAFs and display a range of pro-tumour actions including inhibition of apoptosis, promotion of angiogenesis and metastases (64). MSCs promote an immunosuppressive tumour microenvironment (TME) through polarisation of macrophages to M2 phenotype and negatively affect CD8+ T cell cytotoxicity through expression of programmed death ligand 1 (PDL1) (64, 66). Taken together, these pro-tumour responses of stromal cells in the TME likely contribute to the profoundly poor prognosis seen in patients with stromal-rich GMS2/stromal phenotypic subtype tumours.

1.6 IL6/JAK/STAT3 in CRC

1.6.1 IL6/JAK/STAT3 signal transduction

It is now widely accepted that CRC heterogeneity can be driven by dysregulation of cellular signalling pathways (67). The IL6/JAK/STAT3 pathway is well characterised in inflammatory disorders and is thought to play a role in CRC development and progression in a subset of patients. Signal transducer and activator of transcription 3 (STAT3) is a member of a family of 7 STAT signalling molecules which include STAT1, STAT2, STAT3, STAT4, STAT5A, STAT5B and STAT6 (68). STAT3 is composed of two subunits, STAT3α and STAT3ß (9). STAT3 is constitutively active in the cancer setting and drives tumour progression via its downstream effects on gene transcription (68). Signal transduction occurs when IL6 binds its cognate receptor, IL6R (Figure 1.1). Classical signalling occurs when IL6R is membrane bound, however IL6R can also be found soluble in the cytoplasm, which can initiate *trans*-signalling (69). This *trans*-signalling cascade is enabled by cleavage of the α subunit of the IL6R by proteases ADAM10 and ADAM17 (9). Part of the receptor complex required for signal transduction, gp130, is ubiquitously expressed, which enables IL6 signalling in a huge range of cell types (69). The gp130 tails of IL6R do not have an intracellular kinase domain and so rely on activation of one of four JAK family members to transduce a signal. JAK1, JAK2, JAK3 and TYK2 can all be activated upon IL6 binding IL6R, which results in phosphorylation of STAT3 at tyrosine 705 (70). STAT3 isoforms then dimerise and translocate to the nucleus where STAT3 acts as

a master regulator of genes which promote the hallmarks of cancer (69). STAT3 can become fully activated by phosphorylation at serine 727 after activation at tyrosine 705 as part of post-translational modifications (71).

In addition to its role as a transcription factor, STAT3 is also implicated in metabolic processes, autophagy, and cell mobility (9). In terms of metabolism STAT3 activity promotes aerobic glycolysis, suppresses electron transport chain activity, and promotes the production of reactive oxygen species (ROS) (9). STAT3 can physically interact protein kinase R (PKR) blocking it from binding to autophagy inducer Eukaryotic initiation factor 4A-II (eif2A). Further anti-autophagy roles of STAT3 include promoting the expression of anti-autophagic genes BCL2 and MCL1 (9). Finally, STAT3 can negatively regulate microtubule formation and bind to focal adhesion kinase (FAK) and paxillin to affect cell motility (9).

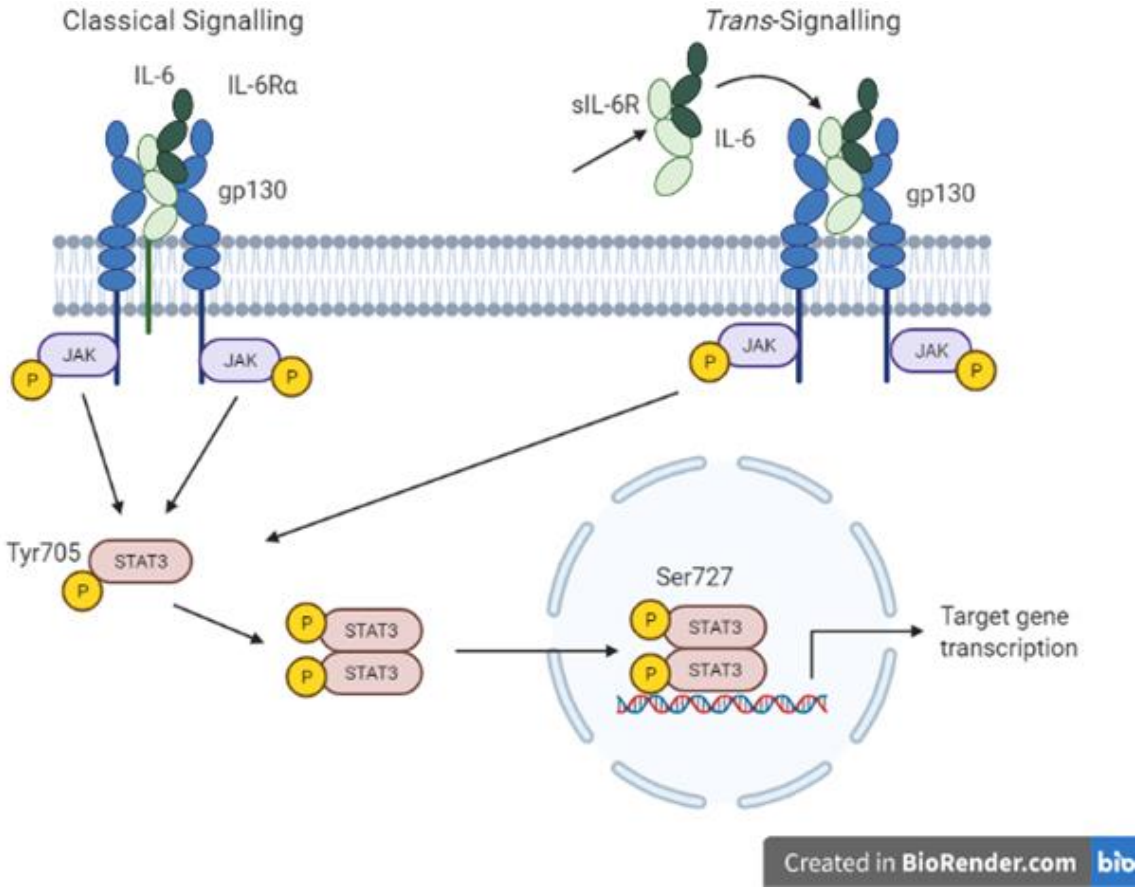


Figure 1.4. IL6/JAK/STAT3 signalling. *Schematic diagram showing IL6/JAK/STAT3 classical and trans-signal transduction. Classical signalling involves IL6 binding membrane bound IL6R initiating a cascade of signal transduction involving JAK1/2/3 or TYK2 activation, leading to phosphorylation of STAT3 at tyrosine 705 and/or serine 727, dimerization of STAT3 and translocation to the nucleus where it promotes expression of pro-tumourigenic gene. Trans-signalling involves IL6 binding to the soluble form of IL6Rwhich then bind to gp130 expressed on cellular membranes causing initiation of the same cascade.*

1.6.2 IL6/JAK/STAT3 signalling in cancer

Elevated systemic levels of IL6 are associated with unfavourable clinical characteristics and poor prognosis (72, 73). In retrospective tissue studies high expression of IL6 and STAT3 are associated with poor prognosis in solid tumours (74, 75). For example, in earlier work published from this lab, phosphorylated STAT3 was shown to associate with poor outcomes in patients with invasive ductal breast cancer (76). However, this pro-tumour role for STAT3 was not seen in an earlier study investigating castrate-resistant prostate cancer (77). IL6 can be produced by a plethora of cells in the TME, with links to inflammatory infiltrates, stromal cells and tumour cells (69). The most prominent source is likely myeloid-derived populations such as tumour associated macrophages (TAMs) or stromal cells such as cancer-associated fibroblasts (CAFs).

1.6.3 IL6/JAK/STAT3 in different compartments of the TME

STAT3 can be dysregulated amongst infiltrating immune cells as well as the tumour cells themselves. Hyperactivation of STAT3 amongst immune cells is associated with promotion of an immunosuppressive microenvironment, with enhanced regulatory T cell action and therefore cancer-promoting. In the tumour cells STAT3 signalling promotes proliferation, invasion, metastases, angiogenesis, and inhibition of apoptosis (78). STAT3 signalling can also actively induce IL6 production, resulting in a feed-forward autocrine feedback loop (79).

IL6 has been shown to directly promote CRC tumour cell proliferation, angiogenesis, survival, and invasion (Figure 1.2)(80, 81). IL6 also promotes tumour development and progression through its effects on specific populations of the inflammatory infiltrate. IL6 polarises macrophages to a pro-tumour phenotype which produce IL10 and show increased PDL1 expression(82). IL6 causes a reduction in TNFα, IL1 and IFN-γ production in tumour associated neutrophils. In terms of dendritic cells, IL6 causes decreased expression of major histocompatibility complex II, resulting in a decrease in antigen presenting, alongside a decrease in CD80/86 expression, decreased IL12 production and increased IL10 production(83). T cells are promoted to differentiate to a FOXP3+ regulatory phenotype by IL6, and cytotoxic CD8+ T cells observe decreased cytolytic activity, decrease chemotaxis to the TME via CXCR3 and increased PD1 expression in the presence of IL6(84, 85). In addition to the pro-tumourigenic effects on immune cells, IL6 can also assist tumourigenesis via promotion of CAF survival, migration, and collagen fibrogenesis(86). Data from the literature has clearly shown a vast array of pro-tumour functions of IL6 with evidence for roles in every compartment of the TME.

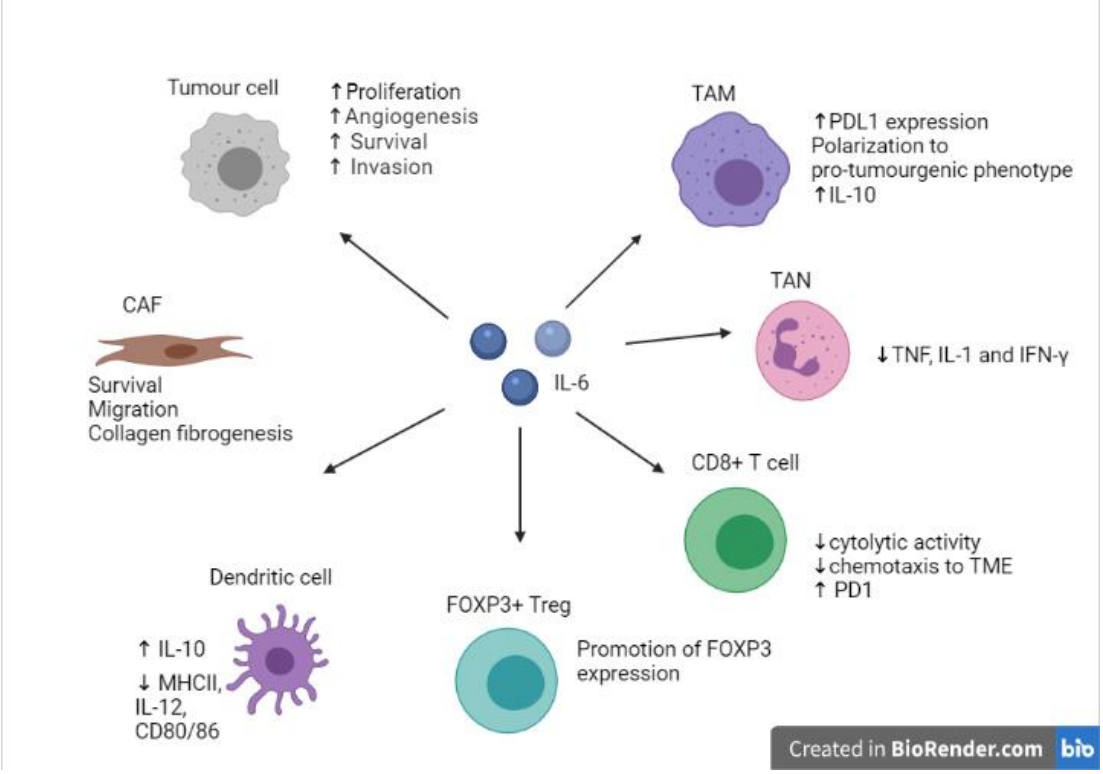


Figure 1.5 IL6 in the tumour microenvironment. *Diagram showing some of the main characterised roles for IL6 in colorectal cancer. IL6 promotes tumour progression through increasing tumour cell survival, angiogenesis, survival, and invasion. IL6 also acts on innate immune cells to increase production of anti-inflammatory cytokines such as IL10 by macrophages and decreases production of TNF-α, IL1 and IFN-γ by tumour-associated neutrophils. In terms of T cells, IL6 dampens the cytotoxic effect of CD8+ cells and promotes the expansion of regulatory T cells. IL6 acts on dendritic cells to cause increased production of IL10 and decreased expression of major histocompatibility complex II. IL6 also effects the surrounding stroma and causes increased survival, migration, and collagen fibrogenesis amongst CAF populations.*

1.6.4 IL6/JAK/STAT3 signalling pathway crosstalk

There is increasing evidence that cellular signalling pathways do not exist in isolation and crosstalk between pathways is common (67). One major signalling cascade associated with CRC development is NFκB signalling. STAT3 and NFκB promote transcription of overlapping cancer-promoting genes that cause proliferation, angiogenesis and production of pro-tumour chemokines and cytokines (87). Additionally, NFκB signalling is responsible for promoting IL6 production in myeloid populations (67). The two master regulators can also physically interact causing retention in the nucleus and thus continual transcription of downstream cancer-associated genes (87). Evidence from prostate cancer models have proven a link between STAT3 signalling with Akt and RAS pathways. RTK signalling via Epidermal growth factor receptor (EGFR) leads to synergistic activation of STAT3, RAS/RAF/ERK and Akt/mTORC1, which all converge to promote the hallmarks of cancer (88, 89).

1.7 Inhibition of IL6/JAK/STAT3 signalling

1.7.1 Inhibitors of IL6/IL6R

Inhibiting STAT3 itself poses challenges so initial efforts were focused on inhibiting upstream pathway components. Inhibition of IL6 binding to IL6R and therefore blocking signal transduction has been successful in certain inflammatory disorders. Tocilizumab is a potent IL6R inhibitor approved for clinical use in refractory Rheumatoid Arthritis (RA) in 2009 and systemic juvenile idiopathic arthritis and cytokine release syndrome in 2017 (90, 91). Siltuximab targets the cytokine itself, and was approved for RA patients in 2017 (92). Siltuximab monotherapy has reached phase I/II clinical trialling for KRAS mutant CRC and was well tolerated but no clinical effects were observed (93). In addition to activating STAT3, IL6 can also activate STAT1, which has opposing anti-tumourigenic functions to STAT3. Therefore, targeting IL6 or IL6R may not be the preferred route for inhibiting STAT3 clinically (69).

1.7.2 Inhibitors of JAK family members

Several small molecule competitive ATP inhibitors of JAK intermediate pathway members (Jakinibs) have been approved by the FDA for clinical use in settings other than cancer. JAK1/2 inhibitor, Ruxolitinib, was approved in 2011 for the treatment of Myelofibrosis. JAK2/3 inhibitor, Tofacitinib, was approved for use in RA and ulcerative colitis patients in 2012. Baficitinib which targets JAK1 and JAK2 is also approved for clinical use in RA (94). TYK2 inhibitor, Deucravasinib, is currently in phase II clinical trials for the treatment of psoriasis (53). Similarly, oral JAK3 inhibitor, PF-06651600, has entered phase IIb/III clinical trials for moderate/severe alopecia areata in a Pfizer sponsored study. One concern of JAK inhibition is the associated toxicity with neutropenia and opportunistic infection commonly noted side effects (9). Jakinibs have however produced excellent results in the cases of polycythaemia and myeloproliferative disorders and offer a promising repurposing option for subsets of many solid tumour types (9).

1.7.3 STAT3 inhibitors

There are challenges with targeting the STAT3 molecule specifically due to its similar chemical structure to STAT 1 and STAT5 (95). To combat this, STAT3 oligonucleotides have been designed for STAT3 to bind as a decoy preventing downstream transcription. In a preclinical xenograft model of head and neck cancer STAT3 oligos reduced growth and caused downregulation of STAT3 target genes (96). However, more recently a phase III clinical trial investigating the effects of first-in-class STAT3 inhibitor Napabucasin in advanced stage CRC found no survival benefit between placebo and treatment arms (97). Further research is required to determine the best way of targeting STAT3

signalling, and whether inhibiting upstream kinases/receptors/cytokines produces better effects than drugs which act on STAT3 itself.

1.8 Research aims and hypotheses

This thesis aimed to investigate IL6/JAK/STAT3 signalling in CRC and determine association between pathway activation with histological subtype. Targeted therapies in the form of small molecule inhibitors and biological therapies are showing increasing promise for combating disease heterogeneity associated with CRC. To investigate the rationale and therapeutic potential for inhibiting STAT3 activation in CRC patients the main objectives were as follows:

1.Use archival tissue from retrospective CRC patients to validate histological subtyping (Glasgow Microenvironment Score and Phenotypic Subtype) in 4 new cohorts ranging from stage I screen-detected CRC to stage IV metastatic disease.

2. Investigate protein expression of key pathway members of IL6/JAK/STAT3 inflammatory signalling pathway within tumour cells of retrospective stage I-IV CRC patient tissue microarrays and assess association with prognosis, clinicopathological features and histological subtypes to establish if targeting IL6/JAK/STAT3 has therapeutic potential.

3. Explore the underlying biology in a subset of these patients to determine any differences in genomic or transcriptomic profiles of patients across GMS subtypes and with upregulated phosphorylated STAT3 by mutation panel DNA sequencing and bulk RNA sequencing.

4. Determine the effects of inhibiting JAK1/2 and JAK2/3 to abrogate STAT3 activation in 2D CRC cell lines and 3D patient-derived organoid models to preliminarily assess therapeutic benefit relative to CMS/histological subtype of the original patient resections.

Chapter 2:
Materials and Methods

2.1 Tissue studies- Tumour Phenotyping

To investigate histological tumour phenotyping methods: Glasgow Microenvironment Scoring (GMS) and Phenotypic Subtyping in stage I-IV CRC, archival tumour resections were stained with Haematoxylin and Eosin (H&E) and for Ki67 by immunohistochemistry (IHC). Histological subtyping was performed in five retrospective CRC cohorts including the Glasgow combined cohort, TransSCOT clinical trial cohort, Australian TMA cohort, Synchronous resection cohort and DM-CRC-TMA screen detected CRC cohort. Utilising these cohorts allowed for validation of histological subtyping in a broad range of CRC stages and use of an international cohort enabled assessment of geographical distinct patients.

2.1.1 Cohorts

2.1.1.1 Glasgow combined cohort

The Glasgow combined cohort consisted of 1030 stage I-V CRC patients undergoing potentially curative resection across Greater Glasgow and Clyde (GGC) hospitals between 1997 and 2007. Tumours were staged with the 5th edition of TNM staging. Clinicopathological characteristics for each patient was available as part of an SPSS database. Clinical follow up data was last updated in 2017 from NHS GGC Safe Haven data. At this time, 324 patients (32%) had died of primary colorectal cancer, 332 patients (32.8%) had died of other causes and 355 patients (35.1%) were still alive. Survival data were missing for 19 patients. Cancer-specific survival (CSS), (date of surgery until last follow up) was used as a clinical endpoint throughout this study. Mean follow-up time was 139 months.

2.1.1.2 TransSCOT clinical trial cohort

The TransSCOT cohort consisted of 2912 CRC patients from the SCOT non-inferiority clinical trial who underwent surgical resection between 2008-2013. Tumours were staged with the 7th edition of TNM staging and patients were all classified as high-risk stage II or stage III. Patients in the trial received a random allocation of either 3- or 6- months of FOLFOX or CAPOX chemotherapy. The mean survival time of patients was 35.4764 months. At the time of writing patients were all followed up for at least 3 years and there were 2221 patients alive and 691 patients who had died of cancer.

2.1.1.3 Glasgow screening cohort

The DM-CRC-TMA cohort consisted of surgical resection CRC tissue from 159 T1/2 patients, screen-detected by faecal occult blood test (FOBT) under NHS GGC 2009-2011. Tumours were

staged with the 5th edition of TNM staging. Clinical follow up data was last obtained in 2014 and was added into an SPSS database along with other clinicopathological information. Median follow up time was 91 months. Of the 159 patients 16 died of the primary cancer and 13 had died of other causes. Due to the lack of events, overall survival was used as a clinical endpoint in this study.

2.1.1.4 Australian cohort

Tissue micro arrays were available via a collaboration with Prof C Soon Lee at Western Sydney University. A total of 410 Australian patients with primary operable colorectal cancer were included in the cohort. Full tumour resections were not available; however, TMA cores were taken from the invasive edge and the tumour centre to allow for TSP and KM grading. Survival data were only available for 156 patients and after 30-day mortalities were excluded this left 144 patients included in analysis. At the time of last follow up 47 (27.6%) of patients were still alive, 68 (40.0%) patients had died of cancer-related causes and 29 (17.1%) had died of other causes. Mean survival time was 39.412 months.

2.1.1.5 Synchronous resection cohort

The synchronous resection cohort consisted of patients (n=46) who had primary colorectal cancer tumours and metastatic liver metastases resected synchronously at Glasgow Royal Infirmary between 2002-2010. Approval for using patient tissue was obtained through NHS GGC Biorepository. Tumours were staged with the 5th edition of TNM staging. Clinical follow up data was updated in 2017 and available in a database with other clinicopathological characteristics. At the time of last follow-up 24 (40.0%) patients were alive, 30 had died of cancer-related causes (50.0%) and 3 patients (5.0%) had died of unrelated causes. Mean follow-up time was 40.14 months.

2.1.2 Tumour phenotyping

Full tumour sections (and TMAs in the case of the Australian cohort) stained with Haematoxylin and eosin and Ki67 (Dako) were scanned at X20 onto NDP Viewer (Hamamatsu, Hertfordshire, UK) at x20 magnification using a Hamamatsu NanoZoomer (Hertfordshire, UK) by the Glasgow Tissue Research Facility (GTRF). In the Glasgow combined cohort GMS was determined by James H Park and phenotypic subtype was determined by Dr Antonia K Roseweir. In the TransSCOT cohort both GMS and Phenotypic subtyping was performed by Dr Antonia K Roseweir. In the Australian TMA cohort, DM-CRC-TMA screen-detected CRC cohort and the synchronous resection cohort GMS and phenotypic subtyping was performed by KP.

2.1.2.1 Haematoxylin and Eosin (H&E) staining

Staining of the combined array was performed by Mr Arfon Powell and staining of the TransSCOT cohort was performed by Dr Jennifer Hay. Staining of the Synchronous resection cohort, DM-CRC-TMA screen-detected cohort, and Australian TMA cohort were performed as outlined below.

2.1.2.1.1 H&E staining protocol

Tumour sections were deparaffinized in Histoclear (Agar Scientific, Essex, UK) (3 x 3 minutes) and then rehydrated through a series of graded alcohols; 2 x 3 minutes 100% ethanol, 1 x 3 minutes 95% ethanol, 1 x 3 minutes 90% ethanol, 1 x 3 minutes 80% ethanol, 1 x 3 minutes 70% ethanol, 1 x 3 minutes 50% ethanol, 1 x 30% ethanol. Sections were then washed in deionised water for 5 minutes and placed in Harris Haematoxylin for 3 minutes. After rinsing in water for 5 minutes slides were dipped in acid alcohol (3% HCl in 70%) ethanol for 3 seconds and rinsed for 2 minutes in water. Sections were then added to Eosin for 30 seconds and passed through a series of graded alcohols to dehydrate; 1 x 2 minutes 30% ethanol, 1 x 2 minutes 50% ethanol, 1 x 2 minutes 70% ethanol, 1 x 2 minutes 80% ethanol, 1 x 2 minutes 90% ethanol, 1 x 2 minutes 95%, 2 x 2 minutes 100% ethanol. Slides were then placed in Histoclear for 3 x 3 minutes. Sections were mounted using dystrene plasticiser and xylene (DPX) and left to dry overnight. Sections were scanned onto NDP Viewer (Hamamatsu, Hertfordshire, UK) at x20 magnification using a Hamamatsu NanoZoomer (Hertfordshire, UK) by Glasgow Tissue Research Facility.

2.1.2.1.2 Klintrup-Makinen Grading

Klintrup-Makinen (KM) grading was performed as previously described (7). Briefly, H&E-stained sections were assessed at the deepest point of tumour invasion for the broad inflammatory infiltrate on a scale of 0 to 3. Tumours were graded 0 if there were no immune cells present and 1 if a patchy band of immune cells was seen. Tumours with a prominent thin band of inflammatory cells were graded 2 and those with a thicker florid cup of immune cells were graded 3. For the purposes of GMS and Phenotypic Subtyping, patients graded 0 or 1 were classed as low (0) and patients graded 2 or 3 were graded high (1) for inflammatory infiltrate.

2.1.2.1.3 Tumour-stroma percentage

Tumour stroma percentage was assessed using H&E-stained tumour sections as previously described (98). The intra-tumoural area was assessed for the level of stromal invasion, excluding necrotic areas. Patients were graded as low for stromal invasion if the intra-tumour areas were made up of <50% stromal cells, and high if the intra-tumour stroma was deemed >50% of total area.

2.1.2.3 Glasgow Microenvironment Scoring

Glasgow microenvironment scores (GMS) were determined as previously described and as outlined and as shown in table 2.1 (98). Patients with a high KM grade were automatically assigned to GMS0. Patients with a low KM grade but high TSP were categorised as GMS2. Patients with a low KM grade and a low TSP were classified as GMS1.

KM Grade (Immune Infiltrate)	TSP (Stromal Invasion)	GMS
High (2-3)	Any	0
Low (0-1)	<50%	1
Low (0-1)	>50%	2

Table 2.1 Glasgow Microenvironment Scores. *Table outlining categorisation of each GMS score based on KM Grade and TSP. Tumours with a high KM grade (2-3) are considered to be GMS0, patients with a high tumour-stroma percentage (>50%) are GMS2 and patients low for both measures are GMS1.*

2.1.2.4 Phenotypic subtype

Phenotypic subtypes were determined as previously outlined by Roseweir et al and in table 2.2 (48). Briefly, patients with a high KM grade were assigned to the Immune subtype. Patients with a low KM grade but high TSP were classified as Stromal. Patients with a low KM grade, low TSP but high Ki67 proliferation index (protocol defined in 2.2.3) were classed as canonical and those with low KM grade, low TSP and low Ki67 proliferation index were classified as latent.

Phenotypic Subtype	Immune Infiltrate (KM Grade)	Stromal Invasion (TSP)	Proliferation rate (%Ki67+)
Immune	High	Any	Any
Canonical	Low	Low	High
Latent	Low	Low	Low
Stromal	Low	High	Any

Table 2.2 Phenotypic Subtyping. *Table outlining categorisation of each phenotypic subtype based on histological assessment of the tumour. Tumours with a high KM grade (2-3) are considered to be Immune, patients with a high tumour-stroma percentage (>50%) are Stromal. Patients low for both measures are divided into Canonical if they have high Ki67 expression (>30%), and Latent if <30% of tumour cells are positive for Ki67.*

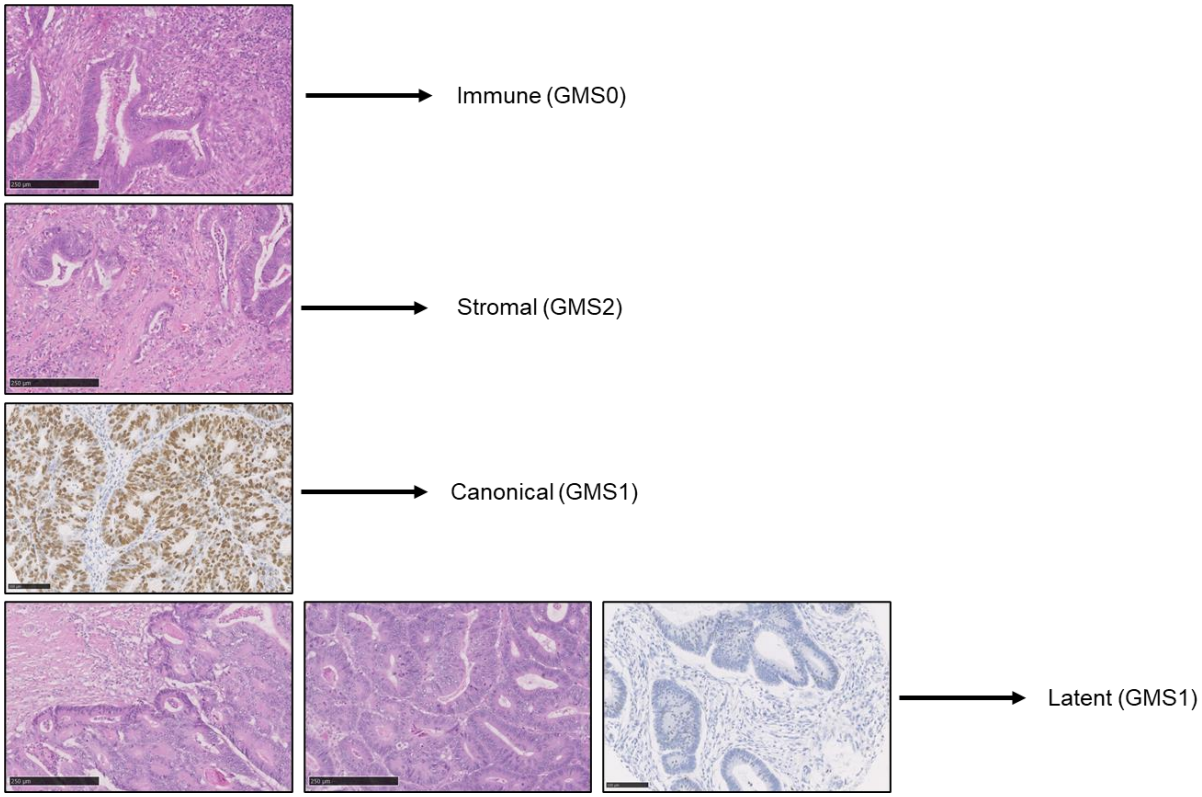


Figure 2.1. Derivation of GMS/Phenotypic histological subtypes. *Tumours are firstly assessed for level of inflammatory infiltrate by Klintrup-Makinen Grade and patients classed as high are automatically assigned to GMS0/Immune subtype. Tumours are then graded for intra-tumour stromal invasion and patients classed as high are considered GMS2 or stromal phenotypic subtype. Patients with a low KM grade and low TSP are automatically GMS1. These patients are split into 2 groups for phenotypic subtype by Ki67 proliferation grade, with tumours with >30% proliferating cells classed as canonical and <30% classed as latent.*

2.1.3 Statistical analysis of tumour phenotype tissue studies

All statistical analyses were completed using IBM SPSS version 22 (IBM, New York, USA). Statistical significance was set to $p<0.05$. Kaplan Meier log rank curves were used to identify associations between protein expression and cancer-specific survival (CSS) within the full cohorts, and then when disease was segregated by GMS, MMR status and tumour subsite (Right-sided, Left-sided, and rectal cases). Cox regression was performed to determine hazard ratios.

2.2 Tissue studies- Investigating IL6/JAK/STAT3 pathway

To assess protein expression of members of JAK-STAT3 signalling in stage I-IV CRC, immunohistochemistry (IHC) was conducted utilising the Glasgow combined colorectal tissue microarray (TMA) of surgically resected tumour specimens. The Glasgow combined array was

previously constructed manually by Dr Jonathan Platt and Dr Arfon Powell, before being reconstructed by KP and Glasgow Tissue Research Facility in 2020 using a TMA Grandmaster (3D Histotech Ltd, Budapest, Hungary).

2.2.1 Antibody Validation

To ensure antibodies utilised were specific for the target protein, specificity testing was performed as outlined in Table 2.3.

2.2.1.1 Western blotting for antibody specificity

Western blotting was used to validate the specificity of IL6R, JAK1, JAK3, TYK2, STAT3, pSTAT3^{tyr705} and pSTAT3^{ser727} antibodies. To make cell lysates, colorectal cell lines SW620 or HT29 cells were seeded at 1x10^5 in 6 well plates and left to attached overnight. Cells were then treated with either 001% DMSO, 1mM Ruxolitinib (JAK1/2 inhibitor, Abcam) or 1mM Tofacitinib (JAK2/3 inhibitor, Abcam). Cells were washed in cold PBS twice on ice before 250µL Laemmli's sample buffer was added to each well. Cells were scraped, sheared using a 21-gauage needle and transferred to Eppendorf tubes. Cells were centrifuged at 15000RPM at 4^{o}C for 15 minutes. Supernatant was carefully removed, transferred to new Eppendorf tubes and heated for 5 minutes at 95^{o}C on a heat block. Cell lysates were used immediately.

Running buffer was prepared by diluting 100mL 10X running buffer, tris/glycine/SDS in 900mL distilled water. Precast protein gels (7.5% mini-PROTEAN ® TGX ™) were utilised (Bio-Rad #4561023). The gel tank was assembled, added to the Bio-Rad tank, and tested for leakage when the reservoir was filled with running buffer. Ladder (10µL) was loaded into lane 1 and 25µL samples were added to remaining wells in triplicate. Gels were run for 90 minutes or until the samples reached the edge of the gel at a constant 120 Volts. Transfer buffer was prepared by adding 200mL methanol and 100mL 10X transfer buffer (tris/glycine) to 700mL distilled water. Polyvinylidene fluoride (PVDF) membranes were soaked in methanol for 5 minutes. Sponges and filter paper were soaked in transfer buffer. Sandwich was prepared as outlined in figure 2.1 and rolled using to remove any air bubbles. The assembled cassette was added to the tank and run at constant 300mA for 90 minutes to transfer protein from gels to the membranes.

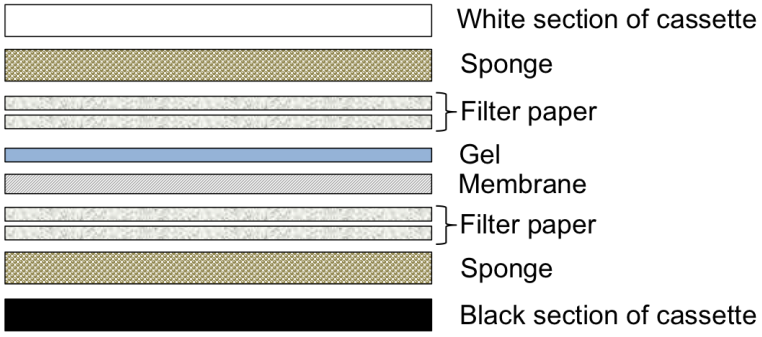


Figure 2.2 Western blot transfer. *Diagram showing the makeup of gel membrane sandwich utilised in western blot transfer step. The cassette is opened, and a piece of sponge is placed at both ends. Two pieces of filter paper are added to each side and the membrane is placed on the side nearest the black end of the cassette. The gel is carefully added on top of the membrane and the sandwich is assembled.*

The sandwich was removed from the tank and membranes were blocked in 3% bovine serum albumin (BSA) for 1 hour at room temperature. Membranes were incubated in appropriate primary antibody (diluted in 0.3% BSA) overnight at 4^{o}C and then washed in tris-buffered saline (TBST) 3 x 10 minutes. Secondary antibodies were diluted in 0.3% BSA with anti-ladder at 1:50000 (1:5000 for anti-rabbit and 1:5000 for anti-mouse antibodies). Membranes were incubated for 90 minutes in secondary antibodies at room temperature and then washed 3 x 10 mins in TBST. After washing, membranes were incubated in horse radish peroxidase (HRP) substrate enhanced chemiluminescence reagent (Pierce ECL) (Thermo Fisher Scientific, Waltham, MA, USA) for 5 minutes at room temperature. Membranes were blotted on blue roll and then imaged using Syngene Gene Sys (Syngene International Ltd, India).

2.2.1.2 Cell pellets for antibody specificity

Cell pellets were used to validate the specificity of JAK2, JAK3, TYK2, pSTAT3^{tyr705} and pSTAT3^{ser727} antibodies. Cells (SW620 or COLO205) were passaged into T75 flasks and incubated overnight. Flasks were treated with 0.01% DMSO, 10µM Ruxolitinib (Abcam, 141346) or10µM Tofacitinib (Abcam, 142068) for 72 hours. For each treatment, 1 x T75 flask was required to make 1 cell pellet. Cells were trypsinised and added to a 15mL falcon tube with medium up to 10mL. Tubes were centrifuged for 3 minutes at 1200RPM. Supernatant was removed using an aspirator and the pellet resuspended in PBS. Cells were

spun for 3 minutes at 1200RPM. Supernatant was removed and cells were resuspended in 1ml 4% paraformaldehyde (PFA) and transferred to Eppendorf tubes. After 20 minutes fixation cells were spun at 1200RPM for 3 minutes and washed in PBS twice. Once dry, the pellet was resuspended in 1% agarose and left overnight at 4^{o}C. The embedded pellet was removed and transferred to a labelled cassette. The pellet was dehydrated through a series of graded alcohols; 1 x 15 minutes 50% ethanol, 1 x 15 minutes 75% alcohol, 2 x 15 minutes absolute ethanol and then 2 x 10 minutes Histoclear. The pellet in the cassette was transferred to melted paraffin wax in a TissueTech embedding centre for 1 hour. Finally, the pellet was removed from the cassette, placed in a mould, embedded in paraffin wax and left on a cold block for 1 hour. After setting overnight embedded pellets were cut at 4µM thick on a microtome and attached to slides by baking overnight at 50^{o}C. Pellets were then stained using the same IHC method outlined in 2.1.4 for respective proteins, aside from heating at pressure for 3 minutes as opposed to 5 minutes for TMAs and full sections.

2.2.1.3 Antibody validation for each protein

For JAK1 (Cell signalling #3344), western blotting was performed on HT29 cell lysates provided by Dr Jean Quinn and PC-3 lysates provided by Dr Milly McAllister. Probing for JAK1 in HT29 cells detected a band at the correct molecular weight of 130Kda in triplicate, with no bands detected in PC3 lysates and tubulin consistent at 52KDa in all samples (Table 2.3). JAK2 (Cell signalling #3773) specificity was tested by staining COLO205 cell pellets treated with either 0.01% DMSO or 1mM Ruxolitinib JAK2 inhibitor (Table 2.3). JAK3 (Abcam, ab45141) antibody specificity was confirmed via western blotting of HT29 cell lysates and PC3 lysates. A single band in triplicate was detected at 115KDa was detected in HT29 cells, with no bands in the PC3 lysates but consistent bands at 52KDa when probed for tubulin (Table 2.3). TYK2 (Abcam, ab223733) specificity was confirmed by probing HT29 and PC3 cell lysates and identifying a single band in triplicate in HT29 cells at 134KDa and no bands in PC3 cells but consistent tubulin detected at 52KDa (Table 2.3). Further validation was performed by staining SW620 cell pellets treated with either 0.01% DMSO or AT2983 JAK3 inhibitor Selleckchem, #S1134). Expression of TYK2 was higher in vehicle control treated cell pellets compared to AT9283-treated cells (Table 2.3). Specificity of STAT3 (Cell Signalling, #9132) was assessed via western blotting of HT29, HeLa and PC3 cell lysates. A single band at 88KDa was detected in HT29 and HeLa lysates with no bands detected in PC3 cells. SW620 cell lysates treated/untreated with JAK inhibitor Ruxolitinib were probed for pSTAT3tyr705 via western blot. Single band in triplicate was observed at 88KDa in the untreated samples and no bands were detected in the Ruxolitinib-treated cells. COLO205 cell pellets treated with 0.01% DMSO showed higher protein expression of $pSTAT3^{tyr705}$ versus 1mM Ruxolitinib COLO205 treated pellets when stained via IHC. Finally, $pSTAT3^{ser727}$ antibody (Cell Signalling #9134) specificity was tested in SW620 cell lysates treated with 0.01% DMSO or 1mM Ruxolitinib via western blot. Single bands in triplicate were detected at 86KDa in vehicle control samples and no bands were observed in JAK1/2 inhibited cells. Ruxolitinib-treated SW620 cell pellets showed lower expression of $pSTAT3^{ser727}$ compared to vehicle control treated cells when stained via IHC.

Protein	Antibody	Validation method
IL6R	Abcam #ab128008	• Single band on Western blot in HeLa lysates at 51KDa with no bands detected at the same molecular weight in LNCaP cell lysates
JAK1	Cell signalling #3344	• Single band on Western blot at 130KDa in HT29 cell lysates and no band detected in true positive PC3- lysates
JAK2	Cell signalling #3773	• Higher JAK2 protein expression observed via IHC staining in COLO205 cell pellets from cells treated with 0.01% DMSO versus COLO205 cells treated with JAK2 inhibitor
JAK3	Abcam #ab45141	• Single band on Western blot at 115KDa in HT29 cells with no bands detected in PC3 lysates. IHC staining of SW620 cell pellets treated with JAK3 inhibitor or 0.01% DMSO.
TYK2	Abcam #ab223733	• Single band on Western blot of HT29 cell lysates at 134KDa with no bands detected in PC3 lysates. • IHC staining of sw620 cell pellets treated with TYK2 inhibitor or 0.01% DMSO.
STAT3	Cell signalling #9132	• Single band at 88KDa detected in HT29 cell lysates and HeLa cell lysates and no bands in PC3 lysates.
pSTAT3^{705}	Cell signalling #9131	• Single band at 88KDa in Sw620 cell lysates treated with vehicle control with no bands in Sw620s treated with Ruxlitinib and consistent Tubulin at 52KDa in all samples. • COLO205 cell pellets treated with 0.01% DMSO or Ruxolitlinib stained via IHC for pSTAT3^{tyr705}
pSTAT3^{727}	Cell signalling #9134	• Single band on western blot detected at 86KDa in Sw620 cells treated with 0.01% DMSO and n bands at the same molecular weight in Ruxolitinib treated SW620 lysates. • IHC staining of pSTAT3^{ser727} in Sw620 cell pellets treated with 1mM Ruxolitinib or Tofacitinib.

Table 2.3: Antibody validation. *The antibody with catalogue number and method of validation outlined for each protein of interest. For each antibody either western blotting to detect a single band in a positive control or staining of cell pellets known to be positive/negative for protein of interest or both was performed.*

2.2.2 Patient TMA

Prior to this study FFPE CRC blocks were retrieved from NHS GGC pathology archives and used to construct TMA blocks. Full sections from each block were stained using Haematoxylin and Eosin and marked by a pathologist for tumour-rich areas. Four 0.6mm cores were taken from the tumour-rich

areas of each full section and embedded in four separate recipient paraffin blocks. The use of a core from four discrete areas aims to account for tumour heterogeneity.

2.2.3 Immunohistochemistry

The Glasgow combined cohort TMA was stained via IHC for IL6R, JAK1, JAK2, STAT3 and pSTAT3^{727} by KP. Staining of pSTAT3^{705} was previously completed by Mr James H Park. Antibody optimisation was performed on surplus colorectal tissue slides prior to staining the patient cohort tissue.

Ki67 staining of the combined array was performed by Mr Arfon Powel and staining of the TransSCOT cohort was performed by the pathology laboratory at the Queen Elizabeth University Hospital, Glasgow. The synchronous resection cohort, DM-CRC-TMA cohort and Australian TMA cohorts were stained as outlined below.

2.2.3.1 Slide preparation

TMA sections were requested from NHS GGC Biorepository and provided by Glasgow Tissue Research Facility. Slides were stored at 4^{o}C before use. Immediately before staining TMA sections were baked at 55^{o}C for 10 minutes to minimise the loss of cores.

2.2.3.2 Dewaxing and rehydration

Sections were dewaxed in Histoclear (2 x 3 minutes) and rehydrated through a series of graded alcohols; (2 x 3 minutes in absolute ethanol, 1 x 2 minutes in 90% ethanol, 1 x 70% ethanol for 2 minutes). Slides were rinsed in running water for 10 minutes.

2.2.3.3 Antigen retrieval

To unmask epitopes blocked during formalin fixation antigen retrieval was performed using Citrate buffer pH6 or TRIS-EDTA buffer pH8 or pH9. Appropriate buffer and pH were determined during antibody optimisation. Citrate buffer was made using 2mM tris-sodium citrate dehydrate and 8mM citric acid in 1L distilled water. TRIS-EDTA buffer was prepared using 5mM Tris Base and 1mM EDTA in 1L distilled water. Prior to antigen retrieval 1L of buffer was heated in an open pressure cooker in a microwave for 13 minutes. Slides were added to the buffer and the pressure cookers seal, lid and topper were secured. The slides were heated for ~3 minutes until pressure reached and then heated for a further 5 minutes. After heating sections were left to cool for 30 minutes in buffer and then rinsed in running water.

2.2.3.4 Blocking endogenous peroxidase activity

Slides were placed in 3% H_2O_2 for 10 minutes (pSTAT3^{tyr705}) or 20 minutes for all other proteins, to quench endogenous peroxidases, thus reducing background staining. Slides were rinsed in running water.

2.2.3.5 Blocking non-specific binding

To prevent off-target non-specific binding sections were incubated in a blocking solution. During optimisation optimum blocking solution was identified for each antibody (Table 2.4). Tumour tissue on each slide was circled using a Dako pen (S2002, Dako, Agilent technologies, Stockport, UK) to ensure the blocking solution covered the entire tumour section and was unable to run off the slide. Blocking solution (5%/10% casein or 5% horse serum) was added and slides were incubated for 1 hour or 30 minutes (only pSTAT3^{tyr705}) at room temperature. Blocking solutions were diluted in antibody diluent (S0809, Dako, Agilent Technologies, Stockport, UK).

Protein	Buffer	Blocking solution	Antibody dilution	Secondary
IL6R	Citrate pH6	5% casein	1:3000	ImmPRESS™
JAK1	Citrate pH6	5% casein	1:100	ImmPRESS™
JAK2	Citrate pH6	5% casein	1:100	ImmPRESS™
JAK3	TRIS-EDTA pH9	5% goat serum	1:200	ImmPRESS™
TYK2	TRIS-EDTA pH9	5% goat serum	1:200	ImmPRESS™
STAT3	Citrate pH6	10% casein	1:300	ImmPRESS™
pSTAT3^{705}	TRIS-EDTA pH9	5% horse serum	1:50	Envision™
pSTAT3^{727}	TRIS-EDTA pH9	10% casein	1:400	ImmPRESS™
Ki67 (S0809 MIB-1 clone, Dako, Agilent, #GA626)	TRIS-EDTA pH8	10% casein	1:50	Envision ™

Table 2.4: IHC conditions for each protein of interest. *Description of antigen retrieval buffer, blocking solution and antibody dilation used for each protein of interest. The buffers utilised were either citrate pH6, TRIS-EDTA pH8 or pH9 all determined during antibody optimisation. The blocking solutions used were 5/10% casein, 5% goat serum or 5% horse serum. The antibody dilution in antibody diluent ranged from 1:3000 to 1:50 dependant on optimisation experiments. ImmPRESS™ or Envision™ were utilised as secondary antibodies.*

2.2.3.6 Primary antibody incubation

Antibodies were diluted to respective concentrations (Table 2.4) in antibody diluent (S0809, Dako, Agilent Technologies, UK). Blocking solution was tapped off, diluted antibodies were added, and sections were incubated at 4^{o}C overnight.

2.2.3.7 Secondary antibody incubation

After incubation sections were washed twice in TBS for 5 minutes and then incubated in ImmPRESS™ reagent (Vector Laboratories Inc, California, USA) or EnVision™ (only for pSTAT3^{tyr705}) (K5007, Dako, Agilent Technologies, Stockport, UK) for 30 minutes at room temperature. Slides were washed twice in TBS for 5 minutes.

2.2.3.8 Visualisation with DAB substrate

Sections were incubated in DAB chromogen substrate (Vector Laboratories Inc, California, USA) for 5 minutes at room temperature and subsequently rinsed in running water for 10 minutes.

2.2.3.9 Counterstaining, dehydration and mounting

Slides were counterstained in Harris Haematoxylin for 5 minutes, dipped in 1% acid alcohol (HCl in ethanol) for 3 seconds and blued in Scott's Tap Water Substitute (80mM Magnesium sulphate, 40mM sodium hydrocarbonate in distilled water) for 45 seconds. Sections were dehydrated through a series of graded alcohols; 1 minute in 70% ethanol, 1 minute in 90% ethanol, 2 x 1 minute in absolute ethanol and then 2 x 1 minutes in Histoclear. Coverslips were mounted onto slides using histological mounting medium Omnimount (HS-110, SLS, Nottingham, UK).

2.2.3.10 Slide scanning and visualisation

After staining slides were scanned onto NDP Viewer (Hamamatsu, Hertfordshire, UK) at x20 magnification using a Hamamatsu NanoZoomer (Hertfordshire, UK).

2.2.4 Scoring of IHC staining

Scoring was performed by a single observer blinded to clinical outcome data, with 10% scored by a coinvestigator. For cytoplasmic and nuclear stains, weighted histoscoring was performed and for membrane staining histoscoring was performed. Weighted histoscore assessment was performed at X20 magnification using: 0 x (% of cells not stained/negative) +1 x (% cells weakly stained) +2 x (% cells moderately stained) + 3 x (% cells strongly stained), giving a range of scores from 0-300.

Membrane histoscores were determined by allocating strong staining a score of 3, moderate staining 2, weak staining 1 and no staining 0. For membrane staining, modified histoscoring was performed with tumours graded as absent or present for expression. Validation of manual scoring was performed using QuPath digital pathology software v0.2.3 (QuPath, Edinburgh, UK) for 10% of cores by Sara Al Badran for IL6R, JAK1, JAK2, JAK3, STAT3 and by KP for TYK2, and both pSTAT3 proteins(99). To validate reliability of scores, interclass correlation coefficient (ICCC) was calculated using SPSS software between both observers scores. If an ICCC>0.7 and a correlation coefficient of >0.7 were observed scoring was deemed valid. Scatter plots and Bland Altman plots were constructed to visualise the correlation between scores.

IL6R, JAK3, TYK2, pSTAT3^{ser727} were scored within the tumour cytoplasm. JAK1 and JAK2 were assessed within the tumour cytoplasm and membrane. STAT3 was scored within the tumour cytoplasm and nuclei. Expression of pSTAT3^{tyr705} was assessed only in the tumour nuclei.

For the combined array and TransSCOT cohorts, scoring of Ki67 was performed by Dr Antonia Roseweir using an automated nuclear algorithm in Slidepath digital image hub (Leica Biosystems, Wetzlar, Germany. Scoring of the synchronous cohort, DM-CRC-TMA cohort and Australian TMA cohorts was performed by a single observer blinded (KP). Hot spot areas with the greatest intensity of nuclear staining were selected and the percentage of positive cells within three predetermined fields of view at X400 were assessed. Receiver operator curves (ROC) were previously utilised to determine the optimal cut point for high and low Ki67 expression at 30%. Patients with a proliferation index of <30% were classified as low proliferation and those with a score off >30% were highly proliferative (49).

2.2.5 Statistical Analysis of IHC tissue-based studies

To set threshold values for high and low expression of each protein log rank statistics were performed in R Studio (RStudio, Boston, MA, USA) using survminer, survival, tidyverse and maxstat packages. All further statistical analyses were completed using IBM SPSS version 22 (IBM, New York, USA).

Kaplan Meier log rank curves were used to identify associations between protein expression and cancer-specific survival (CSS) within the full Glasgow combined cohort, and then when disease was segregated by GMS, MMR status and tumour subsite (Right-sided, Left-sided and rectal cases). Cox regression was performed to determine hazard ratios. Chi-squared testing was performed to assess association between expression of each protein and clinicopathological characteristics. Chi-squared testing was also utilised to determine association between main pathway activation marker pSTAT3^{tyr705} and the other pathway members. Statistical significance was set to $p<0.05$.

2.3 RNAscope®

To detect the soluble cytokine that activates JAK-STAT3 signalling, interleukin-6 (IL6), novel RNA in situ hybridisation RNAscope (ACD Bio, California, USA) was performed by Colin Nixon at the CRUK Beatson Institute on the Glasgow combined cohort TMAs. This technique enabled quantitative detection of IL6, which were unable to be detected by IHC to a sufficient quality. At the same time of IL6 probing, staining of housekeeping gene, PPIB, was also performed in the Glasgow combined cohort. IL6 was quantified using Halo digital pathology software (Indica Labs, Albuquerque, NM, USA) in copies per μM^2. A classifier to distinguish between tumour and stromal rich areas was built to measure IL6 in different TME areas. Thresholds were set to measure the intensity of brown staining indicative of the level of IL6 mRNA copies. Raw scores for IL6 expression within the tumour and stroma were normalised to PPIB scores. Data were entered into IBM SPSS version 22 software (IBM, New York, USA), and statistical analyses were performed in the same manner as for the IHC protein pathway data. Cut offs for high and low expression were determined using survminer, survival, maxstat and tidyverse packages in R studio based on cancer-specific survival (RStudio, Boston, MA, USA).

2.4 Genomics

Mutational profiling was performed on a subset of patients from the Glasgow cohort of patients from the Glasgow combined cohort (n=252). DNA was previously extracted from FFPE sections by NHS molecular diagnostics, Dundee and stored at -80^{o}C. DNA quality and concentration were determined using the Qubit assay (ThermoFisher, Massachusetts, USA). Samples with a DNA concentration of >10ng/µL were included in the study. Sequencing was outsourced and performed by Dr Susie Cooke and the Glasgow Precision Oncology Laboratory (GPOL) using a custom in-house designed panel of 151 cancer-associated genes. Statistical analyses of broad mutational patterns relative to GMS and $pSTAT3^{tyr705}$ high/low groups were performed using maftools packages of R Studio. Summary plots were used to visualise the top variant types and classifications, mutation burden and SNV classes. Oncoplots were used to visualise the top 10 mutations within each patient group. Pathway enrichment analysis was performed to identify patterns of mutation in genes assigned to distinct cancer-associated signalling pathways. Fishers' exact tests were performed to identify any significantly different mutations across groups and results were visualised in forest plots. Copy number analysis was performed in collaboration with Dr Akhill Yerdusi using Python. Swarm and point plots were used to visualise the most differentially mutated genes between each GMS group and $pSTAT3^{tyr705}$ high/low groups.

2.5 Transcriptomics

Whole genome RNAseq was performed on a subset of patients from the Glasgow combined cohort (n=100). To utilise available FFPE full section tissue, novel RNAseq technique, TempO-Seq was employed (Biospyder Technologies, Carlsbad, CA, USA). TempO-Seq gene expression profiling was performed according to manufacturer's directions. Briefly, whole sections (~100 mm^2 x 0.5 µm) were excised from FFPE slides of colorectal cancer resection samples and placed into wells of a PCR plate. TempO-Seq Lysis Buffer was added, and the sample was overlaid with mineral oil. After deparaffinization by heating, the tissue was lysed using TempO-Seq Protease mix. The lysate was then combined with a mixture of detector oligonucleotides (DOs), designed as pairs that anneal adjacent to one another on the target RNAs. After a hybridization step, unbound DOs were degraded in an enzymatic step, and the bound DOs were ligated into a complete probe sequence. The ligated probes were amplified in a PCR step, purified, and combined in an indexed multiplex library which was sequencing using an Illumina instrument (Illumina, CA, United States) to count the relative amount of each target DO pair representing each gene's expression level.

Statistical analyses were performed in R Studio (RStudio, Boston, MA, USA). Raw counts were normalised using DESeq2 and analysed for both GMS and $pSTAT3^{tyr705}$ high/low groups. Analysis was performed using the full 22,000 gene probes by Bioclavis (Bioclavis, Glasgow, UK). DESeq2 was used to construct tables of differential gene expression. Box plots were constructed by KP in SPSS (IBM, NY, USA) for differentially expressed genes relative to GMS groups. Over-representation analysis was performed in R Studio using go_enrich to analyse interactions between differentially expressed genes by KP. PCA plots were constructed to determine any clustering of gene expression between histological groups. Volcano plots were plotted by Dr Gerard Lynch to visualise and differentially expressed genes between $pSTAT3^{tyr705}$ groups. For the purposes of this study significance was set to padj <0.05.

2.6 *In vitro* cell line studies

2.6.1 Colon cancer cell line culture

Seven colon cancer cell lines were cultured in ATCC recommended growth medium and FBS, as outlined in table 2.5. Using seven different lines allowed for a broad spectrum of mutational backgrounds and MSI statuses. Cells were grown in T75 flasks at 37ºC 5% CO_2 and medium was changed every 2-3 days. Upon reaching 70% confluence cells were passaged using 0.05% Trypsin (Thermo Fisher Scientific, Waltham, MA, USA) and split at a 1:3 ratio. Mycoplasma testing was performed once every 3 months. Brightfield images of COLO205, T84, HT29, SW620, HCT116 and SW480 were taken at X40 magnification pre-treatment and are shown in figure 2.2 to visualise

distinct morphological differences between each cell line. Images were not available for DLD-1 or SW48 lines.

Cell Line	CMS	MSI	Mutations	Medium
COLO205	Immune	MSS	BRAF	Dulbeccos Modified Eagle Medium (DMEM) Glutamax (61965240, Life Technologies, Paisley, UK), 10% FBS
HT29 (ATCC HTB-38™)	Metabolic	MSS	BRAF, PI3K, TP53	McCoy's 5A (Modified Medium, GlutaMAX™ Supplement (36600021, Life Technologies, Paisley, UK), 10% FBS
HCT116	Mesenchymal	MSI-h	KRAS, PI3K	RPMI 1640 Media (12633012, Life Technologies, Paisley, UK), 10% FBS
DLD-1	Immune	MSI-h	KRAS PI3K, TP53	RPMI 1640 Media (12633012, Life Technologies, Paisley, UK), 10% FBS
T84 (ATCC, CCL-248™)	Canonical	MSS	KRAS	Dulbecco's Modified Eagle Medium (DMEM)/F12 (Ham) (11320033, Life Technologies, Paisley, UK), 5% FBS
SW620	Mesenchymal	MSS	KRAS, TP53	RPMI 1640 Media (12633012, Life Technologies, Paisley, UK), 10% FBS
SW480	Mesenchymal	MSS	KRAS, TP53	Dulbeccos Modified Eagle Medium (DMEM) Glutamax (61965240, Life Technologies, Paisley, UK), 10% FBS
SW48	Immune	MSI-h	WT	Dulbeccos Modified Eagle Medium (DMEM) Glutamax (61965240, Life Technologies, Paisley, UK), 10% FBS

Table 2.5: Colorectal cancer cell lines. *Table outlining colorectal cell lines used in drug screening assays including COLO205, HT29, HCT116, DLD-1, T84, SW620, SW480 and SW48. For each cell line CMS, MSI status and mutation profile was available in the literature (11). Culture medium was utilised in accordance with ATCC guidelines as shown in the table.*

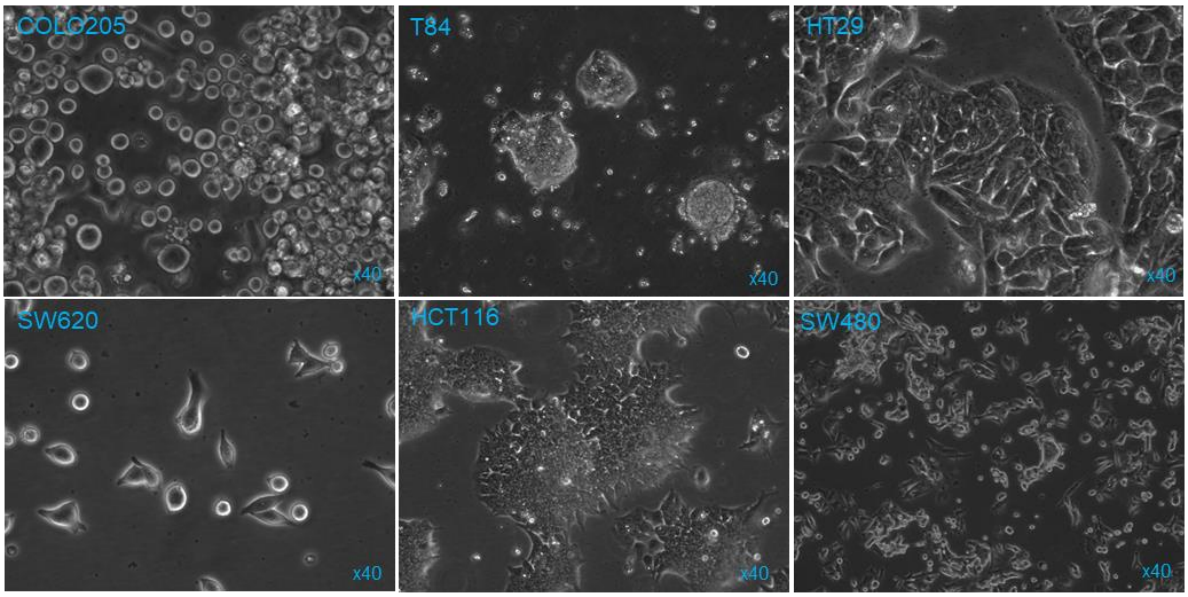


Figure 2.3 Brightfield images of CRC cell lines utilised. *Light microscopy images taken at X40 magnification of COLO205, T84, HT29, SW620, HCT116 and SW480 colorectal cell lines showing distinct morphologies. DLD-1 and SW48 cell lines were also utilised within the project however images were unfortunately not available.*

2.6.2 Inhibition of JAK-STAT3 signalling in colon cancer cell lines

2.6.2.1 ELISA to detect cellular pSTAT3^{tyr705}

To confirm that Ruxolitinib and Tofacitinib were targeting IL6/JAK/STAT3 signalling a solid phase sandwich enzyme linked immunosorbent assay (ELISA) was performed to detect levels of pSTAT3^{tyr705} in treated and untreated 2D cell line samples (RND Systems DYC4607B-2). SW620 cell lines were seeded at 1x10^6 per well of a 6 well plate. Cells were treated 24 hours after plating with either 0.01% DMSO, 1mM Ruxolitinib or 1mM Tofacitinib in triplicate and incubated at 37^{o}C for 72 hours. Plates were rinsed with PBS. Cells were lysed in lysis buffer (1 mM EDTA, 0.5% Triton X-100, 5 mM NaF, 6 M Urea, 25µg/mL Leupeptin, 25µg/mL Pepstatin, 100µM PMSF, 3.0µg/mL Aprotinin, 2.5 mM Sodium Pyrophosphate, 1 mM activated Sodium Orthovanadate in PBS, pH 7.2-7.4) and left on ice for 15 minutes before being frozen at -80^{o}C. Prior to performing the assay, tubes were centrifuged at 2000 x g for 5 minutes and supernatant was transferred to fresh tubes. Samples were diluted 6-fold in diluent (1 mM EDTA, 0.5% Triton X-100, 5 mM NaF in PBS). The ELISA plate was incubated overnight at room temperature with carrier protein provided in the kit. Wells were washed with 0.05% Tween in PBS before being blocked with 1% BSA, 0.05% NaN3, in PBS for 2

hours at room temperature. Plates were washed with 0.05% Tween in PBS and then samples, and standards were added to the plate in triplicate and duplicate, respectively. Plates were incubated for 2 hours at room temperature, washed, and incubated for 2 hours at room temperature with 100uL detection antibody. Plates were washed and 100uL Streptavidin-HRP was added for 20 minutes in the dark. Plates were washed again and 100uL Substrate solution added to each well for 20-minute sin the dark. Finally, 50uL stop solution was added to each well and optical density was read at 450nM using a TECAN Infinite 200 PRO plate reader (Tecan Group Ltd, Switzerland). Optical density read outs for standards were utilised to determine concentration of $pSTAT3^{tyr705}$ in the SW620 samples by plotting a standard curve in Microsoft Excel (Microsoft, Albuquerque, NM, USA). Paired T tests were used to determine any statistical difference between DMSO vehicle control and JAK inhibitor samples with α set to 0.05.

2.6.2.2 Expression of $pSTAT3^{tyr705}$ in patient derived explants

Patient-derived explants were utilised to perform further validity testing of Ruxolitinib and Tofacitinib targets. The explant model was developed by Dr Lesley Stark (University of Edinburgh) and adapted for applying to surplus available tissue. Normal and tumour resections from colorectal cancer patients undergoing surgery with curative intent were obtained from NHS GGC tissue biobank. Samples were stored at 4ºC in PBS until processing. Small biopsy sized chunks of tissue were removed from larger resection specimens. For each patient the aim was to obtain 3 "biopsies/explants" each from the tumour and normal resection. Samples were washed in phosphate buffered saline-antibiotic (PBS-ABS) twice. Each explant was placed in separate wells of a 6-well plate. One well of each plate was treated as control wells and 0.5mL Dulbecco's Modified Eagles Medium with Nutrient Mixture F12 Hams medium (DMMEM) (Invitrogen #12634) (with no additives) was added to each well. The remaining explants were immersed in either 0.5mL 10µM Tofacitinib or 10µM Ruxolitinib. Explants were incubated overnight at 37ºC 5% CO_2 and then fixed in 4% paraformaldehyde in universal tubes overnight. After fixation explants were washed in PBS and immersed in ethanol overnight. Paraffin embedding was performed by Colin Nixon and the CRUK Beatson histology service. Sections for IHC staining (4µM thick) were cut using a microtome and baked overnight on slides at 50ºC. Explant slides were stained for $pSTAT3^{tyr705}$ using the same protocol as for TMA sections to compare protein expression between treated and untreated samples. IHC was performed exactly as outlined in section 2.2. Sections were scanned onto NDP Viewer (Hamamatsu, Hertfordshire, UK) at x20 magnification using a Hamamatsu NanoZoomer (Hertfordshire, UK) by Glasgow Tissue Research Facility.

2.6.2.3 Jakinib drug screening

To perform inhibition of JAK-STAT3 signalling, cell lines were grown to 70% confluence and then seeded into 96 well plates. Due to the difference in growth rates optimum seeding densities (Table 2.6) were determined prior to any drug screening experiments. The volume of cells/medium per well was 200µL. After plating, cells were left to attach overnight at 37ºC 5% CO_2. Medium was removed using an aspirator and replaced by either fresh medium, 0.01% DMSO in medium, 1mM, 100µM, 10µM, 1µM, 0.1µM Ruxolitinib (Abcam, Jak1/2 inhibitor) or Tofacitinib (Abcam, Jak2/3 inhibitor) in triplicate. Cells were incubated for 48 hours, 72 hours or 96 hours. This was repeated for each time point 3 times to give n=3.

Cell Line	Seeding density (96 well plate)
COLO205	15000/well
HT29	5000/well
HCT116	5000/well
DLD-1	10000/well
T84	10000/well
SW620	10000/well
SW480	10000/well
SW48	25000/well

Table 2.6 Cell line seeding densities. *Table outlines number of cells for each cell line seeded in each well of a 96 well plate in preparation for drug screening. Density ranged from 25,000 cells per well to 5000 per well depending on growth rates determined during assay optimisation.*

2.6.3 Cell viability of colon cancer cell lines following JAK-STAT inhibition

At the appropriate time points, 20µL WST-1 reagent was added to each well in the dark. Plates were incubated for 1-2 hours at 37ºC 5% CO_2 and the absorbance was read at 450nM using a TECAN infinite 200 PRO plate reader (Tecan Group Ltd, Switzerland). Averages of the 3 wells were calculated and treated wells were normalised to control read outs. Normalised values for each plate were averaged for the 3 experiments. Microsoft excel (Microsoft, Albuquerque, NM, USA) was used to perform paired T-tests (2-tailed) to compare cell viability between experimental conditions and

vehicle control wells. Bar charts were plotted to show the data graphically and significance was set to $p<0.05$.

2.7 Colorectal cancer models*; in vitro* 3D patient-derived organoids

2.7.1 Patient-derived organoid establishment

For the purposes of this thesis, patient-derived organoids (PDOs) were either established and grown from fresh surplus surgical resection tissue (Glasgow patients) or obtained through Professor Owen Sansom's laboratory at the CRUK Beatson institute from Sanger centre PDO lines.

PDOs from tumour resections were grown using a protocol developed at Memorial Sloan Kettering Cancer Center by Mr Josh Smith and Dr Chao Wu. Additional help was initially received from Dr Kevin Myant at the University of Edinburgh.

Tumour resections were obtained from NHSGGC tissue biobank covered by ethics for surplus tissue, with consent from patients undergoing potentially curative resection of primary CRC. Resections for organoid culture were immediately place in cold PBS and kept on ice/at 4^{o}C until processing.

Tumour specimens were washed in cold PBS-BS twice and chopped in PBS-DTT using two sterile scalpel blades. Chopped tissue was transferred to a 15mL falcon tube with 10mL digestion media (100mL ADMEM + 2mL FBS + 2mL penicillin streptomycin + 10mg Collagenase XI + 12.5mG Dispase II). Samples were incubated at 37^{o}C on a shaker for 40 minutes, after which 1 volume TryPLE (Thermo Fisher Scientific, Waltham, MA, USA #12605028) was added and the samples incubated for a further 20 minutes at 37^{o}C. The sample was washed through a 40µM filter to remove any larger chunks and spun for 5 minutes at 1200RPM for 5 minutes. Pellet was washed in PBS, spun, and resuspended in 0.2-0.6mL Cultrex Reduced Growth Factor Basement Membrane Extract (BME) (Rnd Systems 3433-001) depending on pellet size. Resuspended cells were transferred to a pre-warmed 24 well plate (20µL/well), inverted and incubated for 30 minutes at 37^{o}C 5% CO_2. After the BME had solidified 0.5mL ADMEM with additives (Table 2.7) was added to each well. Medium was changed every 2-3 days and organoids passaged upon reaching 70% confluence.

Component	Supplier	Final volume/concentration
Advanced Dulbecco's Modified Eagles Medium with Nutrient Mixture F12 Hams (ADMEM)	Invitrogen #12634	11.5mL/plate
FBS	Thermo #26140	5mL/500mL ADMEM
Glutamax	Invitrogen 35050-079	10mL/500mL ADMEM
B27	Gibco 12587-010	1X
N2	R&D Systems AR009	1X
Normacure	Invivogen	1mL/500mL ADMEM
Wnt3a	R&D Systems 5036- WN	10 µM
Gastrin	Sigma-Aldrich G9145	10nM
N-Acetyl-L-cysteine	Sigma-Aldrich A9165-5G	1.25mM
Nicotinamide	Sigma-Aldrich N0636	10mM
H recombinant Rspondin	Peprotech 120-38	1µg/mL
P38 MAPK inhibitor (SB202190)	Sigma S7067	10µM
Noggin	Peprotech 120-10C	100ng/mL
EGF	Gibco PMG8043	50ng/mL
TGFβ Receptor inhibitor (A83-01)	Tocris 2939	500nM
Prostaglandin E2	Sigma P0409	10nM
Rhok inhibitor (Y027632)	Merk Millipore SCM075	10µM

Table 2.7: Medium for derivation of patient-derived organoids. *Table describing the constituents of PDO medium including manufacturer and final concentrations. Protocol for making medium was obtained from Mr Josh Smith and Dr Chao Wu at Memorial Sloan Kettering, New York.*

2.7.2 Characterisation of patient-derived organoids using immunofluorescence

To confirm the presence of a mixture of cell types in organoid structures cells were stained for marker of tumour cells (CD326/EpCam1) and fibroblasts (α-SMA) using immunofluorescence. Antibody concentrations and manufacturers are listed in table 2.8. Protocol was adapted with help from Dr Dustin Flannagan (Sansom group, CRUK Glasgow) using Nature Protocols Broutier et al (100).

Organoids were grown in 24 well plates (20µL BME 'dot' per well) to 70-80% confluence. For each protein of interest 100-200 organoids were required (2-3 wells) and during each IF experiment a negative control (no primary antibody) was used. Medium was removed using a P1000 Gilson pipette, taking care not to disrupt the BME. Wells were washed with 1mL of PBSB (1X PBS with 0.1% BSA) and organoids were transferred to 15 mL tubes using a PBSB-coated Pasteur pipette (1 tube per stain). Tubes were topped up with ~10mL cold PBS and left on ice for 10 minutes to allow organoids to settle. PBS was carefully removed, and cells were washed in cold PBS by inverting the tube and settled on ice for 30 minutes. Organoids were fixed in 4% paraformaldehyde on ice for 30 minutes and then washed in PBS on ice for 10 minutes. Blocking was performed using PBSDT (PBS 1X 0.5% TritonX 1% DMSO 1% BSA 1% FBS) for 1-3 hours at room temperature with gentle agitation. Organoids were transferred to Eppendorf tubes and incubated in primary antibody (α-SMA (mouse), Epcam-1/CD326 (rat) (Invitrogen) or pan-Cytokeratin (mouse) diluted 1:200 in PBSB) for 24-48 hours at 4^{o}C with gentle agitation. Cells were washed in PBSB and left to settle for 10 minutes at room temperature. Pellets were resuspended in appropriate secondary antibody diluted 1:250 in PBSB as outlined in table 7 and incubated at room temperature in the dark for 2 hours. Organoids were washed in PBSB and allowed to settle for 10 minutes at room temperature. Pellets were resuspended in 200µL Vectashield with DAPI using a PBSB coated pipette and placed onto the centre of a glass slide. Coverslips were added and slides were kept in the dark at 4^{o}C overnight. Visualisation was performed using a Zeiss LSM 780 confocal microscope (Carl Zeiss AG, Oberkochen, Germany). The 10x objective lens was used to locate organoids on the slide and images were taken with the x20 objective lens. Zeiss Zen 2 software (Carl Zeiss AG, Oberkochen, Germany) was used to adjust the gain and save images.

Protein of interest	Primary antibody dilution	Secondary antibody	Secondary antibody dilution
CD326/EpCam1(eBioscience™, #14-4321-82)	1:200	Goat anti rat 488 Abcam	1:250
α-SMA (Invitrogen #MA5-11547)	1:200	Donkey anti mouse 568 Invitrogen	1:250

Table 2.8: Antibodies used for PDO characterisation via IF. *Table detailing primary and secondary antibodies used to identify tumour and fibroblast cells, product codes and dilutions used. CD326 was utilised to detect tumour cells and α-SMA was used to identify the presence of stromal population in the organoid cultures. Antibodies were diluted as shown in table in PBS with 0.1% bovine serum albumin.*

2.7.3 Inhibition of JAK-STAT3 in patient-derived organoids

Organoids were grown to 70-80% confluence in 24 well plates. Wells were washed in cold PBS and PDOs were incubated at 37ºC 5% CO_2 in PBS-EDTA for 20 minutes. PDOs from 2 wells were combined, transferred to an Eppendorf tube, and spun at 300RPM for 2 minutes. The pellet was resuspended in ~200µL BME and seeded into a pre-warmed 96 well plate with 5µL added per well. Plates were inverted and incubated for 37ºC for 30 minutes. ADMEM plus additives (Table 6) was added to each well (200µL/well). Samples were incubated at 37ºC 5% CO_2 for ~3-5 days to allow for organoids to reform. Medium was removed and replaced with either fresh medium, 0.01% DMSO, 10µM-1mM Tofacitinib, 10µM-1mM Ruxolitinib or 10µM-1mM 5FU in triplicate. Plates were incubated for 72 hours at 37oC 5% CO_2.

2.7.3.1 Cell viability of patient-derived organoids following JAK/STAT3 inhibition

After incubation images were taken of each treatment using a Zeiss light microscope (Carl Zeiss AG, Oberkochen, Germany) using the x20 objective lens to record morphological changes. Next, 20µL WST-1 reagent was added to each well in the dark. Plates were incubated at 37ºC 5% CO_2 for 2 hours and absorbance was read at 450nM on the TECAN infinite 200 PRO plate reader (Tecan Group Ltd, Switzerland). Organoids from each patient were screened in triplicate and average calculated where availability of samples allowed. Paired T-tests were performed in Microsoft Excel (Microsoft,

Albuquerque, NM, USA) to determine any changes in cell viability between vehicle controls and treated wells.

2.7.3.2 IF staining of patient-derived organoids following JAK/STAT3 inhibition

After drug treatments, 3 patient-derived organoid Sangerlines were expanded sufficiently to perform IF staining for markers of mid-phase apoptosis (Caspase-8, Novus, discontinued) and proliferation (Ki67, Dako #M7240). IF staining was performed as previously outlined in 2.7.2. DMSO-treated sample Ki67 and Caspase-8 expression was compared to Ruxolitinib-treated organoids in Sanger31, 25 and 37 cases.

2.7.3.3 Histology of tumour resections from tumouroid patients

To determine histological subtypes of patient tissue from which Sangerand Glasgow organoids were derived the original resection blocks were requested from NHS GGC Biorepository. H&E staining was performed by Glasgow Tissue Research Facility and Ki67 IHC was performed by KP using the protocol outlined in 2.1.2.2. KM grading, TSP and KI67 index were all assessed as previously outlined. GMS and phenotypic subtype were determined by a single observer (KP) and assessed for pattern of response to Jakinib in the PDO model system. Sections were also stained via IHC for $pSTAT3^{tyr705}$ to assess baseline expression of pathway activation. Scoring was performed using weighted histoscore methods identical to that performed on the Glasgow combined cohort TMAs.

2.8 Statistical Analyses summary

2.8.1 Analyses of tumour phenotype, IHC and RNAScope™ data

Statistical analyses of tumour phenotype, IHC and RNAScope™ data were performed in SPSS (IBM, NY, USA). Kaplan Meier survival curves were plotted to determine association between groups and cancer-specific survival. To determine associations with clinicopathological features chi-squared tests were performed. Significance was set to $\alpha<0.05$.

2.8.2 Analysis of Mutational data

Mutational profiling data were analysed in R Studio (IBM, NY, USA) using the maftools package. Fishers' exact tests were utilised to determine any differential patterns of mutation between $pSTAT3^{tyr705}$ and GMS patient groupings. Copy number analysis was performed by KP and Dr Akhill Yeduresi using Python (Python Software Foundation, Wilmington, DE, USA). Significance was set to $\alpha<0.05$.

2.8.3 Analysis of transcriptomics data

RNAseq data analysis was mainly outsourced and performed by Bioclavis (Bioclavis, Glasgow, UK). Data were normalised using DESeq2 in RStudio (RStudio, Boston, MA, USA). Principal component analyses enabled investigation in top gene clustering between GMS and $pSTAT3^{tyr705}$ groups. SPSS v22 (IBM, NY, USA) was utilised to perform Kruskal Wallis tests and construct box plots for differentially expressed genes. Over-representation analyses were performed in R Studio using go_enrich to analyse pathway enrichment and associations of differentially expressed genes. Volcano plots were constricted by Dr Gerard Lynch using ggplot in R Studio to visualise differential gene expression. Significance was set to padj. <0.05.

2.8.4 Analysis of 2D cell line and 3D tumouroid data

Data from cell viability assays were analysed using Microsoft Excel (Microsoft Corporation, Albuquerque, NM, USA). Raw optical density read outs were averaged and normalised to vehicle controls. Paired t tests were used to determine any significant differences between treatment groups. Significance was set to $\alpha<0.05$.

Chapter 3:

Assessment of Glasgow Microenvironment Score and Phenotypic Subtype as prognostic factors in colorectal cancer clinical specimens by haematoxylin and eosin staining and immunohistochemistry

3.1 Introduction

Colorectal cancer is the third most commonly diagnosed cancer worldwide and 5-year survival rates are ~60%. Currently in the clinic, disease is segregated using tumour node metastases (TNM) staging (101). This prognostic classification system assigns patients to one of 4 groups (I/II/III/IV) based on the depth of tumour invasion into the bowel wall, the number of positive local lymph nodes and presence of metastases. Although this method can be used to predict patients' outcomes, there is a vast degree of heterogeneity amongst patients that fall into the same TNM stage (102). Given the advances in precision medicine for oncology, TNM staging is now considered by many in the field as outdated. Research has focussed on identifying new ways of determining patient prognosis with an aim of establishing tools that predict response to different treatment regimes. In 2015 the consensus molecular subtypes were devised, which group patients based on genomic and transcriptomic tumour features to form 4 groups: immune, canonical, metabolic, and stromal (2). Although the CMS are a useful tool for combating tumour heterogeneity, this method relies on complex next generation sequencing technologies that are not yet feasible for clinical laboratory use. In 2015 Park et al developed Glasgow Microenvironment Score for CRC, a simple histological measure based on Klintrup-Mäkinen Grading (KM) and tumour stroma percentage (TSP) of an H&E slide to give 3 prognostic groups (103). Additionally, in 2017 there was further development of histology-based measures with phenotypic subtyping devised by Roseweir et al (104). This method complements GMS with the addition of tumour proliferation index via Ki67 immunohistochemistry to form 4 distinct independently prognostic CRC groups; immune, canonical, latent, and stromal. The laboratory methods required for both phenotypic measures are already used in routine diagnostics and could be easily translated to clinical practise unlike the components of CMS.

KM grade is a measure of the density of immune cells at the invasive edge of the tumour. Although this measure is highly prognostic is it also of importance to understand the composition of specific populations of immune cells. This has been assessed in prognostic histological subtyping method immunoscore which measure the ratio and spatial relationship of total CD3+ T cells, and cytotoxic CD8+ T cells. It is hypothesised that patients with a high KM grade are likely to have a strong influx of anti-tumour lymphocytes. In contrast, the presence of a high TSP is associated with the presence of a myeloid populations, specifically cells functioning in a pro-tumour fashion.

Retrospective patient samples are an invaluable resource in cancer research to investigate the relationship between novel biomarkers, prognosis and other clinicopathological factors. In this study five independent colorectal cancer cohorts were utilised to assess the clinical utility of tumour histopathological phenotype measures, namely the GMS and Phenotypic Subtype method. The Glasgow combined cohort was used as a training cohort to assess tumour phenotype in stage I-IV primary operable CRC. Patient tissue from the SCOT clinical trial was available to be utilised as a

stage II-III CRC validation cohort. To determine utility of GMS and phenotypic subtype in stage IV CRC, the synchronous resection cohort provided access to both primary CRC tumour and matched synchronously resected metastatic liver tissue. Lastly, the DM-CRC-TMA screen-detected cohort was used to investigate tumour phenotypic measures in a stage I CRC screening cohort by looking for associations with prognostic clinicopathological features. An independent cohort consisting of patient tissue from Australia was used to validate GMS and Phenotypic Subtyping methods using a tissue microarray as opposed to full tumour sections. Outcomes were assessed by cancer-specific survival (time from diagnosis to cancer death) in the Glasgow combined cohort, synchronous resections cohort and Australian TMA cohort. Disease-free survival (time to diagnosis to disease recurrence or any death) was used to measure outcome in the TransSCOT cohort.

GMS is determined using an H&E-stained tumour section by combining a measure of the local inflammatory infiltrate (Klintrup-Mäkinen (KM) Grade and intra-tumour stroma percentage (TSP). This gives rise to 3 groups, GMS0 (immune), GMS1 (intermediate) and GMS2 (stromal) (103). Phenotypic subtype is determined using KM Grade, TSP, and proliferation index (% Ki67 positive tumour cells). This segments patients into 4 groups, immune, canonical, latent and stromal (104). Representative images used to determine GMS and phenotypic subtypes are shown in figure 3.1. Published work has shown GMS and phenotypic subtype associate with CSS in stage I-III CRC (49) (105). Patients assigned to GMS0 or the immune phenotypic subtype show the best prognosis, while those with high stromal infiltrates (GMS2/stromal subtype) exhibit the worst outcomes. Current TNM staging used clinically is outdated and subject to inaccuracies. Both GMS and phenotypic subtyping methods would be easily translatable to routine diagnostics and represent promising tools for providing prognostic information for clinicians and patients.

Here, GMS and phenotypic subtyping were performed in 5 independent retrospective colorectal cancer cohorts to assess their prognostic capability across a range of different disease stages, with the hypothesis being that both GMS and phenotypic subtyping would be prognostic in all stage I-IV retrospective CRC cohorts being assessed. This work will also allow for novel biomarkers being studied in these cohorts to be tested for association with tumour phenotype. This chapter aims to validate GMS and phenotypic subtyping in retrospective CRC cohorts and introduce these measures as not only prognostic tools for CRC but also as groups which could guide targeted therapies in the future. It was hypothesised that both histological measures, GMS and phenotypic subtype, would be prognostic in all 5 independent cohorts. Patients with immunologically hot tumours would be predicted to observe the best clinical outcomes and those with stromal-rich tumours likely to exhibit the shortest survival times.

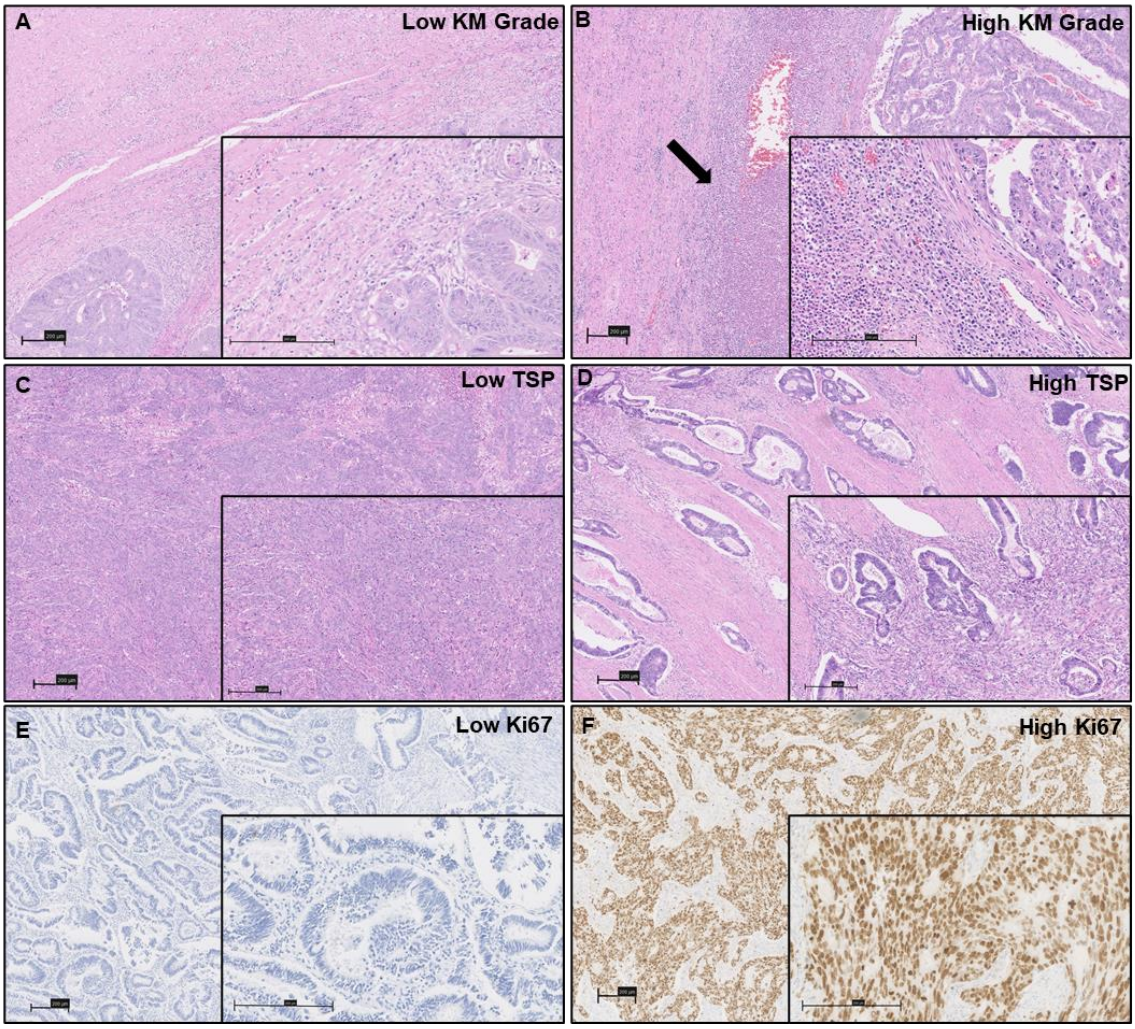


Figure 3.1 Representative images used to determine GMS and Phenotypic subtype. *H&E mages taken at x5 and x20 magnifications showing an example of low (A) and high (B) KM grade at the invasive edge, arrow indicates band of immune cells. H&E images taken at x5 and x20 magnification showing examples of low (C) and high (D) intra tumour stromal percentage. Representative images of CRC tissues stained for Ki67 with an example of low (E) and high (F) proliferation indexes. Scale bar represents 200µM.*

3.2 Clinicopathological characteristics of patient cohorts

3.2.1 Glasgow combined cohort

A total of 1030 patient samples were identified from a retrospective stage I-IV colorectal cancer database held within the Glasgow Royal Infirmary. Clinical specimens came from patients undergoing surgery with curative intent between 1997 and 2007. Most patients (76.9%) had primary colon cancer and 23.1% had primary rectal cancer. There was a relatively even split between sexes with 48.8% males. Patients were predominantly T stage III (55%), with 28.6% T stage IV, 12.1% T stage II and

4.3% T stage I. Most patients were N stage 0 (62.5%), with 26.3% N stage I and 11.2% N stage II. 32.9% of patients received adjuvant chemotherapy and 82.1% were MMR proficient. Patients were excluded from the study if they received neoadjuvant therapy, died within 30 days of surgery, or presented with stage IV disease. The primary endpoint used was cancer-specific survival defined as time in months from the date of surgery to cancer-related mortality. Clinical factors including age ($p<0.001$), T stage ($P<0.001$), N Stage ($p<0.001$) and administration of adjuvant chemotherapy ($p=0.040$) were associated with colorectal cancer-specific survival.

Clinicopathological Characteristic	n (%)	Clinical outcome significance
Age		
<65	293 (31.8)	<0.001
>65	628 (68.2)	
Sex		
Female	447 (48.5)	0.044
Male	474 (51.5)	
T stage		
I	40 (4.3)	<0.001
II	111 (12.1)	
III	507 (55.0)	
IV	263 (28.6)	
N Stage		
0	574 (62.3)	<0.001
I	241 (26.2)	
II	103 (11.2)	
Tumour subsite		
Right	394 (43.1)	0.521
Left	316 (34.5)	
Rectum	205 (22.3)	
MMR status		
pMMR	739 (80.2)	0.185
dMMR	161 (17.5)	
Adjuvant therapy		
No	218 (67.1)	0.040
Yes	107 (32.9)	

Table 3.1 Clinicopathologic characteristics of Glasgow combined cohort of colorectal cancer patients. *Table showing the number (and %) of patients with clinical characteristics and association with cancer-specific survival outcomes in patients from the Glasgow combined cohort including age, sex, T stage, N stage, tumour subsite, MMR status and administration of adjuvant therapy.*

3.2.2 TransSCOT clinical trial cohort

A total of 2913 colorectal cancer patients from the TransSCOT adjuvant chemotherapy clinical trial (ISRCTN no. 59757862) were included in analysis of GMS. For phenotypic subtype, specimens from 1790 patients were available for analysis (due to a reduced number of Ki67 stained sections compared

to H&Es). Tissue was resected between 2008-2013 from high risk TNM II or TNM III patients during surgeries with curative intent undertaken throughout the UK. Patients were mostly T stage III (58.2%) and N stage I (57.1%). 61% of patients were male and most tumours were located in the colon (82.5%). Disease-free survival (DFS) was used as a primary endpoint defined as survival in months from the date of surgery until recurrence or all-cause mortality. Minimum follow-up was 3 years and patients were excluded from analysis if they died within 30 days of surgery. T stage ($p<0.001$), N stage ($p<0.001$) and tumour site ($p=0.003$) were significantly associated DFS (Table 3.2).

Clinicopathological Characteristic	n (%)	Clinical outcome significance
Sex		
Female	1135 (39)	0.436
Male	1778 (61)	
T stage		
I	78 (2.7)	<0.001
II	250 (8.6)	
III	1696 (58.2)	
IV	889 (30.5)	
N Stage		
0	556 (19.1)	<0.001
I	1663 (57.1)	
II	694 (23.8)	
Site		
Colon	2402 (82.5)	0.003
Rectum	511 (17.5)	

Table 3.2 Clinicopathologic characteristics of TransSCOT cohort. *Table showing the number of patients with clinical characteristics and associations with disease-free survival in the TransSCOT cohort including age, sex, T stage, N stage, and tumour subsite.*

3.2.3 Synchronous resection cohort

A total of 44 stage IV colorectal cancer patients who underwent synchronous resection of primary colorectal tumours and matched liver metastases were included in the cohort. Patients were operated on between April 2002 and June 2010 at Glasgow Royal infirmary. 59% of patients were over the age of 65 years, and 59% were male. Primary tumours were mainly T3 or T4 tumours (91%). Primary tumours were mainly located in the colon (55%) and neoadjuvant chemotherapy was administered to 73% of patients. There were 33 cancer-related deaths, 4 non-cancer related deaths and 5-year survival

was 64%. The primary end point used was cancer-specific survival. In terms of association with outcomes, N stage was the only clinicopathological feature associated with cancer-specific survival (p=0.032). No exclusion criteria were applied before analysis due to the size of this cohort (Table 3.3).

Clinicopathological Characteristic	n (%)	Clinical outcome significance
Age		
<65	18 (41)	0.484
>65	26 (59)	
Sex		
Female	18 (41)	1.000
Male	26 (59)	
T stage		
I/II	4 (9)	0.114
III/IV	41 (91)	
N Stage		
0	18 (41)	0.032
I	18 (41)	
II	8 (18)	
Tumour subsite		
Colon	24 (55)	0.896
Rectum	20 (45)	
Adjuvant therapy		
No	18 (42)	0.409
Yes	25 (58)	
Neoadjuvant therapy		
No	32 (73)	0.655
Yes	12 (27)	

Table 3.3: Clinicopathological characteristics of synchronous resection cohort of colorectal cancer patients. *Table showing the number (and %) of patients exhibiting clinical features and association with cancer-specific survival in the synchronous cohort including age, sex, T stage, tumour subsite, adjuvant therapy and neoadjuvant therapy.*

3.2.4 DM-CRC CRC Screening cohort

A total of 159 faecal occult blood test (FOBT) screen-detected CRC patients undergoing surgery within Greater Glasgow and Clyde health board were included. Approximately two-thirds of patients were over the age of 65 years, and most were male (69%). Tumours were mainly TNM stage 1 (142,

89%) with the remaining TNM stage III (17, 11%). Right colon cancer accounted for 60% of cases, with 27% left colon cancer and 25% rectal cancers. 10% of patients underwent adjuvant chemotherapy. Median follow-up was 91 months, and 8 all-cause deaths were recorded to date. Due to the early stage of cancer patients and limited number of events in this cohort, none of the outlined clinicopathological characteristics were associated with cancer-specific survival (Table 3.4).

Clinicopathological Characteristic	n (%)	Clinical outcome significance
Age		
<65	83 (45.6)	0.462
>65	99 (54.4)	
Sex		
Male	127 (69.0)	0.455
Female	55 (29.9)	
T stage		
I	136 (74.7)	0.355
II	46 (25.3)	
N Stage		
0	108 (58.7)	0.761
I	14 (7.6)	
II	4 (2.2)	
Tumour subsite		
Right colon	26 (14.1)	0.129
Left colon	108 (58.7)	
Rectum	48 (26.1)	
Adjuvant therapy		
No	164 (89.1)	0.427
Yes	18 (9.8)	

Table 3.4; Clinicopathological characteristics of DM-CRC screen-detected stage I colorectal cancer cohort. *Table showing the number of patients (and %) with clinical features category and association of clinical features with cancer-specific survival including age, sex, T stage, N stage, tumour subsite and adjuvant therapy.*

3.2.5 Australian TMA cohort

Tissue micro arrays were available via a collaboration with Prof C Soon Lee at Western Sydney University. A total of 410 Australian patients with primary operable colorectal cancer were included in the cohort. Patients were predominantly over 65 years of age (64.4%) and female (65.7%). In terms of T stage, most patients presented at T stage III (39.5%) with 16.6% T stage II, 9.9% T stage I and 8.3% T stage IV. Over half (53.7%) of patients were N stage 0 and tumours were predominantly located in the rectum (74.6%). 43.1% of patients received adjuvant chemotherapy. Patients were excluded from analysis if they died within 30 days of surgery and the primary end point used was cancer-specific survival. T stage ($p<0.001$), N stage ($p<0.001$), tumour subsite ($p=0.046$) and administration of adjuvant chemotherapy ($p=0.018$) were associated with cancer-specific survival (Table 3.5).

Clinicopathological Characteristic	n (%)	Clinical outcome significance
Age		
<65	42 (13.7)	0.735
>65	264 (64.4)	
Sex		
Male	137 (34.3)	0.797
Female	262 (65.7)	
T stage		
I	29 (9.9)	<0.001
II	68 (16.6)	
III	162 (39.5)	
IV	34 (8.3)	
N Stage		
0	159 (53.7)	<0.001
I	70 (23.6)	
II	67 (22.6)	
Tumour subsite		
Colon	104 (25.4)	0.046
Rectum	306 (74.6)	
Adjuvant therapy		
No	156 (56.9)	0.018
Yes	118 (43.1)	

Table 3.5; Clinicopathological characteristics of Australian TMA colorectal cancer cohort. *Table showing the number (and %) of patients with clinical features and associations of clinical features and cancer-specific survival in the Australian TMA cohort including age, sex, T stage, N stage, tumour subsite and adjuvant therapy.*

3.3 Glasgow Microenvironment Score as a prognostic tool in colorectal cancer

3.3.1 Glasgow Microenvironment Score in the Glasgow combined cohort

Histopathological scores for KM Grade, TSP and GMS were available from a previous study by James H Park. After 30-day mortalities, neoadjuvant therapy and stage IV disease patients were

excluded 921 patients were left for GMS analysis. This cohort consisted of 297 GMS0 patients (33.6%), 442 GMS1 patients (48%) and 144 GMS 2 patients (16.3%). To determine whether GMS associated with cancer-specific survival Kaplan-Meier survival curves were plotted (Figure 3.3A). Kaplan-Meier curve survival analysis showed a significant association between GMS and CSS (HR= 1.901, 95%CI; 1.589-2.275, p<0.001 Figure 3.2). Patients classified as GMS0 observed the best outcomes with a mean survival of 173.006 months (95% CI 165.024-180.988) compared to 145.053 months (95% CI 137.051-153.055) for GMS1 group. Patients with GMS2 tumours exhibited the worst prognosis with a mean survival time of 112.725 months (95% CI 97.386-127.614). MMR status performed via IHC in a diagnostic NHS laboratory was available for this cohort. When cases were stratified based on MMR status, GMS was associated with survival in MMR proficient (pMMR) cases (HR= 2.064, 95% CI; 1.701-2.505, log rank p<0.001) and in MMR deficient (dMMR) cases (n=156) (HR= 1.410, 95%CI; 0.896-2.220, log rank p=0.044) (Figure 3.2). Next, the association between GMS and prognosis was investigated with patients segregated by tumour subsite. In patients with primary right-sided colon tumours (n=375) GMS was associated with CSS (HR= 1.863, 95%CI; 1.417-2.450, log rank p<0.001) (Figure 3.2). Similar relationships between GMS and CSS were observed for left-sided colon tumours (HR=1.904, 95%CI; 1.399-2.591, p<0.001, n=299) and rectal cases (HR= 2.803, 95%CI; 1.465-2.963, p<0.001, n=193) (Figure 3.3A) suggesting GMS can be used to predict outcomes regardless of tumour subsite.

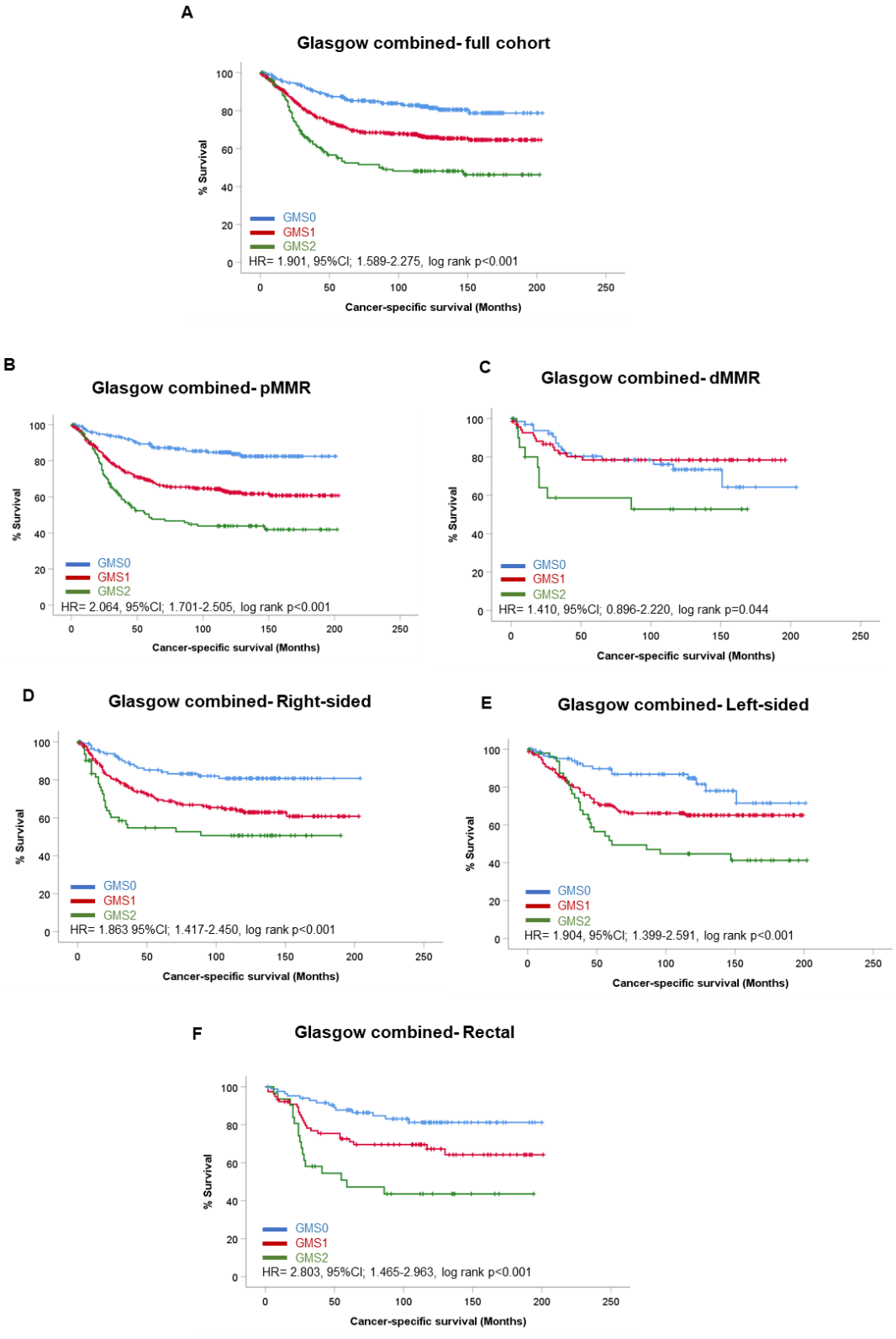


Figure 3.2 Glasgow Microenvironment Score in the Glasgow combined cohort. *Kaplan Meier curves showing association between GMS and cancer-specific survival in the full Glasgow combined cohort (A), MMR proficient patients (B), MMR deficient patients (C), right-sided disease (D), left-sided disease (E) and rectal cases (F). In the full cohort (A) the hazard ratio for GMS was 1.901 (95%CI; 1.589-2.275), log rank p<0.001.*

3.3.2 Glasgow Microenvironment Score in the TransSCOT colorectal cancer clinical trial cohort

Histopathological scores for KM, TSP and GMS were available from a recent study by Mr Peter Alexander and Dr Antonia K Roseweir. After 30-day mortalities were excluded 2887 patients were left for analysis. Of these patients 381 (13.2%) were GMS0, 1848 (64%) were GMS 1 and 658 (22.8%) were GMS2. To determine whether GMS associated with cancer-specific survival Kaplan-Meier survival curves were plotted. GMS was significantly associated with CSS (HR= 1.495, 95%CI; 1.323-1.689, $p<0.001$, Figure 3.3). In validation of the observed results from the Glasgow combined cohort, the GMS0 patients showed the best prognosis with a mean survival time of 69.453 months (95%CI 66.407-72.498). The GMS2 and GMS1 groups exhibited mean survival times of 58.535 months (95%CI 55.697-61.373) and 66.157 months (95%CI 64.247-68.068) respectively. When the association between GMS and DFS were analysed with respect to disease subsite, GMS was only significantly associated with prognosis in the colon cases (HR=1.502, 95%CI; 1.310-1.723, $p<0.001$) and not in rectal cases (Figure 3.3). Chemotherapy data was available for the TransSCOT clinical trial cohort as some patients received one of two combinations: FOLFOX (folinic acid, fluorouracil, oxaliplatin) or CAPOX (oxaliplatin & capecitabine). GMS was significantly associated with DFS in patients from both treatment arms ($p<0.001$), validating the observations seen in the Glasgow combined cohort (Figure 3.3).

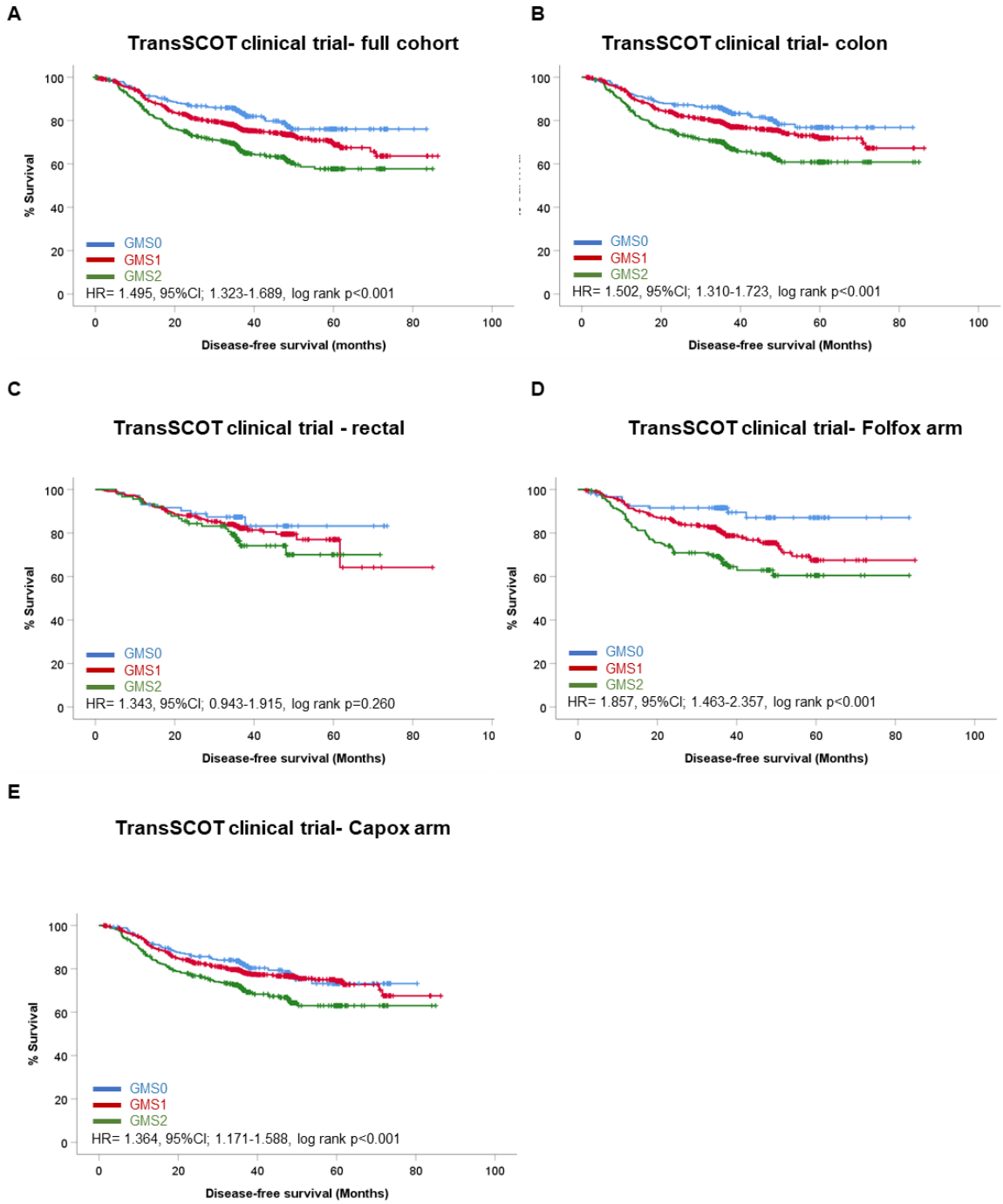


Figure 3.3 Glasgow microenvironment score in the TransSCOT clinical trial cohort. *Kaplan meier curves showing association between GMS and CSS in the full TransSCOT cohort (A), colon cases (B), rectal cases (C), FOLFOX-treated patients (D) and CAPOX-treated patients (E). In the full cohort the hazard ratio for GMS was 1.495 (95%CI; 1.323-1.689) log rank p=0.260.*

3.3.3 Glasgow Microenvironment Score in the Synchronous resection cohort

KM Grade and TSP scoring was undertaken by a single observer (KP) as previously described (106). GMS was calculated for a total of 42 primary tumours and 32 metastatic liver tumours, with some patients excluded due to missing/damaged sections. In terms of the primary tumours 22 (51.2%) were GMS0, 11 (25.6%) were GMS1 and 10 (23.3%) were GMS2. Metastatic tumours followed a similar pattern of segmentation with 19 patients categorised as GMS0 (59.4%), 5 as GMS1 (15.36%) and 8 as GMS2 (25.0%). To determine whether GMS associated with cancer-specific survival Kaplan-Meier survival curves were plotted. Kaplan-Meier curve survival analysis showed a significant association between GMS and CSS in stage IV primary CRC tumours (HR= 1.488, 95%CI; 0.961-2.304, p=0.004, Figure 3.4). Patients with GMS0 tumours had a mean survival time of 80.938 months (95%CI 61.070-100.807) compared to GMS2 patients at 38.247 months (95% CI 29.328-47.166). GMS1 patients showed the worst prognosis in this cohort with a median survival time of 27.947 months (95% CI 21.148-34.476). GMS of matched liver metastases was also significantly associated with CSS (HR=1.403, 95%CI; 0.896-2.196, p=0.017, Figure 3.4). Metastatic tumours graded GMS0 showed the best prognosis with a median survival time of 74.623 months compared to GMS2 tumours at 38.196 months (95% CI 28.426-47.967). GMS2 patients had the worst prognosis with median survival time 35.085 months (95% CI 16.956-53.214). Due to the smaller number of patients GMS was not investigated in relation to tumour subsite in this cohort as analysis would be under powered.

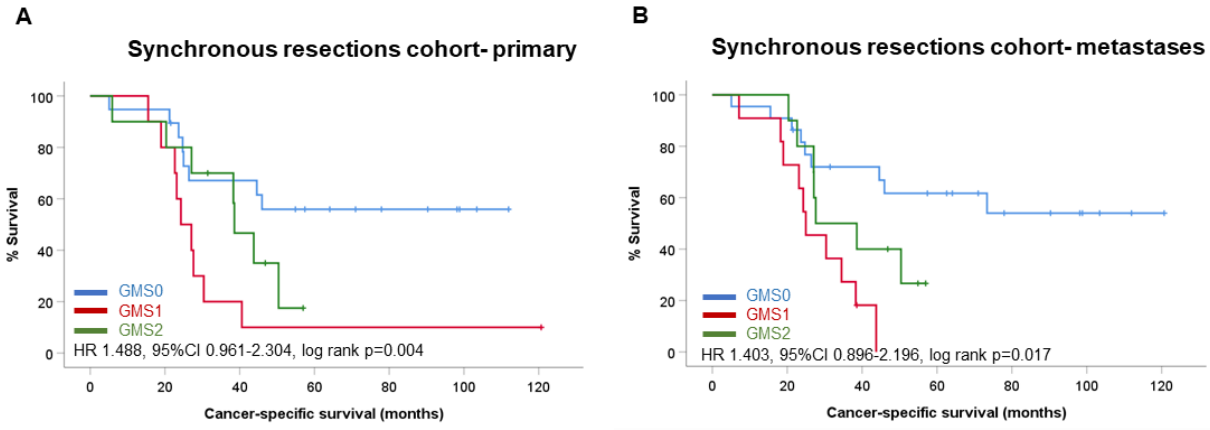


Figure 3.4 Glasgow microenvironment score in the synchronous resection stage IV CRC cohort. *Kaplan Meier curves showing association between GMS and CSS in the primary colorectal tumours (A) and synchronously resected matched liver metastases (B). In primary tumours the hazard ratio for GMS was 1.488, (95%CI; 0.961-2.304), log rank p= 0.004) and in liver metastases the hazard ratio was 1.403, 95%CI; 0.896-2.196), log rank p=0.017.*

3.3.4 Glasgow Microenvironment Score in the DM-CRC screen-detected CRC cohort

In the DM-CRC screen-detected cohort the clinical utility of GMS was assessed by testing for associations with known prognostic clinicopathological features due to the limited number of events. KM grade and TSP were scored manually by a single observer (KP). Due to missing/damaged sections and 30-day mortality exclusions a total of 110 patients were analysed. In this cohort, 33 (42.9%) patients had GMS0 tumours, 33 (42.9%) had GMS1 tumours and 11 (14.3%) had GMS2 tumours. Chi-squared analysis showed no statistically significant associations between GMS and clinical features tested (Table 3.6).

Clinical feature	GMS			p
	0 (n=33)	1 (n=33)	2 (n=11)	
Age				
<65	17 (51.5)	18 (54.5)	4 (36.4)	0.571
>65	16 (48.5)	15 (45.5)	7 (63.6)	
Sex				
Female	21 (63.6)	24 (72.7)	10 (90.9)	0.354
Male	12 (36.4)	9 (27.3)	1 (9.1)	
T stage				
I	15 (45.5)	14 (42.4)	4 (36.4)	0.867
II	18 (54.5)	19 (57.6)	7 (63.6)	
N Stage				
0	26 (78.8)	25 (75.8)	9 (90.0)	0.603
I	6 (18.2)	5 (15.2)	1 (10.0)	
II	1 (3.0)	3 (9.1)	0 (0.0)	
Site				
Colon	21 (63.6)	26 (78.8)	8 (72.7)	0.392
Rectum	12 (36.4)	7 (21.2)	3 (27.3)	
Vascular Invasion				
Absent	27 (81.8)	21 (65.6)	7 (63.6)	0.260
Present	6 (18.2)	11 (34.4)	4 (36.4)	
Adj. Chemotherapy				
No	26 (78.8)	27 (81.8)	9 (81.8)	0.947
Yes	7 (21.2)	6 (18.2)	2 (18.2)	

Table 3.6 Glasgow microenvironment score and clinicopathologic characteristics in the DM-CRC cohort. *Table showing chi-squared analysis for associations between clinical factors and GMS scores including age, sex, T stage, N stage, tumour subsite, vascular invasion and adjuvant chemotherapy.*

3.3.5 Glasgow Microenvironment Score in the Australian CRC TMA cohort

The Australian TMA cohort was used to determine if using tissue microarrays to score tumours for KM grade, TSP and GMS would be as effective as using full tumour sections. KM grade and TSP were scored manually by a single observer (KP). A total of 133 patients were included in the analysis with some patients excluded due to 30-day mortality, missing survival data and/or missing/damaged cores. The TMA consisted of two cores taken from the intra tumour area, which were used to determine TSP and two cores taken from the tumour edge were used to determine KM grade. There

were 40 patients (25.6%) categorised as GMS0, 62 patients (39.7%) GMS1 and 54 (34.6%) of patients with GMS2 tumours. To determine whether GMS associated with cancer-specific survival Kaplan-Meier survival curves were plotted. GMS determined from Australian cohort TMAs was not associated with cancer-specific survival (HR= 1.006, 95%CI; 0.738-1.372, p=0.968, Figure 3.5), suggesting full section analysis is needed to perform GMS accurately. When patients were stratified by disease subsite GMS was not associated with CSS in primary colon or primary rectal cases (Figure 3.5).

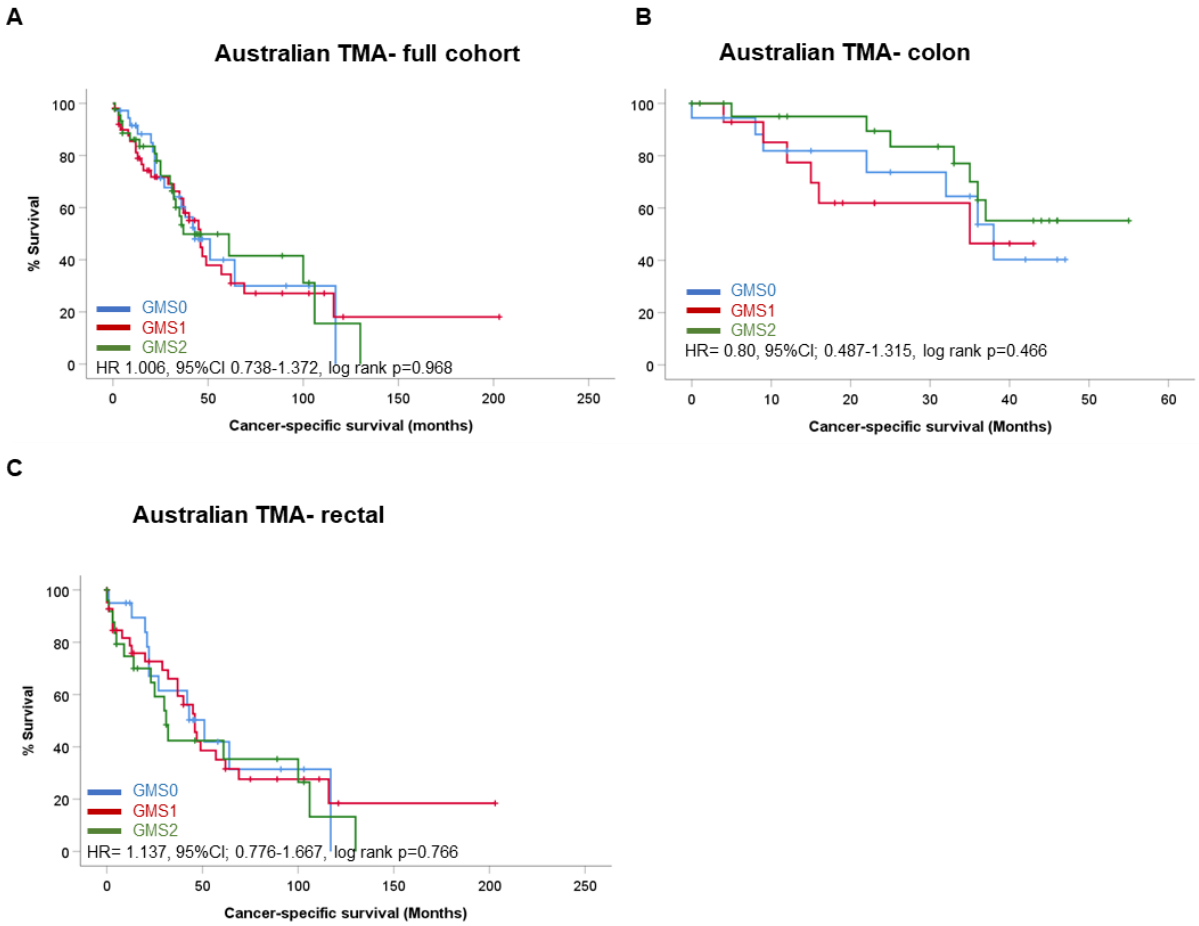


Figure 3.5 Glasgow microenvironment score in the Australian TMA cohort. *Kaplan Meier curves showing association between GMS and cancer-specific survival in the full Australian TMA cohort (A), primary colon cases (B), and primary rectal cases (C). In the full cohort the hazard ratio for GMS was 1.006, (95%CI; 0.738-1.372), log rank p=0.968.*

3.4 Phenotypic Subtyping as a prognostic tool in colorectal cancer

3.4.1 Phenotypic subtype in the Glasgow combined cohort

Data for Klintrup-Mäkinen grade, TSP, Ki67 and phenotypic subtype were already available from a prior study conducted by Dr Antonia K Roseweir. After exclusion criteria were applied there were 921 patients left. Of these patients 299 (34.1%) were grouped as immune, 250 (28.5%) were canonical, 186 (21.2%) were latent and 143 (16.3%) were assigned to the stromal phenotypic subtype. To determine whether phenotypic subtype associated with cancer-specific survival, Kaplan-Meier survival curves were plotted. As previously shown in published work, phenotypic subtype was

significantly associated with cancer-specific survival in this cohort of stage I-III CRC patients (HR= 1.507, 95%CI;1.346-1.687, p<0.001 Fig 3.6). Patients in the immune subtype observed the best prognosis with a mean survival time of 173.228 months (95%CI 165.295-181.160) versus the worst prognostic group, the stromal subtype, which had a mean survival time of 112.037 months (95%CI 97.093-126.980). The canonical and latent subtypes observed mean survival times of 149.525 months (95%CI 139.013-160.038) and 125.810 months (95%CI 125.810-150.429), respectively. When patients were stratified by MMR status, 161 were MMR deficient and 739 were MMR proficient. Phenotypic subtype in pMMR patients was significantly associated with CSS (HR=1.592, 95%CI; 1.411-1.795, p<0.001), however phenotypic subtype was not associated with prognosis in dMMR patients (Figure 3.6). When disease was stratified by tumour subsite, 394 patients presented with right-sided colon tumours, 316 had left-sided primary colon tumours and 205 had tumours located in the rectum. Phenotypic subtype was significantly associated with CSS regardless of tumour subsite, slightly potentiated in rectal cases with the highest HR of 1.613 (95% CI; 1.278-2.036, p<0.001) (Figure 3.6).

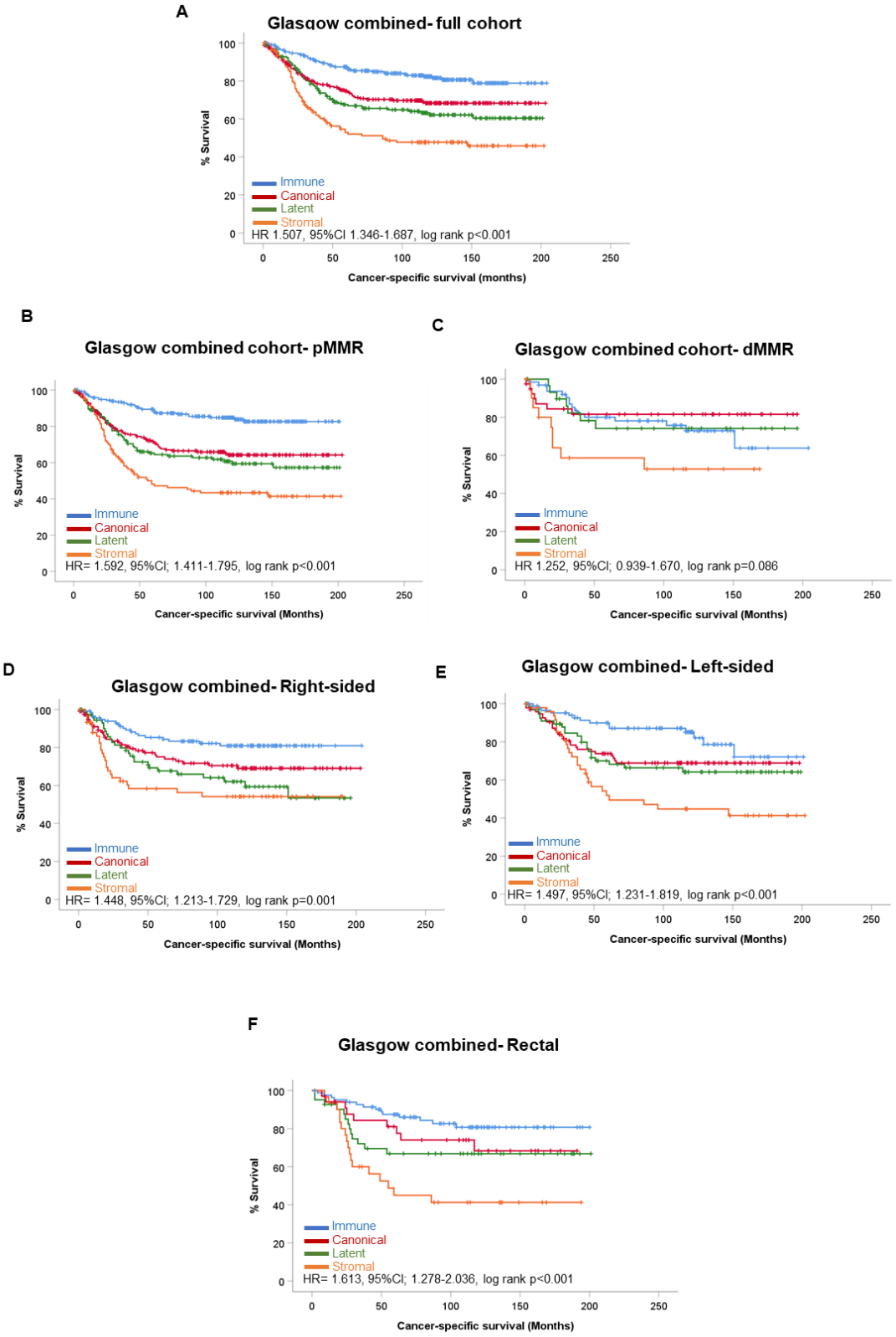


Figure 3.6 Phenotypic subtyping in the Glasgow combined cohort. *Kaplan Meier curves showing association between GMS and cancer-specific survival in the full Glasgow combined cohort (A), MMR proficient patients (B), MMR deficient patients (C), right-sided disease (D), left-sided disease (E) and rectal cases (F). In the full cohort the hazard ratio for phenotypic subtype was 1.507, 95%CI; 1.346-1.687), log rank p<0.001.*

3.4.2 Phenotypic subtype in the TransSCOT colorectal clinical trial cohort

Histological scores for KM Grade, TSP, Ki67 and phenotypic subtype were provided from previous work by Dr Antonia K Roseweir. After excluding patients who died within 30 days of surgery 1776 patients were available for analysis. Of the patients included, 381 (21.5%) were immune, 542 (30.5%) were canonical, 195 (11.0%) were latent and 658 (37.0%) were classified as the stromal phenotypic subtype. To determine whether phenotypic subtype associated with cancer-specific survival Kaplan-Meier survival curves were plotted. As previously described by Roseweir et al 2020, phenotypic subtype was associated with disease-free survival in the TransSCOT colorectal cohort (HR= 1.274, 95%CI; 1.179-1.376, $p<0.001$, Figure 3.7). Patients in the immune subtype showed the best prognosis with a mean survival time of 69.543 months (95%CI 66.407-72.498) compared to the worst prognostic group, the canonical subtype, which observed a mean disease-free survival of 56.816 months (95% CI 53.144-60.487). The latent and stromal subtypes showed mean disease-free survival times of 55.306 months (95%CI 51.290-59.322) and 58.622 months (95%CI 55.785-61.460) respectively. In concordance with GMS, phenotypic subtype was only associated with DFS in colon cases (HR= 1.282, 95%CI; 1.180-1.393, log rank $p<0.001$) and not in rectal cases (Figure 3.7). In terms of treatment regime, phenotypic subtype was significantly associated with DFS in both FOLFOX and CAPOX-treated patients ($p<0.001$, $p=0.002$, respectively, Figure 3.7). However, of note, in the FOLFOX group patients in the immune subtype had a mean survival time of 75.240 months (95%CI; 70.772-79.708 months) compared to the CAPOX-treated patients' immune group who observed a mean survival time of only 65.934 months (95%CI; 62.352-69.515). These data show that patients in the immune subtype respond better to FOLFOX chemotherapy over CAPOX, as previously discussed (47). Patients in the latent subtype observed a trend towards the opposite effect, with the mean survival time of FOLFOX-treated patients 45.395 months versus 57.865 months for CAPOX-treated patients ($p=0.062$).

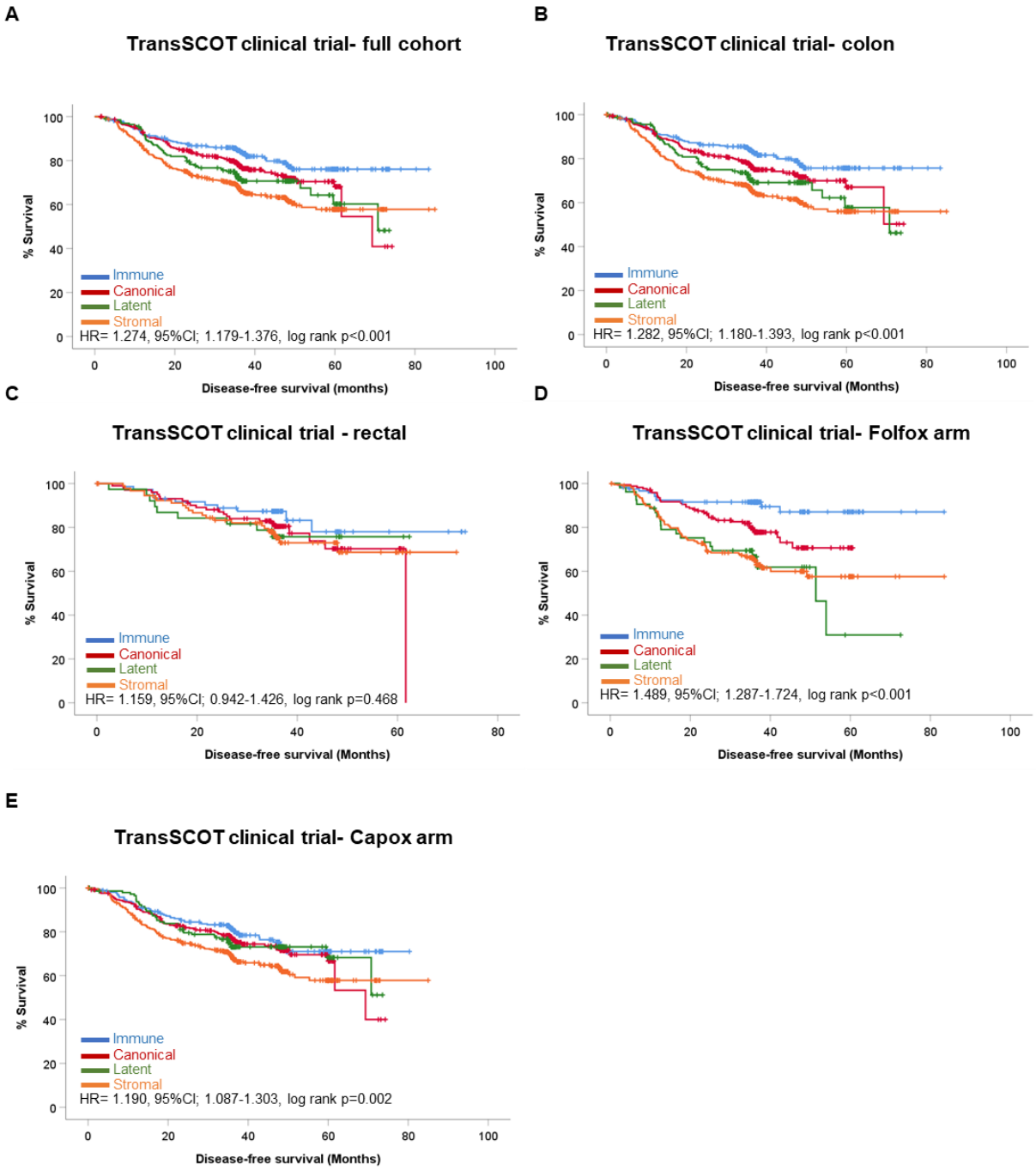


Figure 3.7 Phenotypic subtype and prognosis in the TransSCOT clinical trial cohort. *Kaplan Meier curves showing association between phenotypic subtype and disease-free survival in the full TransSCOT cohort (A), colon cases (B), rectal cases (C), FOLFOX-treated patients (D) and CAPOX-treated patients (E). In the full cohort the hazard ratio for phenotypic subtype was 1.274, (95%CI; 1.179-1.376) log rank p<0.001.*

3.4.3 Phenotypic subtype in the Synchronous resection cohort

KM grade and TSP were scored manually by a single observer (KP) to obtain the components of GMS. Ki67 proliferation index (% tumour cells positive for Ki67) was scored digitally using the Slidepath digital pathology platform (Leica Microsystems, Wetzlar, Germany). The total number of sections scored for KM grade, TSP and Ki67 proliferation index was 42 for primary tumours and 32 for liver metastases due to damaged or sections not being scoreable (total cohort n=44 paired samples). In terms of primary tumours, 20 patients (48.8%) were immune, 5 (8.3%) were canonical, 4 (6.7%) were latent and 12 (29.3%) were stromal. When phenotypic subtyping was applied to metastatic tumours 18 patients (54.5%) were immune, 2 patients were canonical (6.1%), 3 patients were latent (9.1%) and 10 patients were categorised as stromal subtype (30.3%). To determine whether phenotypic subtype associated with cancer-specific survival Kaplan-Meier survival curves were plotted. Phenotypic subtype of primary tumours significantly associated with CSS (HR= 1.272, 95%CI; 0.954-1.697, p=0.001, Figure 3.8). Patients in the immune subtype observed the best prognosis with a mean survival time of 80.938 months (95%CI; 61.070-100.87) compared to the stromal subtype group which had a mean survival time of 38.247 months (95%CI; 29.328-47.166). The mean survival time for the canonical subtype was 22.472 months (95%CI; 10.64-34.251) and the latent subtype showed a mean survival of 29.199 (95%CI; 22.525-35.873).

In terms of metastatic liver tumours, phenotypic subtype was trending towards an association with CSS (HR=1.252, 95%CI; 0.926-1.692, p=0.068, Figure 3.8). Patients within the immune group observed the best outcomes with a mean survival time of 74.623 months (95%CI; 54.931-94.315) compared to the worst prognostic group, canonical, which observed a mean survival time of 26.327 months (95%CI; 24.495-28.359). The stromal subtype has a mean survival time of 38.196 months (95%CI; 28.426-47.967) and the latent subtypes mean survival time was 42.152 months (13.377-70.927).

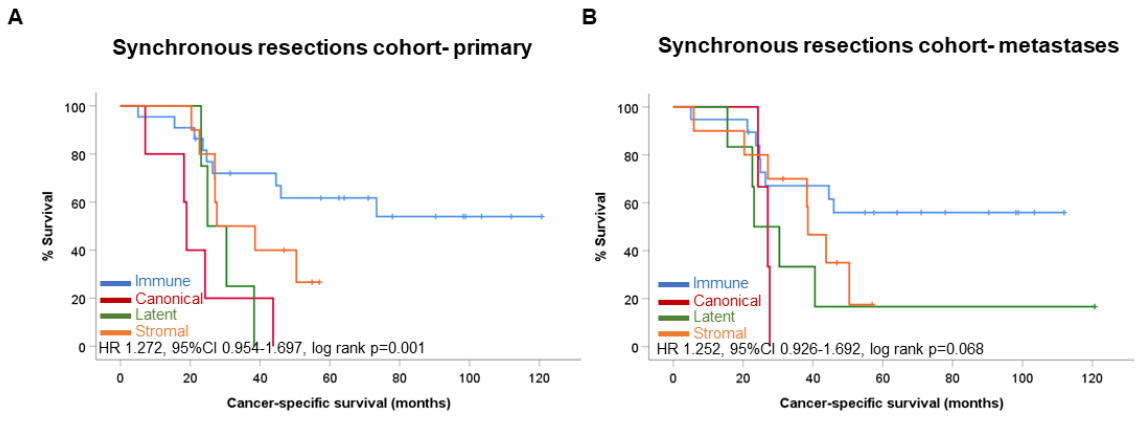


Figure 3.8 Phenotypic subtyping in the synchronous resection stage IV CRC cohort. *Kaplan Meier survival cures showing association between phenotypic subtype and CSS in primary colorectal tumours (A) and matched liver metastases (B). In primary colorectal tumours the hazard ratio for phenotypic subtype was 1.272, (95%CI; 0.954-1.697), log rank p=0.001 and in liver metastases the hazard ratio was 1.252, (95%CI; 0.926-1.692), log rank p=0.068.*

3.4.4 Phenotypic subtype in the DM-CRC screen-detected CRC cohort

KM grade and TSP were scored manually by a single observer (KP). Ki67 proliferation index (% tumour cells positive for Ki67) was scored digitally using the Slidepath digital platform (Leica Microsystems). Due to the small number of events in this cohort, phenotypic subtype was assessed in relation to clinicopathological features and not survival outcomes. When phenotypic subtyping was applied to this cohort, 33 patients were immune (45.2%), 10 patients were canonical (13.7%), 19 patients were latent (26%), and 11 patients were categorised as stromal (15.1%). Chi squared analysis revealed a significant association between phenotypic subtype and tumour site (p=0.041, Table 3.7). No other statistically significant associations were observed between phenotypic subtype and clinicopathological features analysed.

Clinical factor	Phenotypic Subtype				p
	Immune (n=33)	Canonical (n=10)	Latent (n=19)	Stromal (n=11)	
Age					
<65	17 (51.5)	5 (50)	10 (52.6)	4 (36.4)	0.826
>65	16 (48.5)	5 (50)	9 (47.4)	7 (63.6)	
Sex					
Female	21 (63.6)	6 (60)	14 (73.7)	10 (90.9)	0.319
Male	12 (36.4)	4 (40)	5 (26.3)	1 (9.1)	
T stage					
I	15 (45.5)	4 (40)	9 (47.4)	4 (36.4)	0.931
II	18 (54.5)	6 (60)	10 (52.6)	7 (63.6)	
N stage					
0	26 (78.8)	8 (80)	14 (73.7)	9 (90)	0.767
I	6 (18.2)	1 (10)	3 (15.8)	1 (90)	
II	1 (3)	1 (10)	2 (10.5)	0 (0)	
Site					
Colon	21 (63.6)	6 (60)	18 (94.7)	8 (72.7)	0.041
Rectum	12 (36.4)	4 (40)	1 (5.3)	2 (27.3)	
Vascular invasion	27 (81.8)	8 (80)	12 (63.2)	7 (63.6)	0.395
Absent	6 (18.2)	2 (20)	7 (36.8)	4 (36.4)	
Present					
Adj. Chemotherapy	26 (78.8)	8 (80)	17 (89.5)	9 (81.8)	0.785
No	7 (21.2)	2 (20)	2 (10.5)	2 (18.2)	
Yes					

Table 3.7 Phenotypic subtype associations with clinical features. *Table showing Chi squared analysis comparing phenotypic subtype and clinicopathological features including age, sex, T stage, N stage, tumour subsite, vascular invasion and adjuvant chemotherapy.*

3.4.5 Phenotypic subtype in the Australian CRC TMA cohort

KM grade and TSP were scored manually by a single observer (KP). Ki67 proliferation index (% tumour cells positive for Ki67) was scored digitally using the Slidepath digital platform (Leica Microsystems). Phenotypic subtyping was performed in 145 patients, with some patients excluded due to missing survival data, mortality within 30 days of surgery or missing/damaged TMA cores. Of the remaining patients, 40 (27.6%) were categorised as the immune subtype, 29 patients (20.0%) to the canonical subtype, 7 patients (4.8%) as latent subtype and 69 patients (47.6%) to the stromal subtype. To determine whether phenotypic subtype associated with cancer-specific survival Kaplan-Meier survival curves were plotted. Phenotypic subtype was significantly associated with CSS (HR= 1.023, 95%CI; 0.848-1.234, p=0.014 Figure 3.9), however the results did not reflect the observations in the Glasgow combined cohort. Patients in the canonical subtype showed the best prognosis with a mean survival time of 69 months (95%CI32.057-105.943) compared to group with the worst prognosis (latent) with mean survival of 16 months (95%CI 3.556-28.444). The immune subtype had a mean survival time of 43 months (95%CI 29.053-56.947) and the stromal subtype showed a mean survival time of 46 months (95%CI 30.038-61.962). When patients were stratified by disease subsite, phenotypic subtype was not associated with CSS in primary colon (B) or primary rectal (C) cases. This is evidence that TMA sections are not sufficient for performing phenotypic subtyping and full tumour sections are needed.

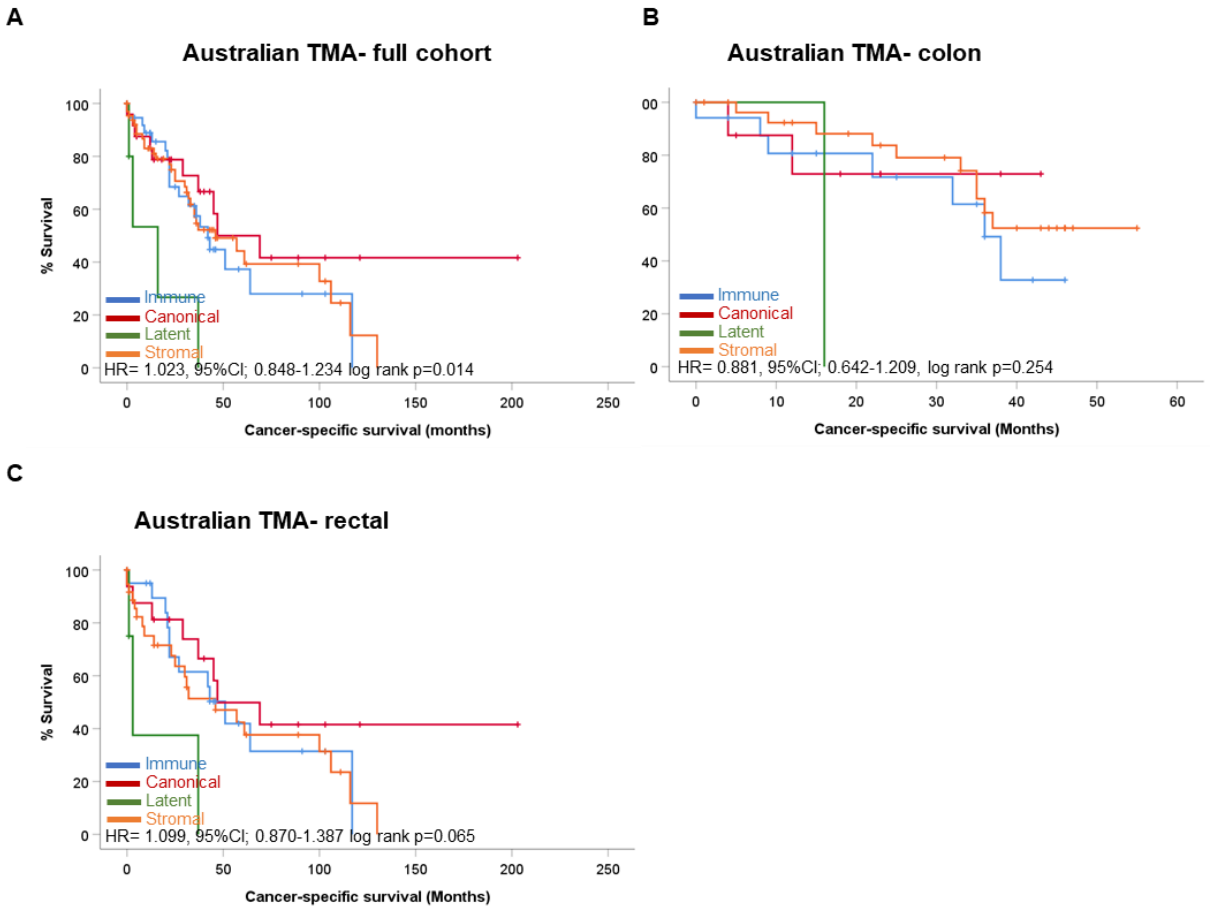


Figure 3.9 Phenotypic subtyping in the Australian TMA cohort. *Kaplan Meier survival cures showing association between phenotypic subtype and CSS in the full Australian TMA cohort (A), colon cases (B) and rectal cases (C). In the full cohort the hazard ratio for phenotypic subtype was 1.023, (95%CI; 0.848-1.234), log rank p=0.014.*

3.5 Concordance of tumour phenotype in metastatic CRC

3.5.1 Comparison of Glasgow Microenvironment Score and Phenotypic Subtyping across matched tumours from the Synchronous resection cohort

The synchronous resection cohort provided unique access to matched primary and metastatic tumours removed during the same surgical procedure. To determine if phenotype of primary tumours was reflective of phenotype of metastatic tumours Spearman correlation analysis was performed. There was a high concordance of GMS between primary tumours and matched metastases (rho= 0.719, p<0.001, Table 3.8). Similarly, phenotypic subtype of primary lesions was also strongly associated with phenotypic subtype of matched liver metastases (rho=0.688, p<0.001). KM grade of primary tumours was strongly associated with KM grade assigned to matched metastases (rho=0.723,

p<0.001). Ki67 proliferation index of primary tumour (<30%/>30%) was strongly associated with matched liver metastases Ki67 proliferation index (rho= 0.675, p<0.001). However, TSP showed discordance between sites (rho=0.186, p=0.326).

Phenotypic measure	rho	p
GMS	0.719	<0.001
Phenotypic Subtype	0.688	<0.001
Klintrup-Mäkinen Grade	0.723	<0.001
Tumour-stroma %	0.186	0.326
Ki67 proliferation index	0.675	<0.001

Table 3.8 Degree of concordance between phenotype of primary and matched metastatic colorectal tumours. *Table showing Spearman correlations for phenotypic measures applied to primary colorectal cancers and matched metastatic liver tumours.*

3.6 Discussion

Colorectal cancer is one of the most common malignancies worldwide and still harbours poor 5-year survival outcomes (12). Research over the past decade has focused on moving treatment towards individualised regimes by subtyping tumours. A lot of this work has focused on genomic and transcriptomic measures, which are currently not translationally relevant to the clinic (104). Phenotyping tumours based on histopathological markers represents an exciting, cost-effective, easily translatable method that could be used initially to predict patient prognosis and potentially guide treatment regimes in the future. Glasgow microenvironment scoring requires a simple H&E section, already available as part of routine clinical diagnostics. Phenotypic subtyping additionally requires a Ki67 IHC stained section, which could easily be adopted to practise as it is already used for subtyping other cancer types. It is not yet clear which is a better measure of tumour phenotype between GMS and phenotypic subtype. It could be argued that GMS is easier to perform, requiring only 1 tumour section, however phenotypic subtype may segment disease better by accounting for heterogeneity present within GMS1.

As previously published, GMS and phenotypic subtype both strongly associated with cancer-specific survival in the Glasgow combined stage I-III CRC cohort, and this relationship was recently validated in the TransSCOT clinical trial cohort (49, 105, 106). Access to the TransSCOT tissue provided, for the first time, an ability to look for associations between phenotypic measures and response to specific chemotherapeutic regimes. The SCOT trial was a non-inferiority trial that compared the use of two chemotherapy types (CAPOX (capecitabine and oxaliplatin) and FOLFOX (bolus and infused fluorouracil and oxaliplatin)) for either 3- or 6-months duration (107). Patients categorised as GMS0 or immune phenotypic subtype, had better clinical outcomes when they received FOLFOX over

CAPOX regimes (105). This work represented a first step towards understanding how histological phenotypic measures could be used clinically to guide optimal treatment for stage II-III CRC patients. Future work is needed to validate these findings and ultimately to look at predicting response to more targeted therapies.

In terms of stage IV disease, utilising the Synchronous resection cohort, both GMS and phenotypic subtype significantly associated with CSS in primary tumours with the immunologically hot tumours observing best clinical outcomes. At the metastatic site a similar trend was observed for both phenotypic measures. The data suggests that in stage IV CRC, presence of a strong inflammatory infiltrate is the most important aspect of tumour phenotype with the immune subtype showing the best prognosis in primary and secondary disease. Although phenotypic subtype only trended towards an association with CSS in metastatic liver tumours in this cohort, spearman correlations showed a significant link between tumour phenotype (both GMS and phenotypic subtype) of primary site and matched metastases. This result is promising as it suggests future targeted therapies designed for each phenotype may be effective against tumours at both sites in patients with stage IV disease.

Data acquired from the Australian TMA suggests that to accurately depict GMS and phenotypic subtype, full tumour section H&Es as opposed to tissue-microarray cores are a necessity. In terms of moving GMS forward into the clinic, it would be optimal for scoring to be performed on pre-treatment tumour biopsies. A TMA core is arguably more similar to a biopsy than a full-face tumour resection. CRIS, but not CMS subtypes have previously been shown to be effectively determined from a tumour biopsy alone (Dunne 2018). Similarly, a surrogate IHC based method of determining CMS was identified in 2017, and this proved prognostic when assessed in TMAs. The method was 87% concordant with matched transcriptomic CMS classification and involved staining for 5 proteins (CDX2, FRMD6, HTR2B, ZEB1, and KER) plus MSI status showing 87% concordance with transcriptionally defined CMS (Trinh et al 2017 CCR). This was an interesting step forward towards making CMS clinically translatable, however is still more complex than GMS and although prognostic, not independently prognostic like GMS.

To tackle the limitations of needing a full face section for determining GMS, in 2019 Park et al showed that using a CD3 IHC stained biopsies to determine immune grading for GMS could be used instead of KM grade, therefore it could be hypothesised that using a TMA would be similarly effective (108). Unfortunately, it was not possible to obtain CD3 stained sections or histological T cell counts for the Australian cohort to validate this finding. It may be of interest to repeat GMS and phenotypic subtyping in this cohort using full sections to determine if lack of prognostic pow(109)er was due to geographical and environmental differences between UK-based and Australian-based patients rather than limitations with sample format.

GMS was not associated with any clinicopathological features within the DM-CRC-TMA gFOBT screen detected cohort, suggesting that phenotype might not be an important measure for early CRC. Phenotypic subtype was significantly associated with tumour subsite in the DM-CRC-TMA cohort ($p=0.041$), with almost all patients in the latent group presenting colonic tumours (94.7%) compared to the canonical group, which consisted of 40% rectal origin. Given that high TSP within tumours predicts higher risk of recurrence in stage II-III disease (110), it may be interesting to monitor the follow up data on these DM-CRC-TMA patients to look for any patterns of recurrence and association with GMS/Phenotypic subtype.

In conclusion, GMS and phenotypic subtyping measures are promising histopathological methods for assessing CRC patient prognosis. Routine diagnostics could easily introduce these scores to clinical practice. Future work should focus on identifying underlying biological factors responsible for driving the observed phenotypes to identify novel and repurposed therapeutics optimally suited to each group. From these data, the presence of a dense inflammatory infiltrate is of profound prognostic significance in stage II through to stage IV CRC. Therefore, investigation into inflammatory signalling pathways will be important for understanding the mechanisms of CRC development, progression, and phenotype.

Chapter 4: Expression of IL6, IL8 and IL6R in colorectal cancer clinical specimens

4.1 Introduction

Interleukin-6 (IL6) is a pleiotropic cytokine associated with inflammation, tissue repair and, tumourigenesis. In the colorectal cancer setting IL6 activates several downstream transcription factors to promote tumour proliferation, differentiation, and metastases (111). Signal transduction is initiated by IL6 binding its cognate receptor, IL6R, which can either be soluble or membrane bound. Ligation results in a cascade of signals ultimately leading to activation of STAT3. Systemic IL6 levels are well-characterised in the literature to be associated with poor prognosis in CRC, however the role of IL6/IL6R expression within CRC tumour tissue is less well characterised. In cancer IL6 is mainly produced by tumour-associated macrophages (TAMs), tumour cells or adjacent stromal cells, and IL6R can be found within/on tumour cells and inflammatory cells (112).

IL6 is generally associated with promoting tumour development and progression through its effects on tumour cells, tumour infiltrating leukocytes and cancer-associated stromal cells. In terms of direct effects on the tumour, IL6 promotes many of the hallmarks of cancer including angiogenesis, proliferation, invasion, (80, 81, 113) The effects of IL6 on the inflammatory infiltrate involves polarization of Macrophages to pro-tumour phenotypes, which produce IL10 leading to an immunosuppressive TME (82). This is perpetuated via IL6 promoting T cell differentiation to a regulatory phenotype via inducing FOXP3 expression, and these Tregs have shown enhanced suppressive capacities (84). Presence of IL6 hinders Th1 cell and dendritic cell differentiation which hinders the cytotoxic functions of CD8+ T cells(83). Additionally, IL6 promotes CAF and tumour cell survival and migration(86). IL6 can promote cancer through a number of mechanisms and has been implicated in each compartment of the TME of CRC.

In addition to IL6, TAMs often produce inflammatory chemokine IL8/CXCL8. The main roles of IL8 identified in tumour progression include recruitment of neutrophils and promotion of angiogenesis and metastases (114). There is a degree of crosstalk between IL6/IL6R and IL8 signalling with overlap between STAT3, NFκB and AKT pathways (115). For example, IL8 activates the canonical NFκB cascade which results in production of IL6. IL8 has also been shown to phosphorylate JAK resulting in activation of STAT3 (114). Additionally, IL6 can activate NFκB resulting in a feedback loop. Both IL6 and IL8 have been implicated in a number of tumour types, including CRC development and progression (112).

The aim of this chapter was to semi-quantitively measure IL6, IL8 and IL6R within patient samples from the Glasgow combined cohort with the hypothesis that high expression of each marker would associate with reduced survival outcomes and unfavourable clinical characteristics. Due to the soluble nature of IL6 and IL8 it was difficult to access protein expression via IHC, as expression is not localised with the cell and appears as a brown blush and non-specific staining. Therefore RNAScope® was utilised to detect mRNA copies of each cytokine per μM^2, enabling quantification

of mRNA expression and the cellular location. IL6R protein expression was assessed in the same patient samples via IHC. For RNAScope®, images were analysed digitally using HALO software within tumour epithelium-rich and stromal areas of each core. For IL6R expression was measured manually by performing weighted histoscoring, with 10% validated by digital pathology software QuPath (116). Expression of IL6, IL8 and IL6R were related to clinical outcome and clinicopathological characteristics. Data were analysed in relation to GMS phenotypes as a step towards understanding if a particular subtype of patients may benefit from therapeutic intervention targeting the IL6/JAK/STAT3 pathway.

4.2 Expression of IL6, IL8 and IL6R in CRC patient specimens

The Glasgow combined array was stained via RNAScope™ for IL6, IL8 and IL6R. A consort diagram was constructed to show the number of patients included in the final set of analysis due to exclusion criteria of damaged/missing cores, mortality within 30 days of surgery and/or administration of neoadjuvant chemotherapy (Figure 4.1).

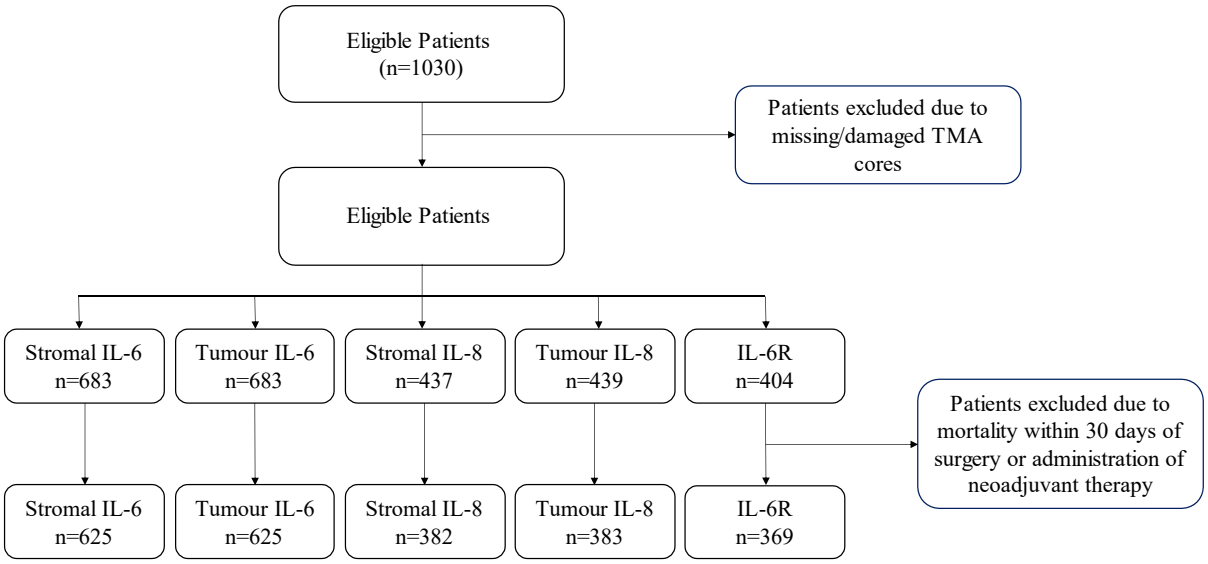


Figure 4.1 Consort diagram showing patient exclusions. *Consort/flow diagram showing the number of patients included in analysis for each marker based on exclusion criteria of missing/damaged cores, mortality within 30 days of surgery and/or administration of neoadjuvant therapy. There were 1030 patients in the full cohort prior to exclusions. After exclusions there were 625 patients analysed for stromal and tumour IL6 mRNA, 382 patients for stromal IL8 mRNA, 383 for tumour IL8 mRNA and 369 patients analysed for intensity of IL6R staining.*

4.2.1 Expression of IL6 within the tumour-surrounding stroma and clinical outcome

RNAScope® was utilised to probe for IL6 mRNA in the Glasgow combined cohort TMA cores. Staining was performed by Colin Nixon at the Beatson Histology unit, Glasgow. In addition to IL6, housekeeping gene PPBIB was probed for in the same cohort. Expression was quantified using Halo digital analysis platform to determine copies per μM^2. A classifier was built to distinguish between tumour epithelium and stromal tissue within cores. Thresholds for intensity of DAB chromogen staining were set and kept constant for each TMA section. Scores were obtained for IL6 in the tumour areas and stroma-rich areas from 3 cores for each patient and an average was taken. Data for IL6 were normalised to housekeeping gene PPIB expression. Normalised data were then entered to the master SPSS database and analysed for association with clinical outcomes and clinical features. Representative images of positive and negative staining of IL6 are shown in Figure 4.2.

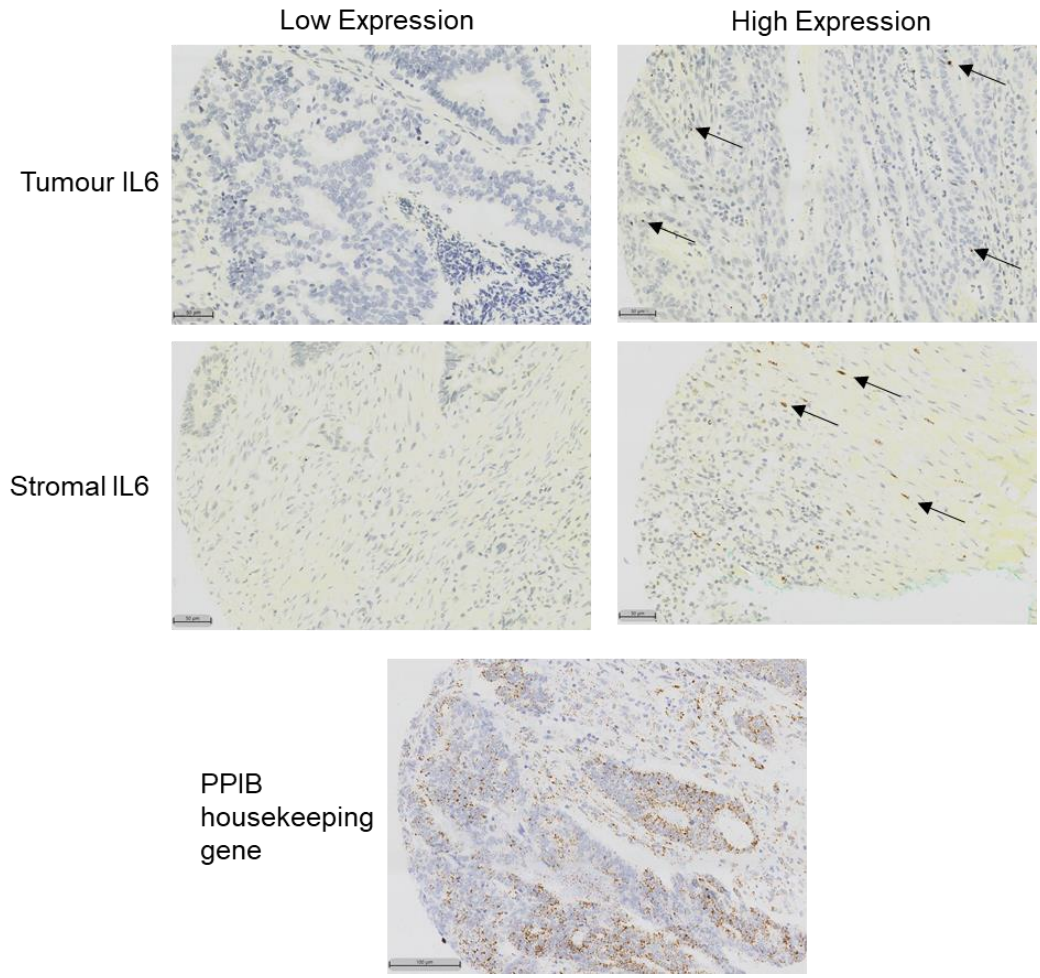


Figure 4.2 Representative images of RNAScope™. Representative images of RNAScope™ high and low staining for IL6 within tumour and stromal areas, and a representative image of PPIB housekeeping gene.

In terms of IL6 within the stromal regions, scores were available for 683 patients from the 1030 Glasgow combined patients due to missing or damaged cores. Patients who died within 30 days of surgery or received neoadjuvant therapy were excluded leaving 625 patients included in further analysis. A histogram was plotted to assess distribution pattern of the data, which showed a positive skew (Figure 4.3). The mean score was 1.117, with a range of 0-29.13 copies per μM^2. Maxstat and survminer packages in R studio were used to determine optimal cut points for high and low expression by log rank statistics based on CSS. Density and scatter plots were constructed to visualise the optimal cut point determined to be 2.48 (Figure 4.4). This resulted in 20 patients classified as high stromal IL6 expression and 605 patients as low for stromal IL6.

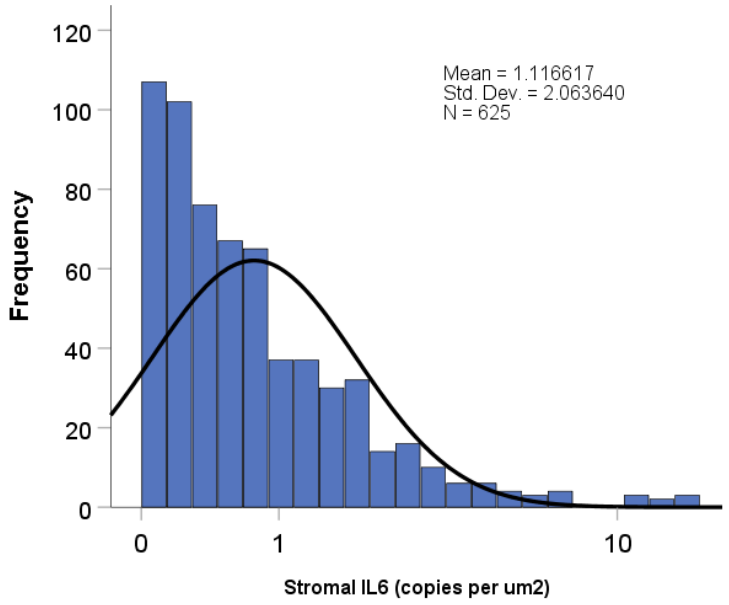


Figure 4.3 Distribution of stromal IL6. *Histogram showing positively skewed distribution pattern of stromal IL6 scores for 625 patients from the Glasgow combined cohort. Mean score for stromal IL6 was 1.117 copies per μM^2 and scores ranged from 0-29.13 copies per μM^2.*

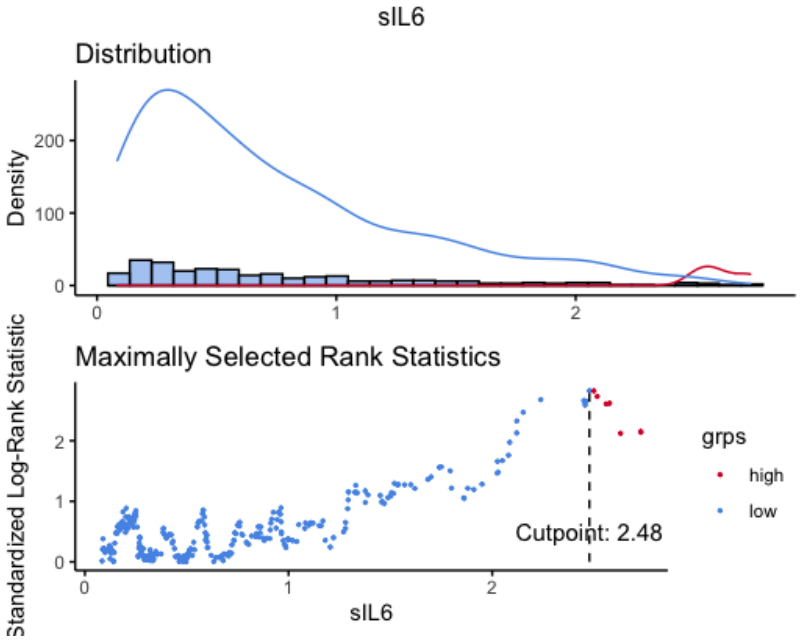


Figure 4.4 Defining optimal cut point for high and low stromal IL6 expression. *Density and scatter plot showing optimal cut off point for high and low expression of stromal IL6 groups based on cancer-specific survival. For stromal IL6 the optimal cut point determined was 2.48 copies per μM^2, so patients with a score of ≥ 2.48 were classified as high and those with a score of ≤ 2.48 were classified as low for stromal IL6 expression.*

High stromal IL6 was associated with reduced CSS in the full cohort (HR= 1.957, 95%CI; 1.000-3.827, log rank p=0.045) (Figure 4.5). Patients with low expression observed a mean survival time of 150.507 months compared to 113.812 months for the high expression group. When patients were stratified by GMS there was no association between stromal IL6 mRNA expression and CSS in any subgroup (Figure 4.6). This may be due to the limited number of patients in the high expression group when subdivided into GMS groups.

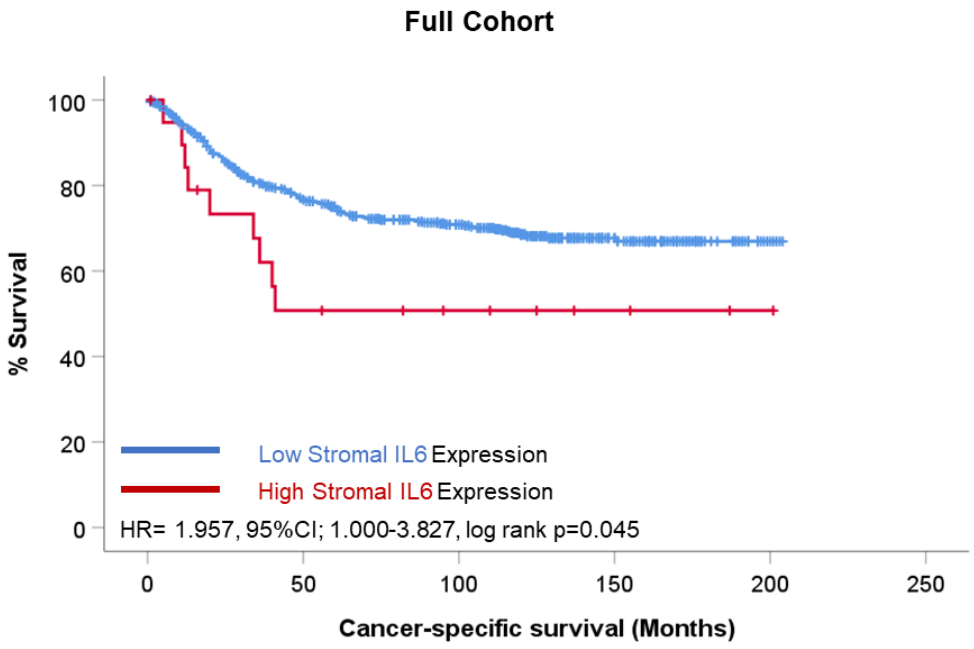


Figure 4.5 Association between stromal IL6 expression and cancer-specific survival.

Kaplan Meier curve showing association between stromal IL6 mRNA expression and CSS in patients from the Glasgow combined cohort with hazard ratio of 1.957 (95% confidence interval 1.000-3.827), log rank p=0.045.

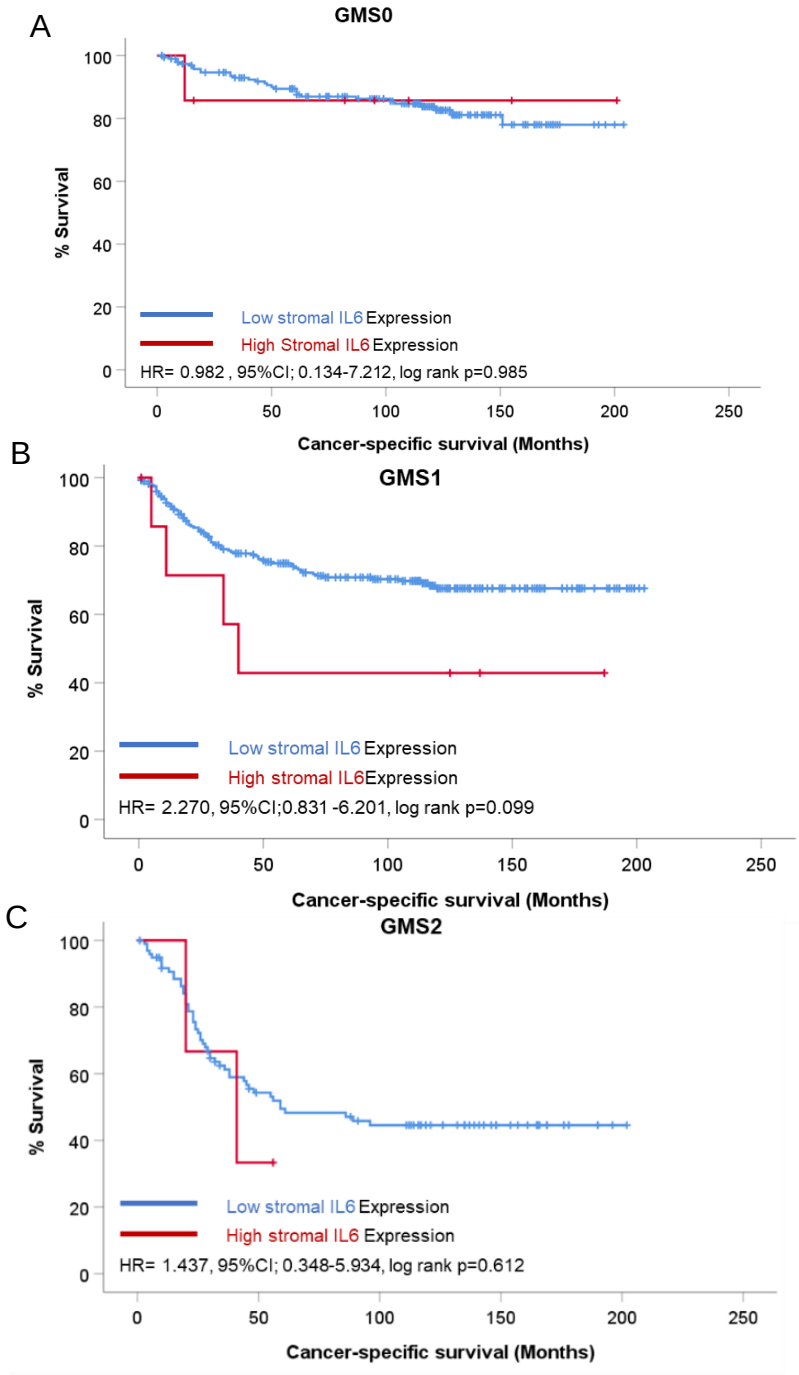


Figure 4.6 Association between stromal IL6 expression and CSS in each GMS classification. *Kaplan Meier survival analysis of IL6 mRNA expression and CSS in patients classified as GMS0 (A), GMS1 (B) and GMS2 (C). Hazard ratio for GMS0 patients was 0.982 (95%CI; 0.134-7.212), log rank p=0.985. GMS1 patients observed a hazard ratio of 2.270 (95%CI; 0.831-6.201), log rank p=0.099 and in GMS2 patients the hazard ratio was 1.437, (95%CI; 0.348-5.934), log rank p=0.612.*

Stromal IL6 expression was strongly associated with reduced CSS in patients with MMR deficient tumours (HR=3.747, 95%CI; 1.114-12.611, log rank p=0.022), but not MMR proficient tumours

(Figure 4.7). Patients with high stromal IL6 expression in dMMR cases had a mean survival time of 66.25 months compared to 161.170 months for the low stromal IL6 expression patients.

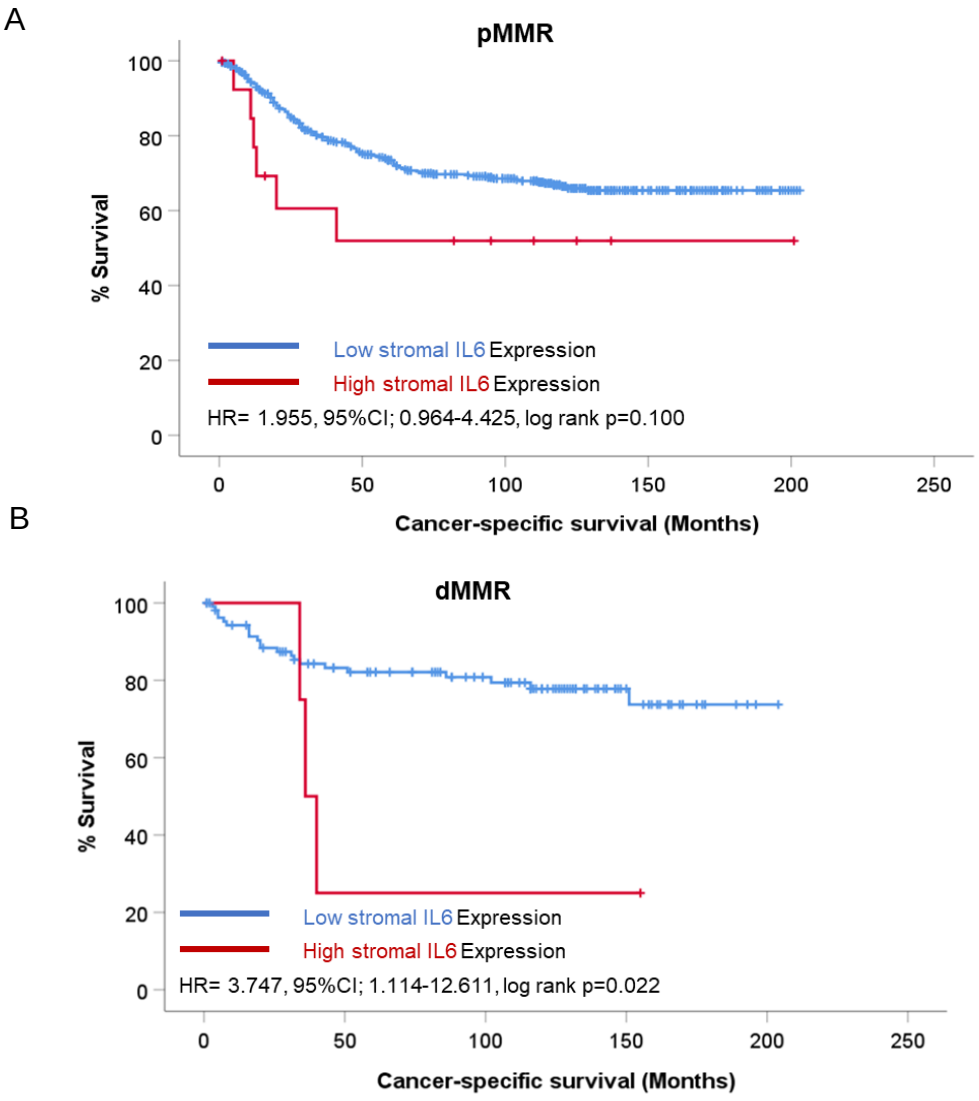


Figure 4.7 Association between stromal IL6 expression and CSS in pMMR and dMMR cases. *Kaplan Meier survival analysis of stromal IL6 mRNA expression and CSS in MMR proficient patients (A) and MMR deficient patients (B). The hazard ratio for stromal IL6 in patients with pMMR disease was 1.955 (95%CI; 0.964-4.425), log rank p=0.100, and in patients with dMMR disease the hazard ratio was 3.747 (95%CI; 1.114-12.611), log rank p=0.022.*

In terms of tumour subsite, there was no significant association between stromal IL6 expression and CSS when Kaplan Meier survival analysis was applied to right-sided, left-sided, or rectal cases (Figure 4.8). Patients with rectal tumours observed a trend towards reduced CSS with high expression of stromal IL6, but the relationship did not reach significance (p=0.052) (Figure 4.8).

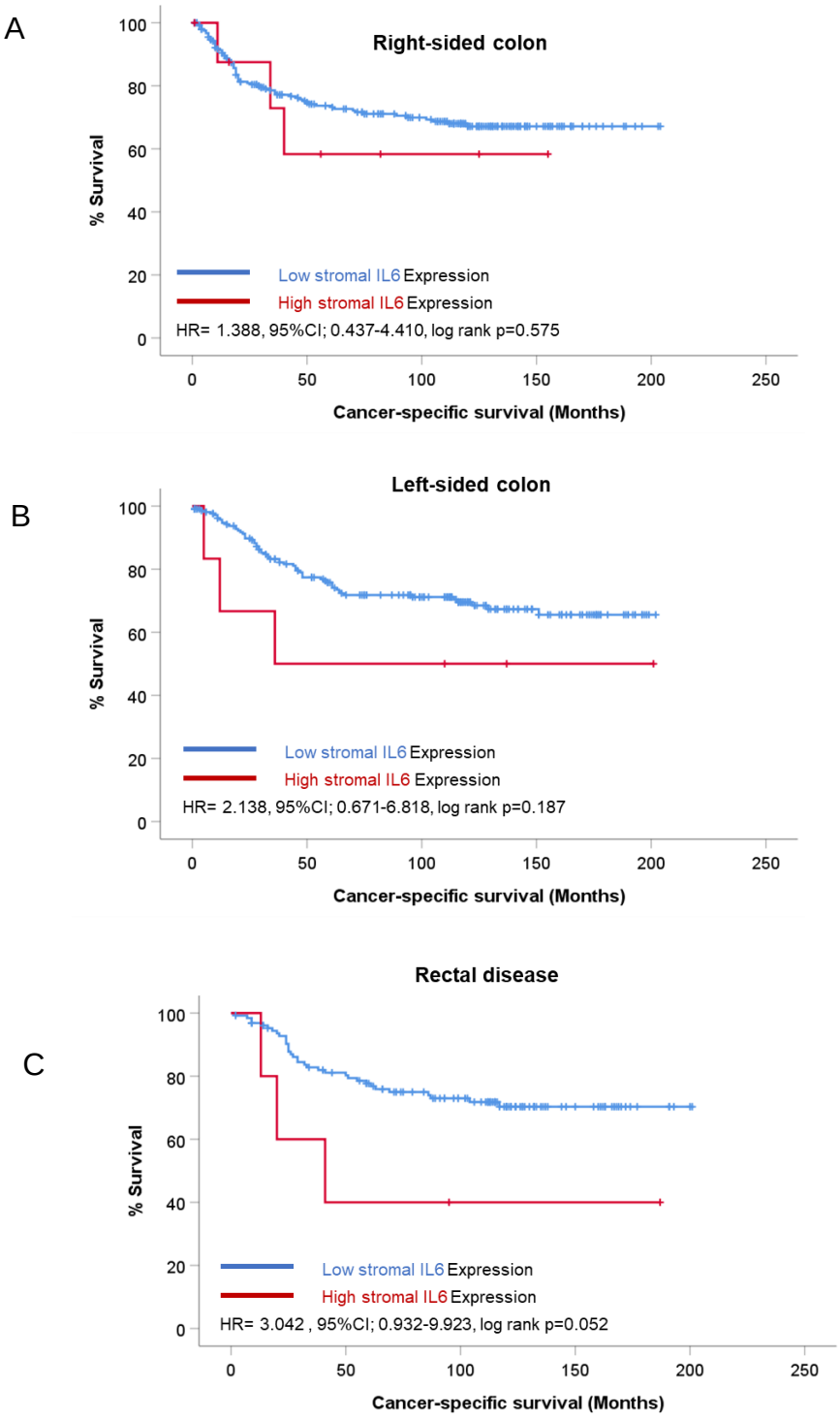


Figure 4.8 IL6 expression, CSS, and tumour subsite. *Kaplan Meier survival analysis showing association between IL6 mRNA expression and CSS in right-sided disease (A), left-sided disease (B) and rectal disease (C). In patients with right-sided disease the hazard ratio for stromal IL6 expression was 1.388 (95%CI; 0.437-4.410) log rank p=0.575 and in left sided cases the hazard ratio was 2.138 (95%CI; 0.671-6.818), log rank p=0.187. In patients with rectal tumours the hazard ratio for stromal IL6 expression was 3.042, (95%CI; 0.932-9.923) log rank p=0.052.*

Chi-squared testing was performed to identify any association between stromal IL6 mRNA expression and clinicopathological characteristics. No significant associations were observed (Table 4.1).

Clinical Factor	Stromal IL6 mRNA Expression		p
	Low (n=605)	High (n=20)	
Age			
<65	197 (32.6)	6 (30.0)	0.512
>65	408 (67.4)	14 (70.0)	
Sex			
Female	298 (59.3)	11 (55.0)	0.391
Male	307 (50.0)	9 (45.0)	
T Stage			
I	24 (4.0)	2 (10.0)	0.451
II	82 (13.6)	1 (5.0)	
III	327 (54.0)	11 (55.0)	
IV	172 (28.4)	6 (30.0)	
N Stage			
0	375 (62.2)	9 (45.0)	0.287
I	158 (26.2)	7 (35.0)	
II	70 (11.6)	4 (20.0)	
Tumour subsite			
Right-sided colon	256 (42.7)	9 (45.0)	0.832
Left-sided colon	217 (36.2)	6 (30.0)	
Rectum	127 (21.2)	5 (25.0)	
GMS			
0	200 (34.3)	7 (38.9)	0.921
1	282 (48.4)	8 (44.4)	
2	101 (17.3)	3 (16.7)	
Phenotypic Subtype			
1	201 (34.5)	7 (38.9)	0.611
2	170 (29.2)	3 (16.7)	
3	111 (19.1)	5 (27.8)	
4	100 (17.2)	3 (16.7)	
mGPS			
0	248 (52.9)	8 (50.0)	0.789
1	136 (29.0)	4 (25.0)	
2	85 (18.8)	4 (25.0)	
MMR status			
pMMR	495 (82.0)	14 (77.8)	0.419
dMMR	109 (18.0)	4 (22.2)	
Tumour differentiation			
0	536 (88.6)	16 (80.0)	0.196
1	69 (11.4)	4 (20.0)	
Marginal involvement			
0	578 (95.5)	18 (90.0)	0.236
1	27 (4.5)	2 (10.0)	
Vascular invasion			
0	406 (67.1)	11 (55.0)	0.185
1	199 (32.9)	9 (45.0)	

Table 4.1 Association between IL6 expression and clinical features. *Chi-squared table of associations for IL6 mRNA expression and clinical prognostic factors including age, sex, T stage, N stage, tumour subsite, GMS, phenotypic subtype, systemic inflammation measured via modified Glasgow Prognostic Score (mGPS), MMR status, tumour differentiation, marginal involvement and vascular invasion.*

4.2.2 Expression of IL6 within the tumour epithelium and clinical outcome

IL6 mRNA was detected within the tumour epithelium in addition to the stromal compartment of TMA cores. Halo software was utilised to quantify IL6 expression in copies per μM^2. Scores were normalised to expression levels of PPBIB housekeeping gene scores obtained for the same core area. Tumoural IL6 data were collected for 683 patients from the Glasgow combined cohort. Patients who died within 30 days of surgery or received neoadjuvant therapy were excluded leaving 625 patients included in the analysis. Scores ranged from 0 to 11.28 copies per $\mu M^{2,}$ and data were positively skewed as can be observed in histogram plot (Figure 4.9). Survminer and maxstat packages in R studio were utilised to determine an optimal cut point for high and low expression of IL6 within the tumour stroma. Density and scatter plots were constructed to visualise the optimal cut point based on CSS (Figure 4.10). There were 191 patients classed as high and 191 classed as low for tumoural IL6 based on a cut point of 0.05 copies per μM^2.

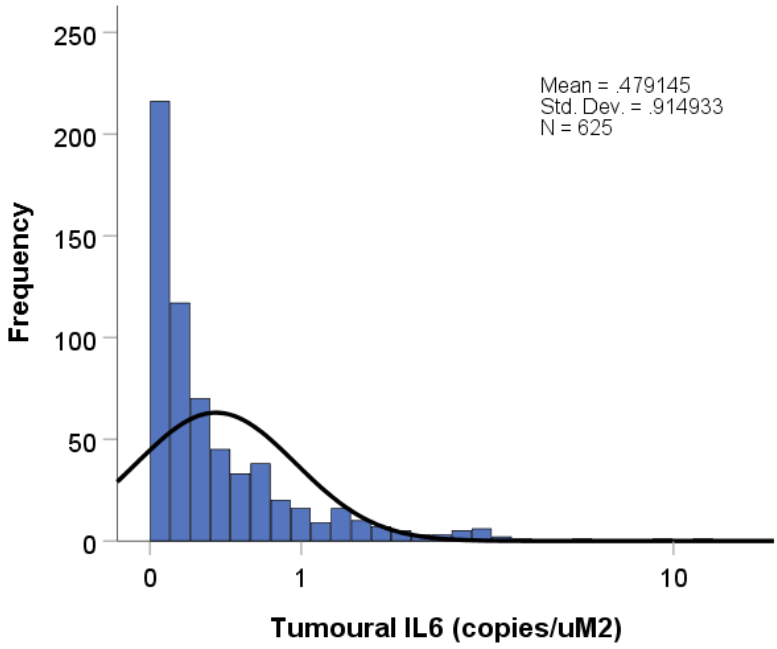


Figure 4.9 Distribution of tumoural IL6. *Histogram showing positively skewed distribution curve for tumoural IL6 scores (n=625). The mean value for tumour IL6 was 0.48 copies per μM^2 and scores ranged from 0-11.28 copies per μM^2.*

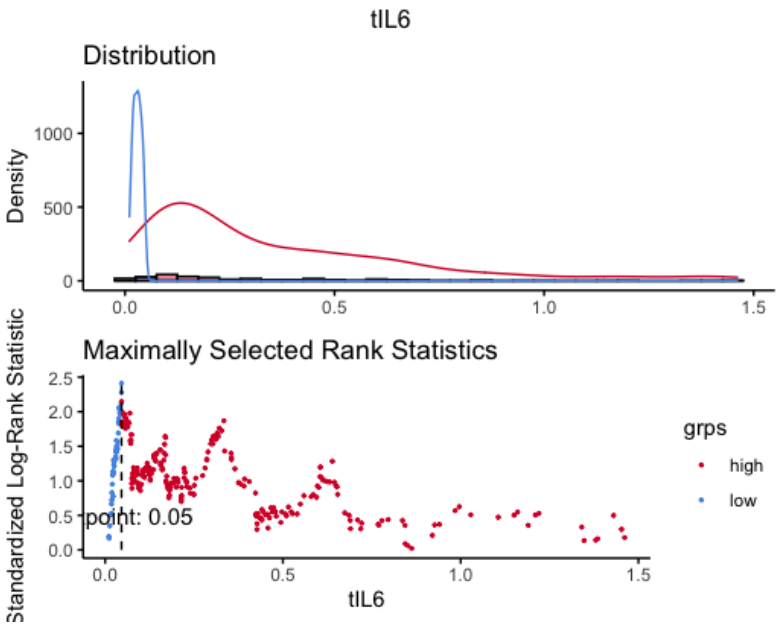


Figure 4.10 Defining optimal cut point for high and low tumour IL6 expression. *Density and scatter plot showing optimal cut off point for high and low expression of tumour IL6 groups based on cancer-specific survival. The optimal cut point determined was 0.05 copies per μM^2 therefore patients with a score of ≥0.05 copies per μM^2* were classified as high expression and patients with a score of ≤0.05 *copies per μM^2 were classified as high for expression of tumour IL6.*

In contrast to IL6 within the stroma, there was no association between tumoural IL6 mRNA expression and CSS in the full cohort (Figure 4.11).

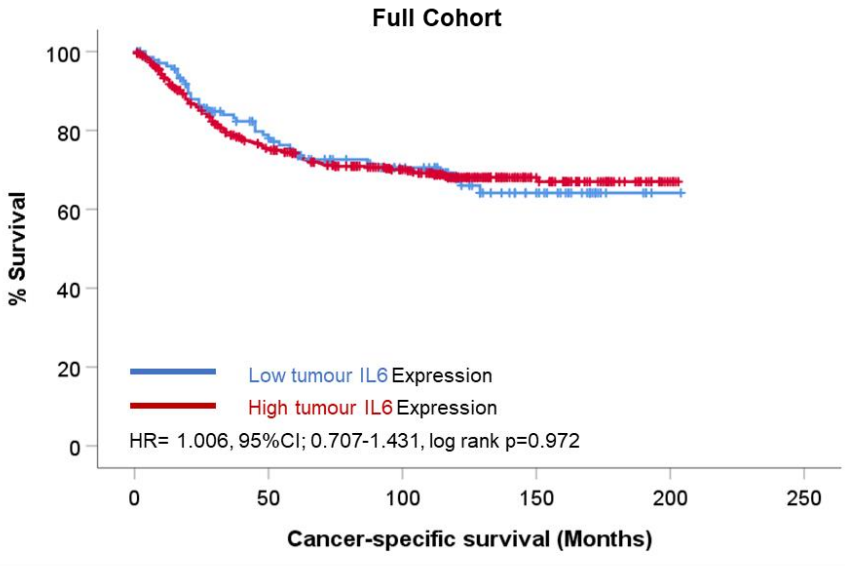


Figure 4.11 Association between tumour IL6 expression and cancer-specific survival.

Kaplan Meier curve showing association between tumour IL6 mRNA expression and CSS in patients from the Glasgow combined cohort. The hazard ratio for IL6 copies per μM^2 within the tumour compartment was 1.006 (95%CI; 0.707-1.431), log rank p=0.972.

When patients were stratified based on GMS classification, there was no association between tumour IL6 mRNA expression and CSS in GMS0, GMS1 or GMS2 (Figure 4.12).

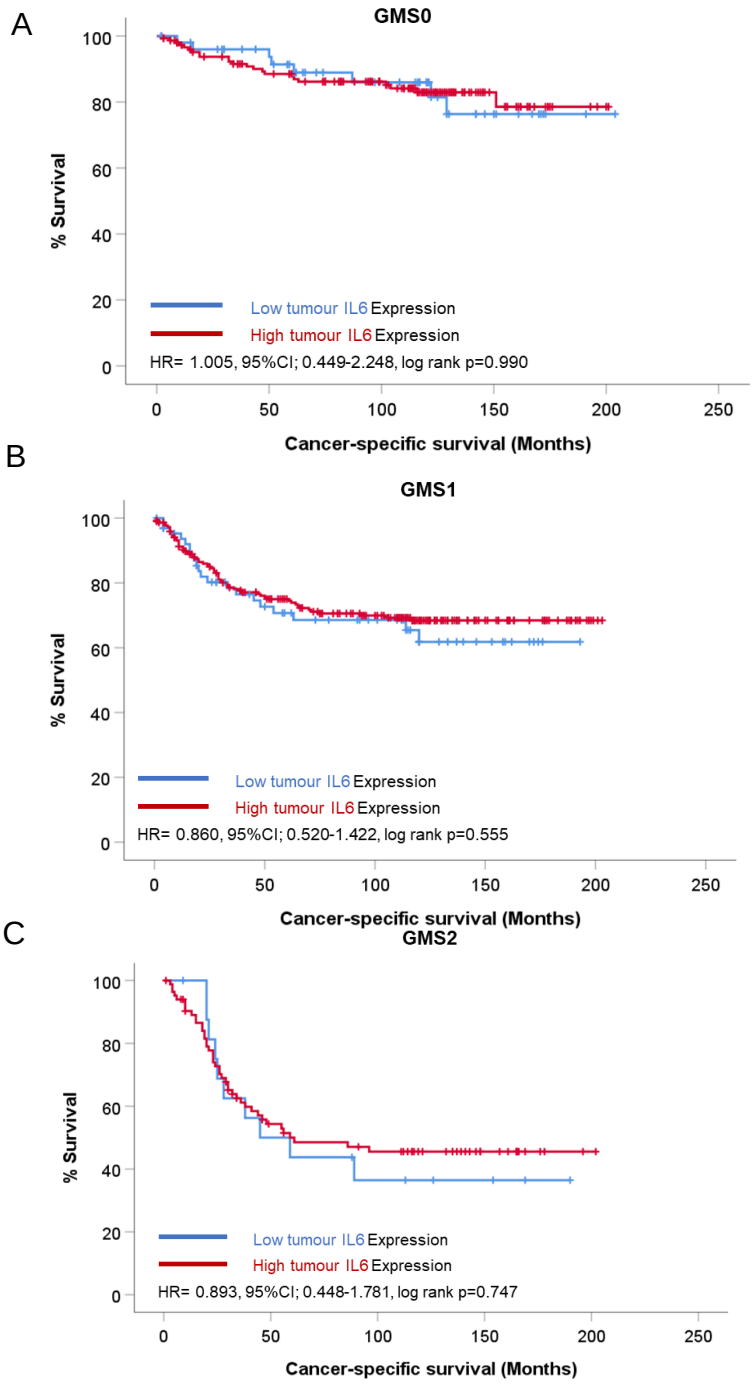


Figure 4.12 Association between tumour IL6 expression and CSS in each GMS classification. *Kaplan Meier survival analysis of tumour IL6 mRNA expression and CSS in patients classified as GMS0 (A), GMS1 (B) and GMS2 (C). The hazard ratio for tumour IL6 expression in GMS0 patients was 1.006 (95%CI; 0.449-2.248), log rank p=0.990 and for patients with GMS1 tumours the hazard ratio was 0.860, (95%CI; 0.520-1.422), log rank p=0.555. In GMS2 cases the hazard ratio for tumoural IL6 expression was 0.893, (95%CI; 0.448-1.781), p=0.747.*

Similarly, when stratified by MMR status, there was no survival difference between patients with high/low tumour IL6 in pMMR or dMMR tumours (Figure 4.13).

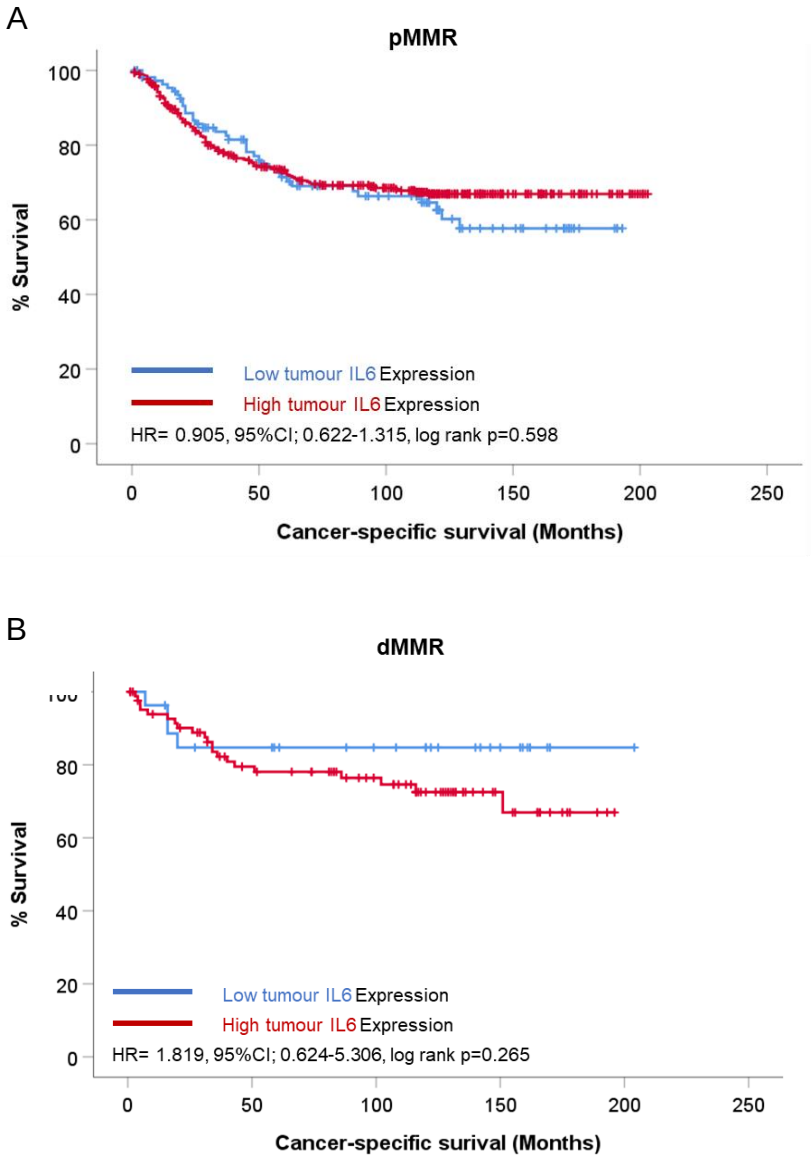


Figure 4.13 Association between tumour IL6 expression and CSS in pMMR and dMMR cases. *Kaplan Meier survival analysis of tumour IL6 mRNA expression and CSS in MMR proficient patients (A) and MMR deficient patients (B). In MMR proficient cases, the hazard ratio for tumour IL6 expression was 0.905, (95% CI; 0.622-1.315), log rank p=0.598 and in MMR deficient cases the hazard ratio was 1.819, (95%CI; 0.624-5.306), log rank p=0.265.*

When patients were subdivided by tumour subsite, no association between CSS and tumour IL6 was observed in right-sided, left-sided, or rectal cases (Figure 4.15). These data suggest IL6 within the stromal compartment is of greater importance than IL6 detected within the tumour-rich areas of CRC.

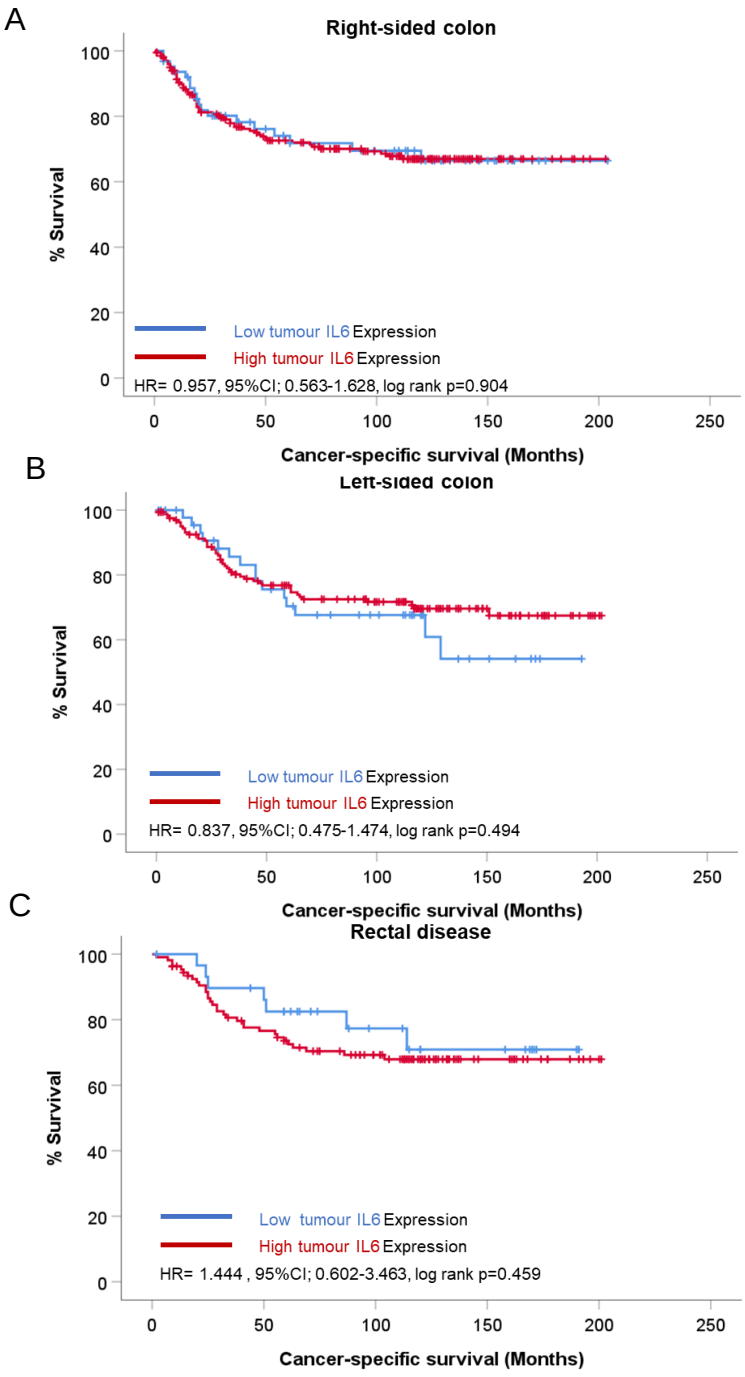


Figure 4.15 Tumour IL6 expression, CSS and tumour subsite. *Kaplan Meier survival analysis showing association between tumour IL6 mRNA expression and CSS in right-sided disease (A), left-sided disease (B) and rectal disease (C). In right sided cases the hazard ratio for expression of tumour IL6 was 0.957, (95%CI; 0.563-1.628), log rank p=0.904 and in left-sided cases the hazard ratio was 0.837, (95%CI; 0.475-1.474), log rank p=0.494. In patients with rectal tumour the hazard ratio for tumoural IL6 expression was 1.444, (95%CI; 0.602-3.463), log rank p=0.459.*

Chi-squared tests were conducted to determine any association between tumour IL6 and clinical features. High expression of IL6 within the tumour epithelium was associated with reduced systemic inflammation assessed via modified Glasgow prognostic score (p=0.040) (Table 4.2).

Clinical Factor	Tumour IL6 mRNA Expression Low (n=191) High (n=191)		p
Age <65 >65	43 (30.3) 99 (69.7)	160 (33.1) 323 (66.9)	0.298
Sex Female Male	64 (45.7) 78 (54.9)	245 (50.7) 238 (49.3)	0.138
T Stage I II III IV	5 (3.5) 18 (12.7) 76 (53.5) 43 (30.3)	21 (4.3) 65 (13.5) 262 (54.2) 135 (28.0)	0.929
N Stage 0 I II	94 (66.7) 34 (24.1) 13 (12.6)	290 (60.2) 131 (27.2) 12.7 (11.5)	0.320
Tumour subsite Right-sided colon Left-sided colon Rectum	62 (3.5) 50 (35.7) 28 (20.0)	203 (42.3) 173 (36.0) 104 (21.7)	0.884
GMS 0 1 2	52 (38.2) 67 (49.3) 17 (12.5)	155 (33.3) 223 (48.0) 87 (18.7)	0.194
Phenotypic Subtype 1 2 3 4	52 (38.2) 44 (32.4) 23 (16.9) 17 (12.5)	156 (33.6) 129 (27.8) 93 (20.0) 86 (18.5)	0.229
mGPS 0 1 2	 51 (44.0) 44 (37.9) 21 (18.1)	 205 (55.6) 96 (26.0) 68 (18.4)	**0.040**
MMR status pMMR dMMR	114 (80.9) 27 (19.1)	395 (82.1) 86 (17.9)	0.408
Tumour differentiation 0 1	122 (85.9) 20 (14.1)	430 (89.0) 53 (11.0)	0.192
Marginal involvement 0 1	185 (96.9) (6.1)	180 (94.2) 11 (5.8)	0.161
Vascular invasion 0 1	92 (64.8) 50 (35.2)	325 (67.3) 158 (32.7)	0.323

Table 4.2 Association between tumour IL6 expression and clinical features. *Chi-squared table of associations for tumour IL6 mRNA expression and clinical prognostic factors including age, sex, T stage, N stage, tumour subsite, GMS, phenotypic subtype, systemic inflammation as measured by modified Glasgow Prognostic Score (mGPS), MMR status, tumour differentiation, marginal involvement and vascular invasion.*

4.2.3 Expression of IL8 in the tumour-stroma and clinical outcome

RNAScope® was performed to detect IL8 within the stroma of the Glasgow combined cohort TMA cores. Due to missing and damaged cores, a total of 439 patients were assessed. After patients who died within 30 days of surgery and those who received neoadjuvant therapy were excluded 382 patients were included in further analysis. obtained for IL8 expression were normalised to PPIB housekeeping gene data. A histogram was plotted to show the distribution pattern of copies, which was positively skewed (Figure 4.15). The mean score was 0.78 with a range of 0-32.35 copies per μM^2. Surminer and maxstat packages in R Studio were utilised to determine the optimal cut point for high and low expression. Scatter and density plots were constructed to visualise the cut-off point of 0.32 (Figure 4.16). This stratified patients into two equal groups with 191 classified as low expression and 191 patients classified as high expression of stromal IL8.

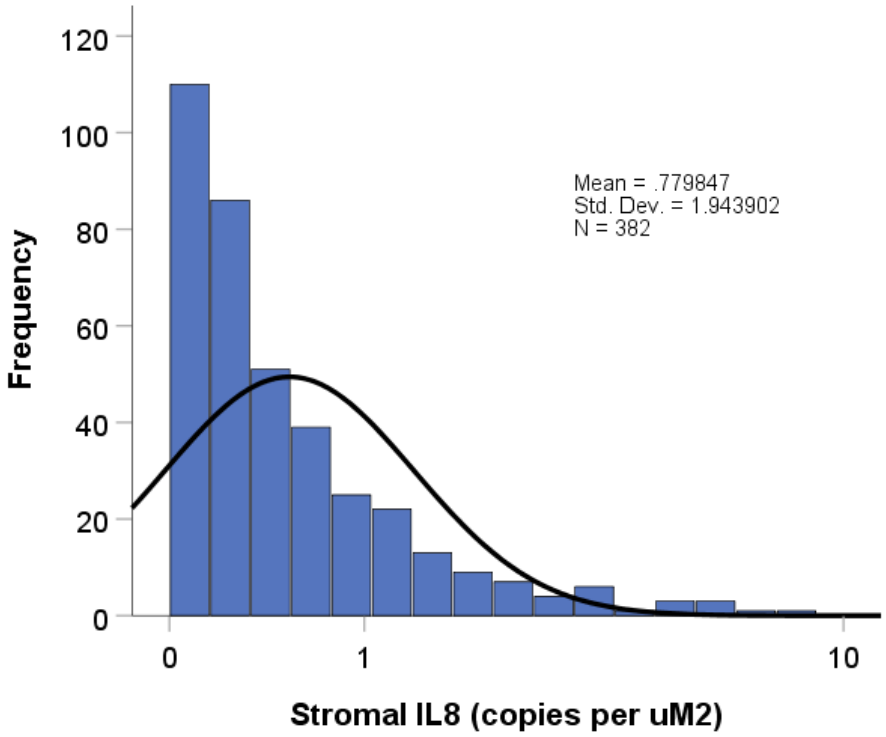


Figure 4.1 Distribution of stromal IL8. *Histogram showing positively skewed distribution of stromal IL8 scores. The mean score for stromal IL8 was 0.78 copies per μM^2 and the scores ranged from 0-32.35 copies per μM^2.*

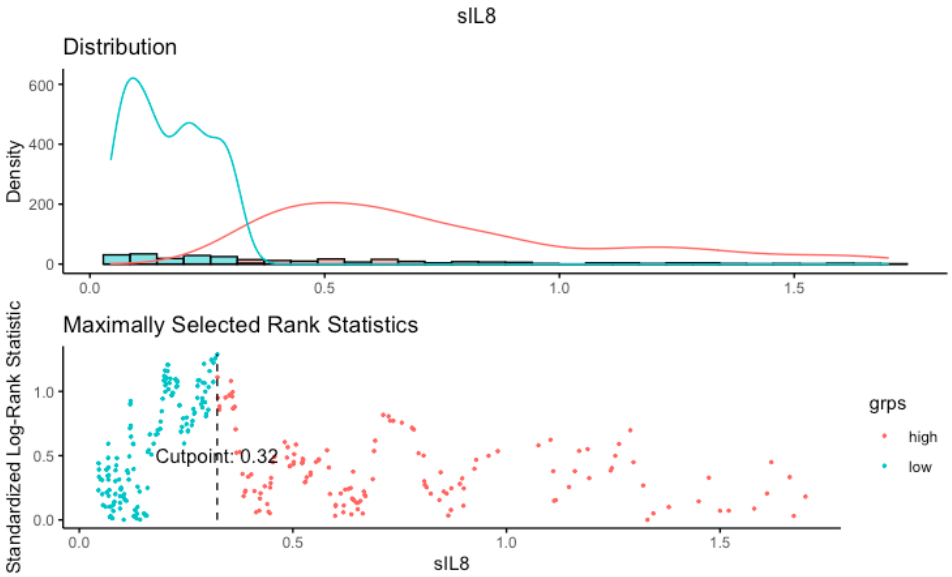


Figure 4.16 Defining optimal cut point for high and low stromal IL8 expression. *Density and scatter plot showing optimal cut off point for high and low expression of stromal IL8 groups based on cancer-specific survival. The optimal cut point determined was 0.32 and therefore patients with a score of ≥0.32 copies per μM^2 were classed as high and patients with a score of ≤0.32 copies per μM^2 were classified as low for expression of stromal IL8.*

Kaplan Meier survival analysis was performed to determine any association between IL8 expression within the stroma and CSS. There was a trend towards high expression of stromal IL8 associating with reduced CSS (p=0.061) (Figure 4.17).

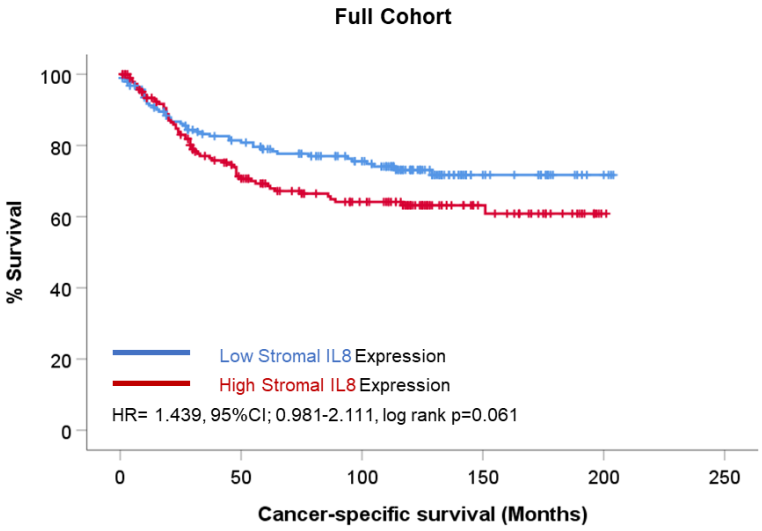


Figure 4.17 Association between stromal IL8 expression and cancer-specific survival.

Kaplan Meier curve showing association between stromal IL8 mRNA expression and CSS in patients from the Glasgow combined cohort. The hazard ratio for stromal IL8 in the full cohort was 1.439, (95%CI; 0.981-2.111), log rank p=0.061.

When patients were stratified by GMS classification there was no association between stromal IL8 mRNA expression and CSS in any group (Figure 4.18).

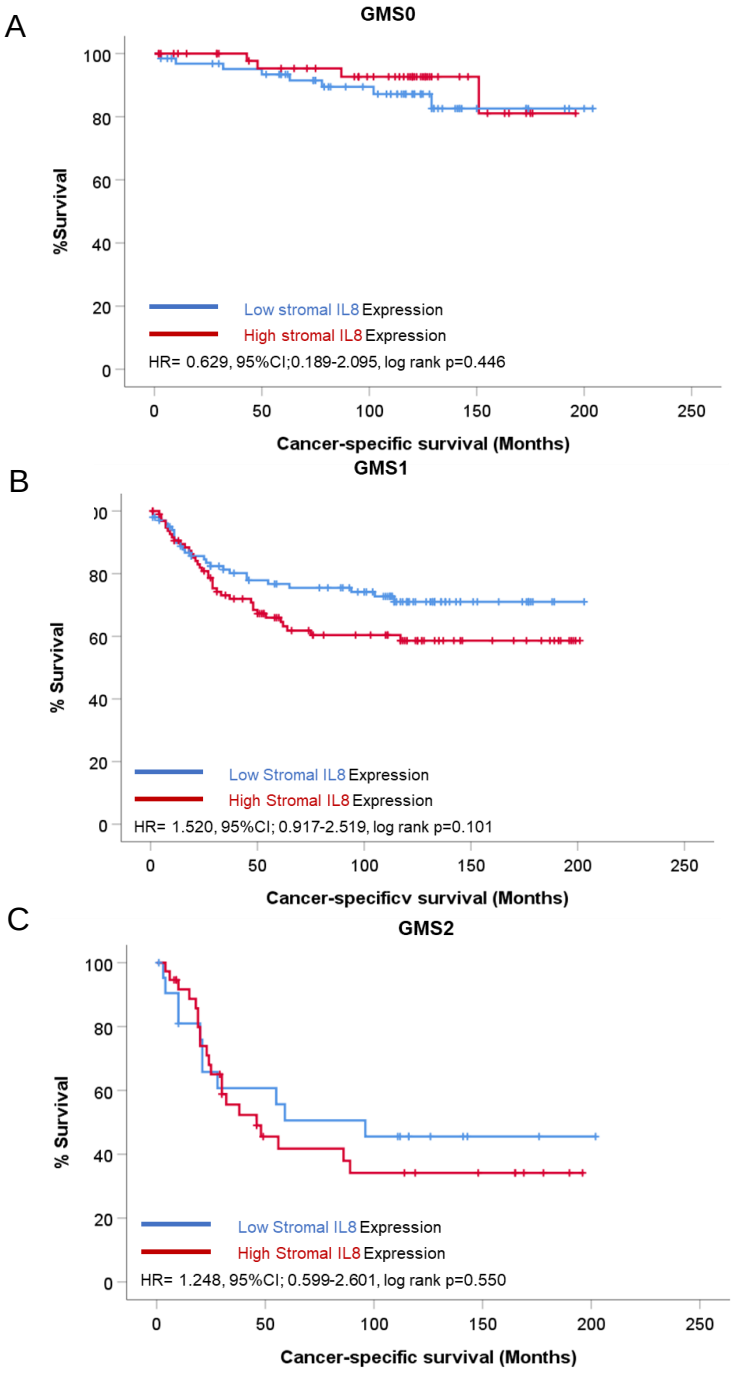


Figure 4.18 Association between stromal IL8 expression and CSS in each GMS classification. *Kaplan Meier survival analysis of stromal IL8 mRNA expression and CSS in patients classified as GMS0 (A), GMS1 (B) and GMS2 (C). The hazard ratio for stromal IL8 in GMS0 cases was 0.629 (95%CI; 0.189-2.095), log rank p=0.446 and in GMS1 cases the hazard ratio was 1.520, (95%CI; 0.917-2.519), log rank p= 0.101. In GMS2 stromal-rich patients the hazard ratio for stromal IL8 was 1.248, (95%CI; 0.599-2.601), log rank p=0.550.*

Similarly, when pMMR and dMMR patient groups were analysed separately there was no significant association between stromal IL8 mRNA expression and CSS in either group (Figure 4.19).

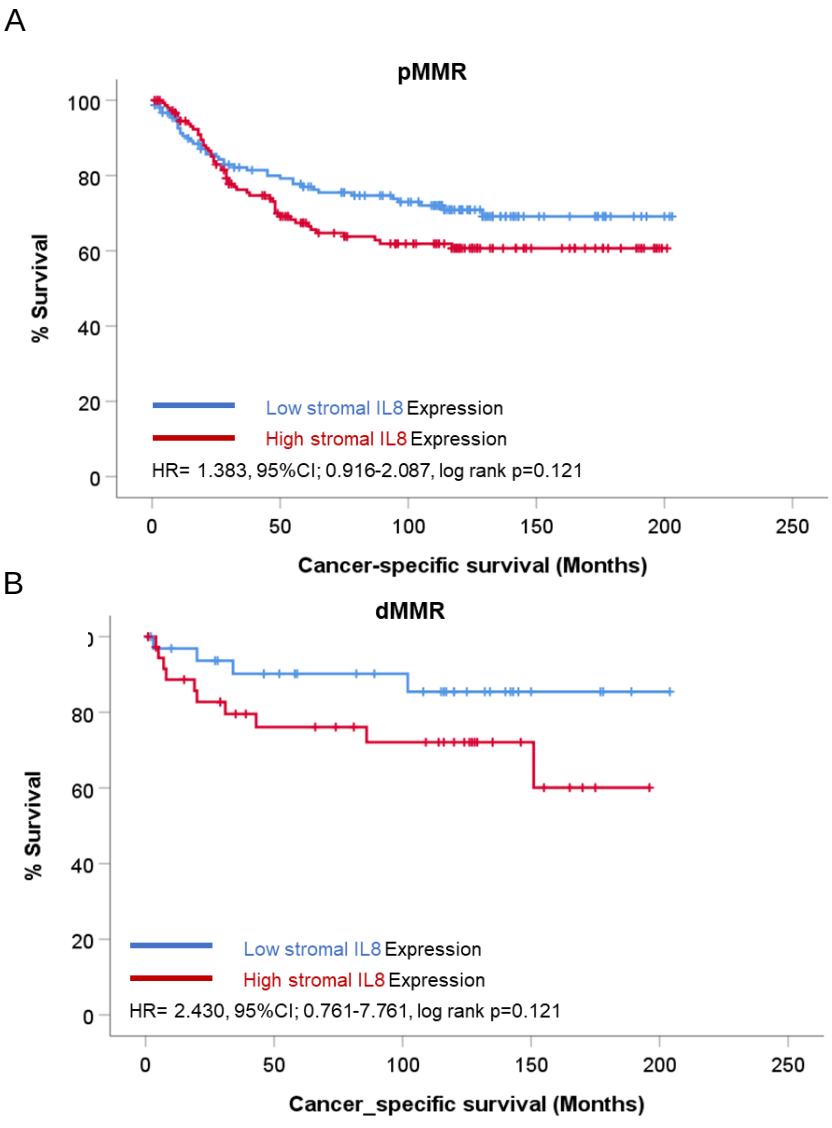


Figure 4.19 Association between stromal IL8 expression and CSS in pMMR and dMMR cases. *Kaplan Meier survival analysis of stromal IL8 mRNA expression and CSS in MMR proficient patients (A) and MMR deficient patients (B). In MMR proficient cases the hazard ratio for stromal IL8 was 1.383, (95%CI; 0.916-2.087), log rank p=0.121 and in MMR deficient cases the hazard ratio was 2.430, (95%CI; 0.761-7.761), log rank p=0.121.*

Kaplan Meier survival analysis was performed to assess stromal IL8 expression within right-sided, left-sided, and rectal cases and no significant association with CSS was observed for any tumour type (Figure 4.20).

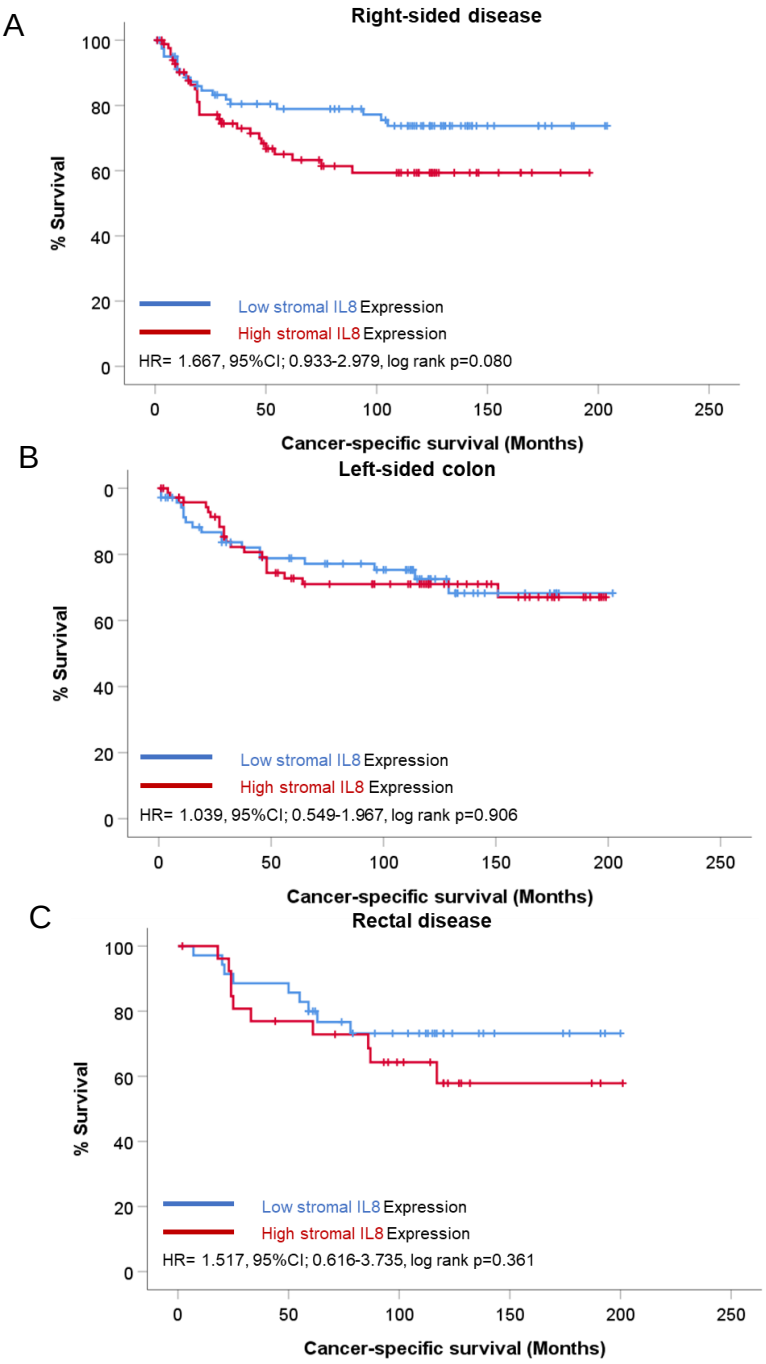


Figure 4.20 Stromal IL8 expression, CSS, and tumour subsite. *Kaplan Meier survival analysis showing association between stromal IL8 mRNA expression and CSS in right-sided disease (A), left-sided disease (B) and rectal disease (C). In right-sided cases, the hazard ratio for IL8 within the stromal compartment of the TME was 1.667, (95%CI; 0.933-2.979), log rank p=0.080 and in left-sided cases the hazard ratio was 1.039, (95%CI 0.549-1.967), log rank p=0.906. In terms of rectal cases the hazard ratio for stromal IL8 was 1.517, (95%CI; 0.616-3.735), log rank= 0.361.*

Chi-squared tests were performed to identify association between clinicopathological characteristics and stromal IL8 mRNA expression (Table 4.3). High expression was enriched in GMS2 (p=0.048) and the stromal phenotypic subtype (p=0.048) (Table 4.3). There was a significant association between stromal IL8 levels and tumour differentiation (p=0.038) (Table 4.3).

Clinical Factor	Stromal IL8 mRNA Expression Low (n=191) High (n=191)		p
Age			
<65	52 (27.2)	52 (27.2)	0.546
>65	139 (72.8)	139 (72.8)	
Sex			
Female	99 (51.8)	97 (50.8)	0.459
Male	92 (48.2)	94 (49.2)	
T Stage			
I	11 (5.8)	4 (2.1)	0.055
II	37 (19.4)	24 (12.6)	
III	94 (49.2)	104 (54.2)	
IV	49 (25.7)	59 (30.9)	
N Stage			
0	130 (68.4)	116 (60.7)	0.154
I	37 (19.4)	53 (27.7)	
II	24 (12.6)	22 (11.5)	
Tumour subsite			
Right-sided colon	84 (44.0)	87 (46.3)	0.579
Left-sided colon	72 (37.7)	74 (39.4)	
Rectum	35 (18.3)	27 (14.4)	
GMS			
0	66 (34.7)	54 (28.3)	**0.048**
1	102 (53.7)	98 (51.3)	
2	22 (11.6)	39 (20.4)	
Phenotypic Subtype			
1			**0.048**
2	67 (35.4)	54 (28.3)	
3	69 (36.5)	59 (30.9)	
4	31 (16.4)	39 (20.4)	
	22 (11.6)	39 (20.4)	
mGPS			
0	75 (55.6)	57 (48.7)	0.546
1	27 (20.0)	28 (23.9)	
2	33 (24.4)	32 (27.4)	
MMR status			
pMMR	154 (81.5)	153 (80.5)	0.458
dMMR	35 (18.5)	37 (19.5)	
Tumour differentiation	163 (85.3)	175 (91.6)	**0.038**
0	28 (14.7)	16 (8.4)	
1			
Marginal involvement			
0	185 (96.9)	180 (94.2)	0.161
1	(6.1)	11 (5.8)	
Vascular invasion			
0	120 (62.8)	128 (67.0)	0.227
1	71 (37.2)	63 (33.0)	

Table 4.3 Association between stromal IL8 expression and clinical features. *Chi-squared table of associations for stromal IL8 mRNA expression and clinical prognostic factors including age, sex, T stage, N stage, tumour subsite, GMS, phenotypic subtype, mGPS, MMR status, tumour differentiation, marginal involvement and vascular invasion.*

4.2.4 Expression of IL8 in the tumour epithelium and clinical outcome

IL8 in the tumour epithelium of the Glasgow combined cohort was probed for by RNAScope® and quantified using Halo® digital pathology platform. Due to missing and damaged cores 439 patients were assessed for tumour IL8 and after exclusion criteria were applied 383 patients were left for further analysis. Scores for tumour IL8 were normalised to PPIB housekeeping gene scores for the same area. A histogram plot was constructed to visualise the distribution pattern of tumour IL8 levels, which appeared positively skewed (Figure 4.21). The mean score for tumoural IL8 was 0.57 with a range of 0-13.80 copies per μM^2. Survminer and maxtast R studio packages were utilised to determine the optimum cut point for high and low expression. Scatter and density plots were constructed to visualise the output with an optimal cut point of 0.65 copies per μM^2 devised (Figure 4.22). This resulted in 289 patients classified as low for tumour IL8 copies, and 92 as high for tumour IL8.

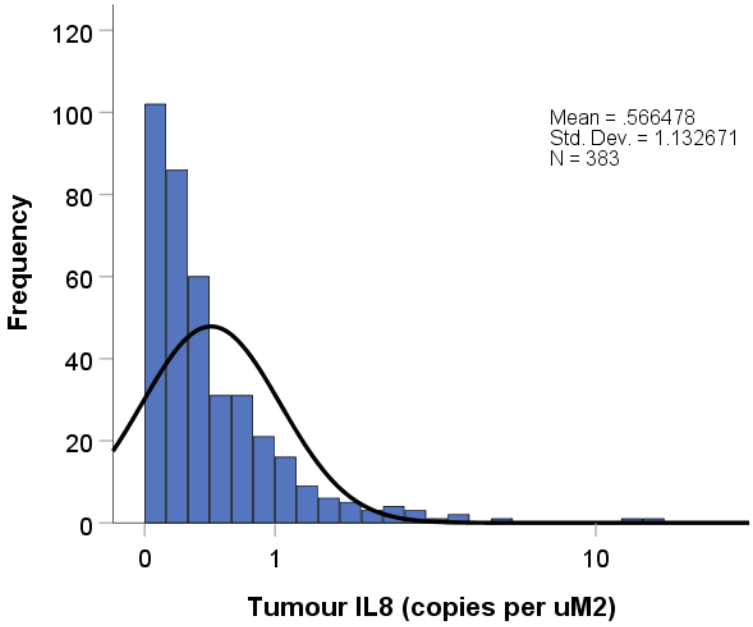


Figure 4.21 Distribution of tumour IL8. *Histogram showing positively skewed distribution pattern of tumour IL8 scores. The mean score for tumour IL8 was 0.57 copies per μM^2 and the scores ranged from 0-13.80 copies per μM^2.*

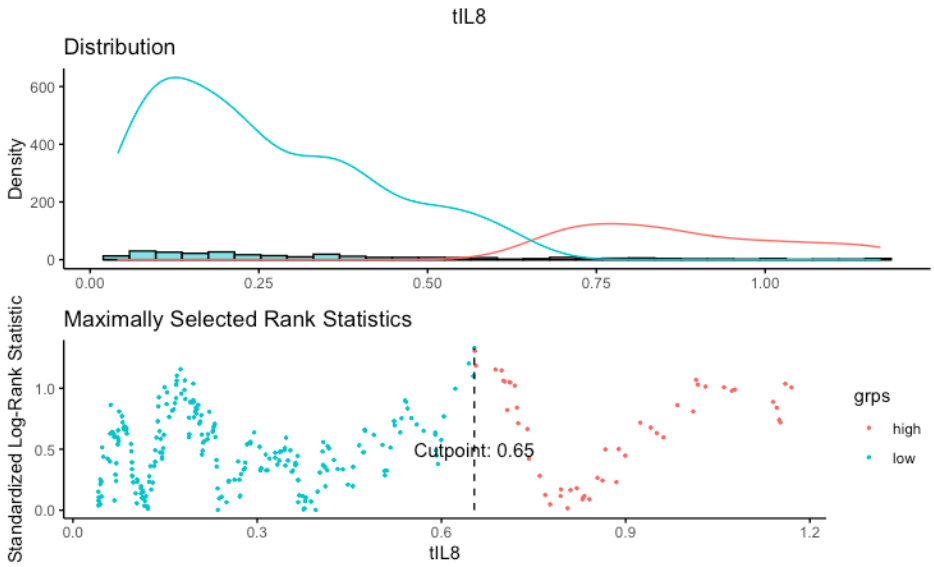


Figure 4.22 Defining optimal cut point for high and low tumour IL8 expression. *Density and scatter plot showing optimal cut off point for high and low expression of tumour IL8 groups based on cancer-specific survival. The optimalt cut off point determined was 0.65 copies per μM^2, therefore patients with a score of ≤0.65 IL8 copies per μM^2 were classified as low and patients with a score of ≥0.65 copies per μM^2 were classed as high for tumour IL8 expression.*

Kaplan Meier survival analysis was performed to investigate association between tumour IL8 and CSS in the full Glasgow combined cohort. There was no difference in survival time between patients with high or low expression of tumour IL8 (Figure 4.23).

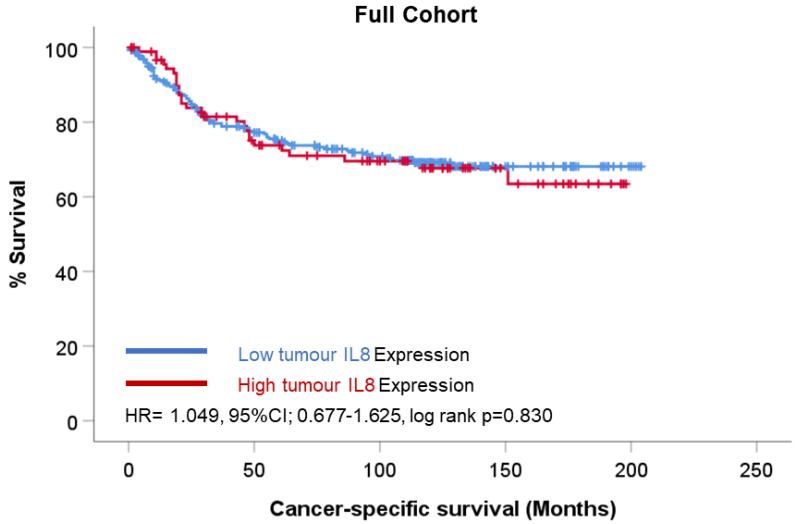


Figure 4.23 Association between tumour IL8 expression and cancer-specific survival. *Kaplan Meier curve showing association between tumour IL8 mRNA expression and CSS in patients from the Glasgow combined cohort. The hazard ratio for tumour IL8 expression in the full cohort was 1.049, (95%CI; 0.677-1.625), log rank p=0.830.*

When survival analysis was performed on patients from each histological GMS classification, no association between tumour IL8 and CSS was observed in any group (Figure 4.23).

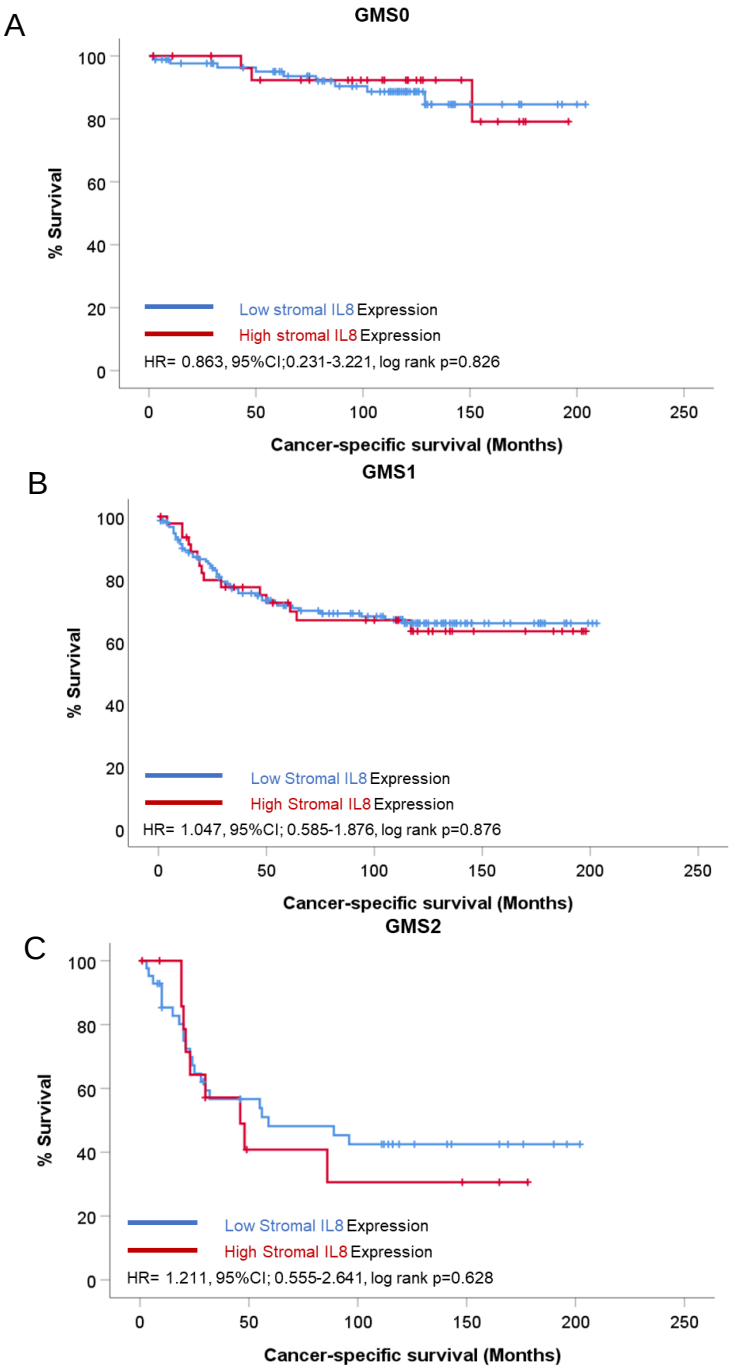


Figure 4.24 Association between tumour IL8 expression and CSS in each GMS classification. *Kaplan Meier survival analysis of tumour IL8 mRNA expression and CSS in patients classified as GMS0 (A), GMS1 (B) and GMS2 (C). In GMS0 cases the hazard ratio for tumour IL8 expression was 0.863, (95%CI; 0.231-3.221), log rank p=0.826 and in GMS1 cases the hazard ratio was 1.047, 95%CI; 0.585-1.867), log rank p=0.876. In gMS2 stromal-rich cases the hazard ratio for tumour IL8 was 1.211, (95%CI; 0.555-2.641), log rank p=0.628.*

When patients were stratified by MMR status, no association between tumour IL8 and CSS was seen in pMMR or dMMR tumours (Figure 4.25).

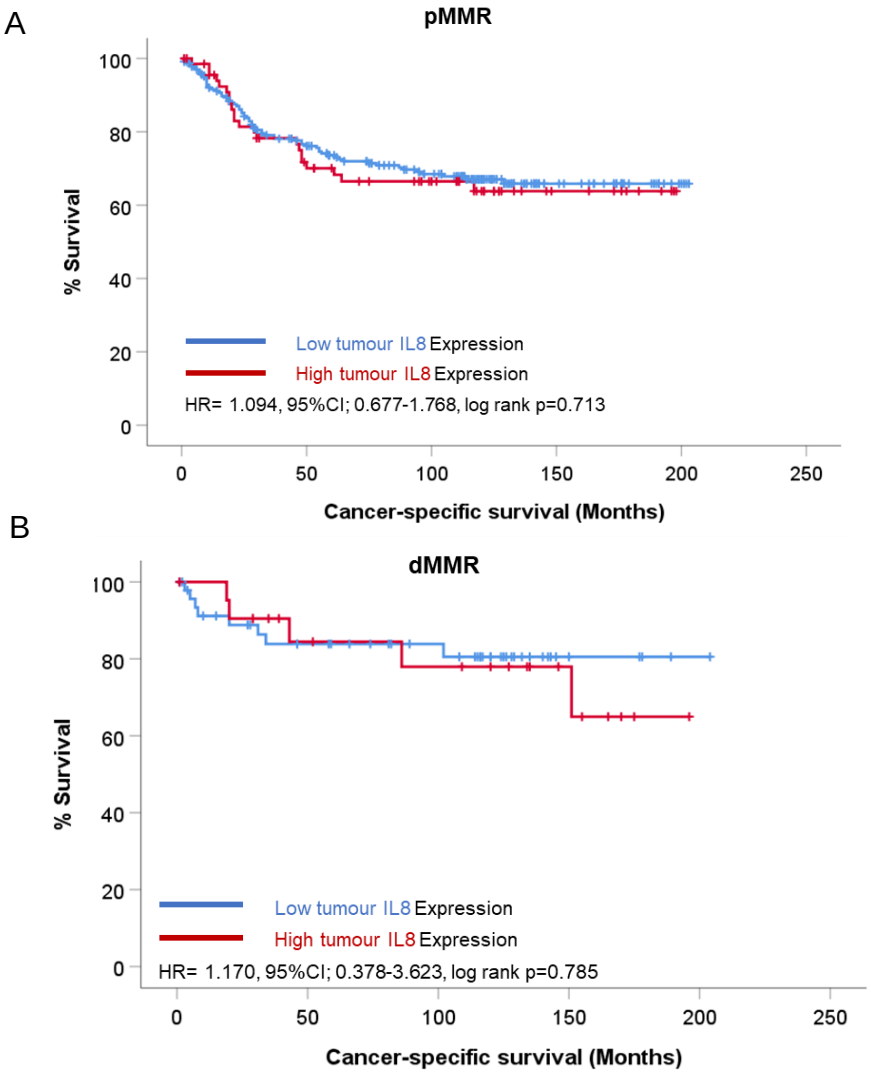


Figure 4.25 Association between tumour IL8 expression and CSS in pMMR and dMMR cases. *Kaplan Meier survival analysis of tumourIL8 mRNA expression and CSS in MMR proficient patients (A) and MMR deficient patients (B). In MMR proficient cases the hazard ratio for tumour IL8 expression was 1.094, (9%CI; 0.677-1.768), log rank p=0.713 and in MMR deficient cases the hazard ratio was 1.170, (95%CI; 0.378-3.623), log rank p=0.785.*

There was no significant association between tumour IL8 levels and CSS in right-sided, left-sided, or rectal cases (Figure 4.26).

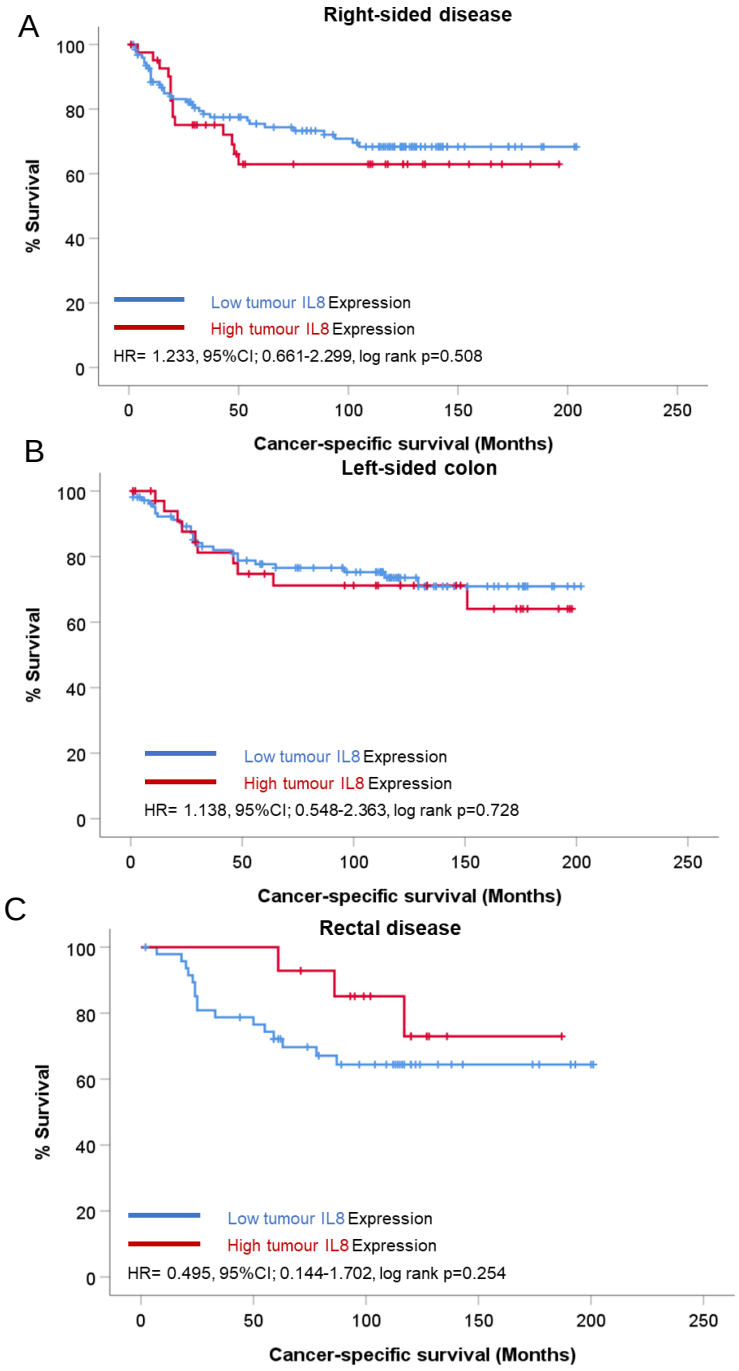


Figure 4.26 Tumour IL8 expression, CSS and tumour subsite. *Kaplan Meier survival analysis showing association between tumour IL8 mRNA expression and CSS in right-sided disease (A), left-sided disease (B) and rectal disease (C). In patients with right-sided tumours the hazard ratio for tumour IL8 was 1.233, (95%CI; 0.661-2.299), log rank =0.508 and in left-sided cases the hazard ratio was 1.138, (95%CI; 0.548-2.363), log rank =0.728. In rectal cases, the hazard ratio was 0.495, (95%CI; 0.144-1.702), log rank p=0.254.*

Chi-squared tests were performed to determine if levels of tumour IL8 were associated with clinicopathological characteristics. No statistically significant associations were identified (Table 4.4).

Clinical Factor	Tumour IL8 mRNA Expression Low (n=289)	High (n=92)	p
Age <65 >65	81 (28.0) 208 (72.0)	22 (23.9) 70 (76.1)	0.264
Sex Female Male	152 (52.6) 137 (47.4)	44 (47.8) 48 (52.2)	0.249
T Stage I II III IV	11 (3.8) 49 (17.0) 152 (52.6) 77 (26.6)	4 (4.3) 12 (13.0) 45 (48.9) 31 (33.7)	0.550
N Stage 0 I II	190 (65.7) 64 (22.1) 35 (12.1)	56 (60.9) 26 (28.3) 10 (10.9)	0.493
Tumour subsite Right-sided colon Left-sided colon Rectum	130 (45.5) 108 (37.8) 48 (16.8)	42 (45.7) 36 (39.1) 14 (15.1)	0.932
GMS 0 1 2	91 (31.6) 153 (53.1) 44 (15.3)	30 (32.6) 46 (50.0) 16 (17.4)	0.841
Phenotypic Subtype 1 2 3 4	92 (32.1) 102 (35.5) 49 (17.1) 44 (15.3)	30 (32.6) 26 (28.3) 20 (21.7) 16 (17.4)	0.546
mGPS 0 1 2	101 (51.0) 44 (22.2) 53 (26.8)	31 (58.5) 12 (22.6) 10 (18.9)	0.461
MMR status pMMR dMMR	238 (82.9) 49 (17.1)	69 (75.8) 22 (24.2)	0.089
Tumour differentiation 0 1	254 (87.9) 15 (5.2)	90 (97.8) 2 (2.2)	0.178
Marginal involvement 0 1	274 (94.8) 3 (5.8)	303 (95.6) 14 (4.4)	0.439
Vascular invasion 0 1	186 (64.4) 103 (35.6)	63 (68.5) 29 (31.5)	0.277

Table 4.4 Association between tumour IL8 expression and clinical features. *Chi-squared table of associations for stromal IL8 mRNA expression and clinical prognostic factors including age, sex, T stage, N stage, tumour subsite, GMS, phenotypic subtype, mGPS, MMR status, tumour differentiation, marginal involvement and vascular invasion.*

Stromal IL6 and IL8 scores were combined to form a new score of low for both, high for one, or high for both markers. Kaplan Meier survival analysis showed no significant association for combined stromal IL6 and stromal IL8 score (p=0.184) (Figure 4.27). This could be due to the number of patients in the high for both markers category being too few.

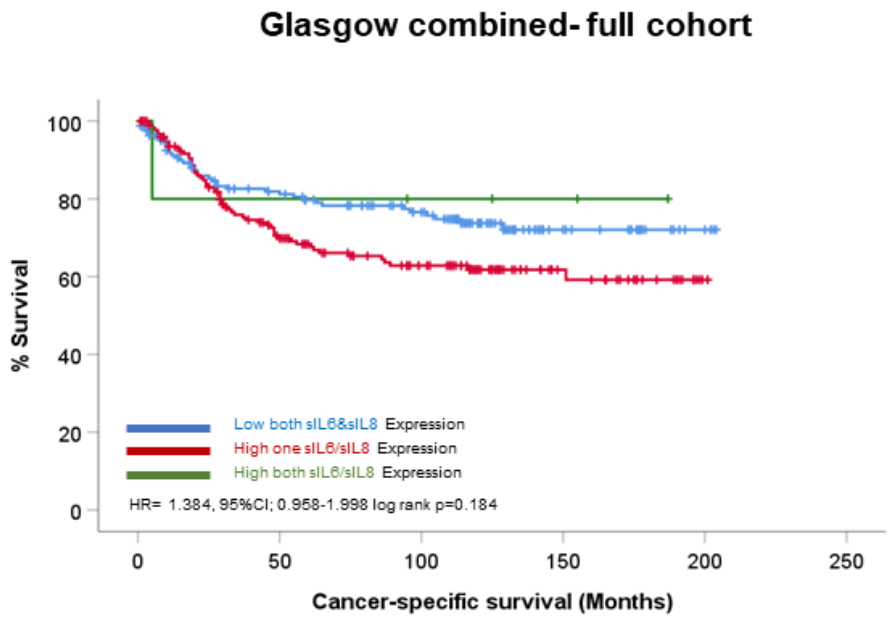


Figure 4.27 Combined stromal IL6 and IL8 expression and CSS. *Kaplan Meier survival analysis showing association between a combined score of stromal IL6 and IL8 mRNA expression and CSS in the full Glasgow combined cohort. The hazard ratio for a combined score of stromal IL6 and IL8 was 1.384, (95%CI; 0.958-1.998), log rank p= 0.184.*

4.2.5 Expression of IL6R in the tumour epithelium and clinical outcome

Cytoplasmic expression of IL6R in tumour cells was assessed by a single observer (KP) in patients from the Glasgow combined array. Representative images of weak, moderate, and strong staining and shown in Figure 4.28. Weighted histoscores ranged from 0 to 200 with a median of 100. True positive control for IL6R was liver tissue and a true negative was prostate tissue (Figure 4.28). Positive control colorectal tissue identified during antibody optimisation and a negative control with no primary antibody representative images are shown in Figure 4.28. Antibody specificity was performed via western blotting, with a single band at 51KDa observed in 3 replicates of HeLa lysates and no protein detected in LNCaP prostate lysates, with tubulin present in all 6 samples at 52KDa (Figure 4.28). Positive staining for IL6R was noted in stromal and immune cell compartments of the TME, with

representative images of low and high expression for each shown in Figure 4.28, however it was outwith the scope of this study to quantify expression in these areas, and analysis was focused on the tumour areas.

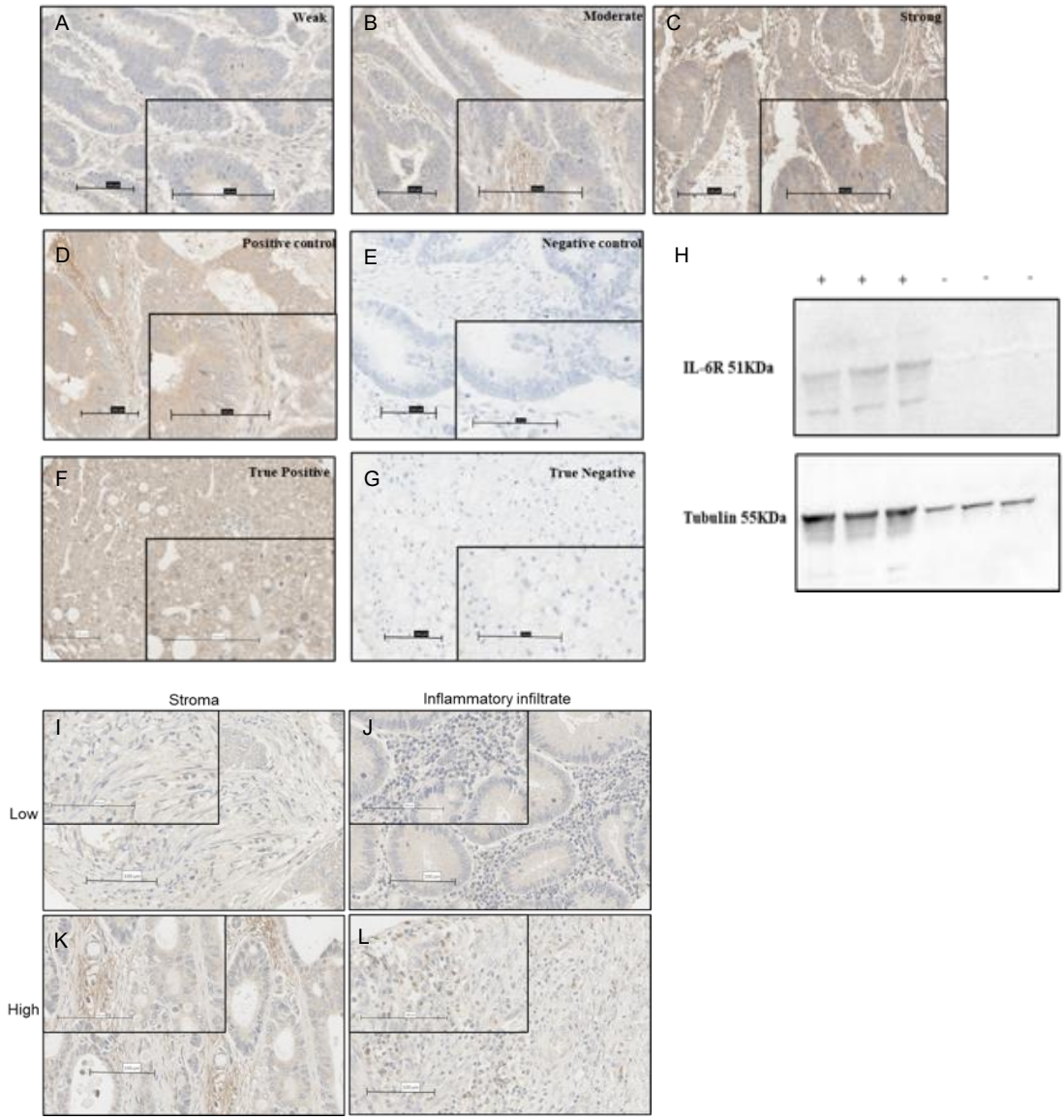


Figure 4.28 Representative images and antibody specificity for IL6R. *Images showing representative images of weak (A), moderate (B) and strong (C) cytoplasmic staining of IL6R within tumour cells. Representative images of positive (D) and negative (E) colorectal tissue, and true positive (F) and true negative tissue (G). Image of western blot showing single band in triplicate for HeLa cell lysates (20μL loaded per well) probed for IL6R at the correct molecular weight of 51KDa and no visible band in triplicate for true negative LNCaP lysates (20μL loaded per well) at 51KDa, but visible bands at 55KDa across all samples when the membrane was probed for tubulin (H). Representative images of low expression of IL6R within the stromal (I) and immune (J) compartments*

of the TME, respectively. Representative images of high expression of IL6R within the stromal (K) and immune (L) compartments of the TME.

Due to missing and damaged cores scores were obtained for 404 patients from the 1030 patient cohort. Exclusion criteria were applied (30-day mortality and neoadjuvant therapy), leaving 369 patients included in downstream analysis. Data were relatively normally distributed as shown in histogram plot (Figure 4.29). Scoring was validated using QuPath digital image analysis software on 10% of TMA cores, with a correlation of 0.720 and an intra-class correlation coefficient of 0.825 obtained. A scatterplot (Figure 4.30) and a Bland-Altman (Figure 4.31) were constructed to visualise the correlation and bias between scores, respectively. Cut offs for high and low expression were determined using survminer and maxstat packages in R studio via log rank statistics based on cancer-specific survival. The optimal cut point determined by this method was 44.9, as shown in density and scatter plot (Figure 4.32). This resulted in 52 patients classified as low for IL6R expression and 317 as high expressers.

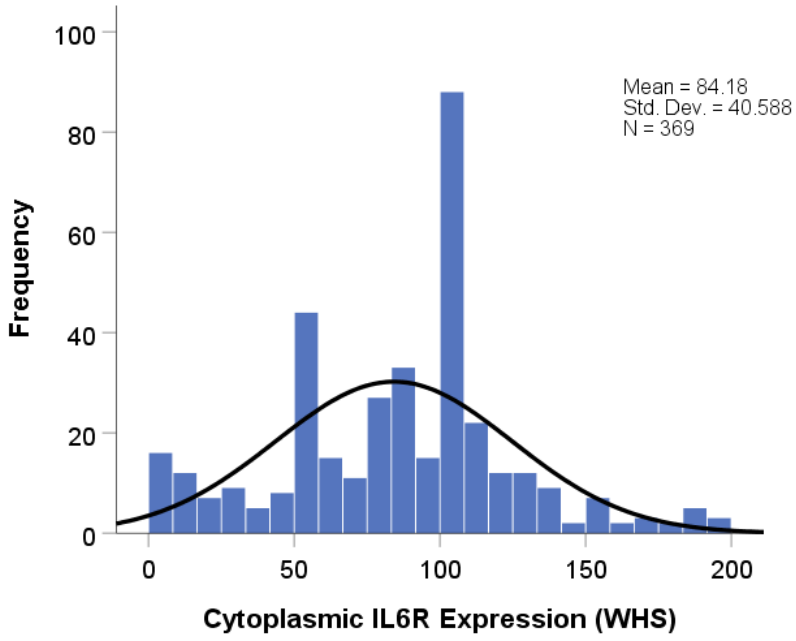


Figure 4.29 Distribution of weighted histoscores for IL6R. *Histogram showing the range of scores obtained for cytoplasmic tumour IL6R expression and normal distribution pattern of the data (n=369). The mean score for IL6R was 84.18 and the scores ranged from 0-200.*

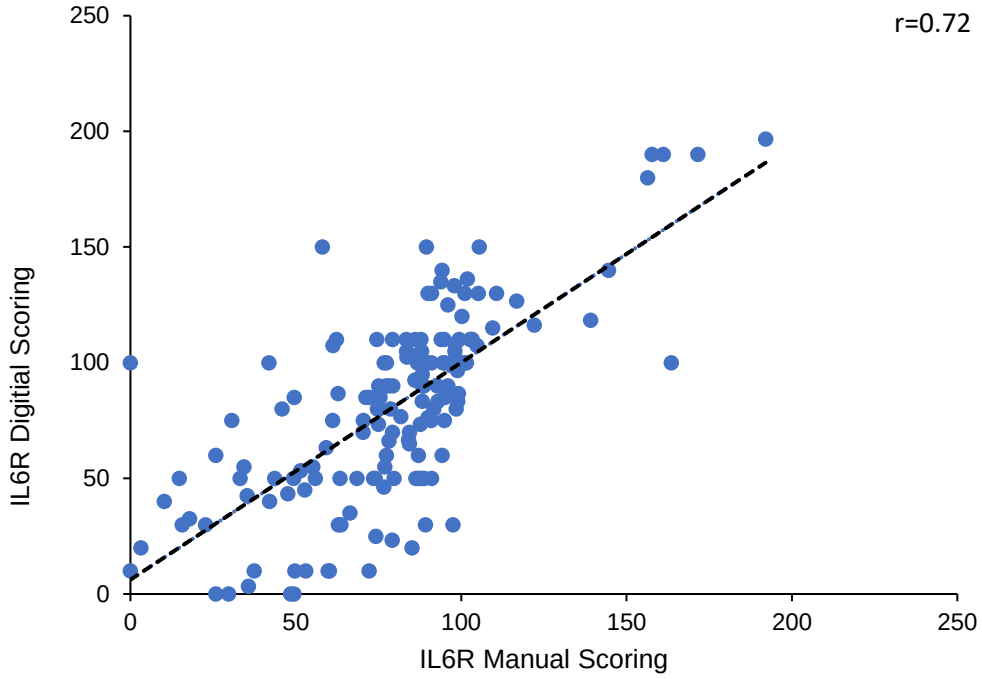


Figure 4.30 Correlation between manual and digital weighted histoscoring of IL6R.
Scatter plot showing correlation between tumour cytoplasm weighted histoscores for IL6R for 10% of co-scored cases. The correlation coefficient was 0.72.

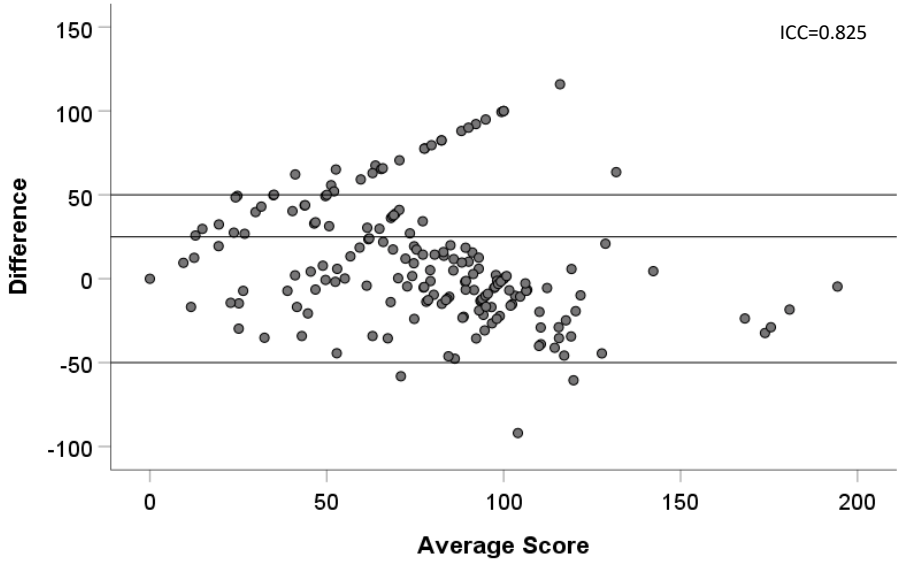


Figure 4.31 Validation of IL6R manual scoring. *Bland Altmann plot showing difference between manual and digital QuPath scores for IL6R in tumour cytoplasm for the 10% of cases co-scored.*

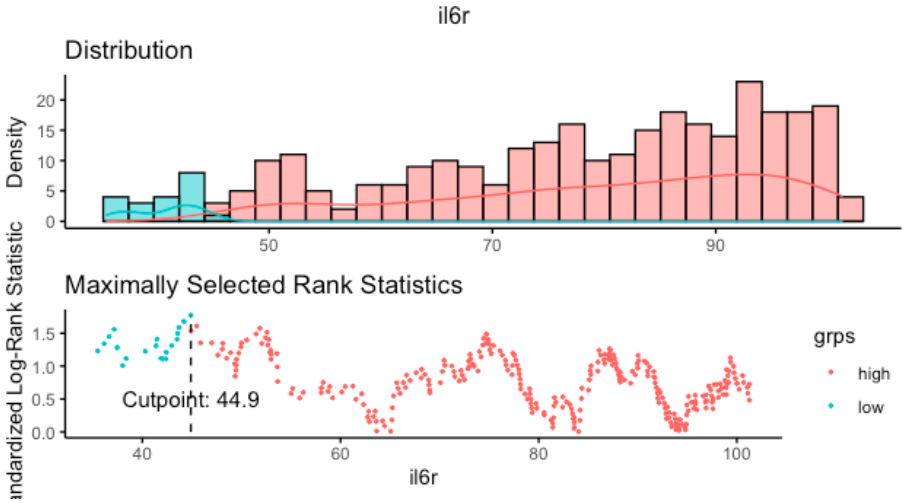
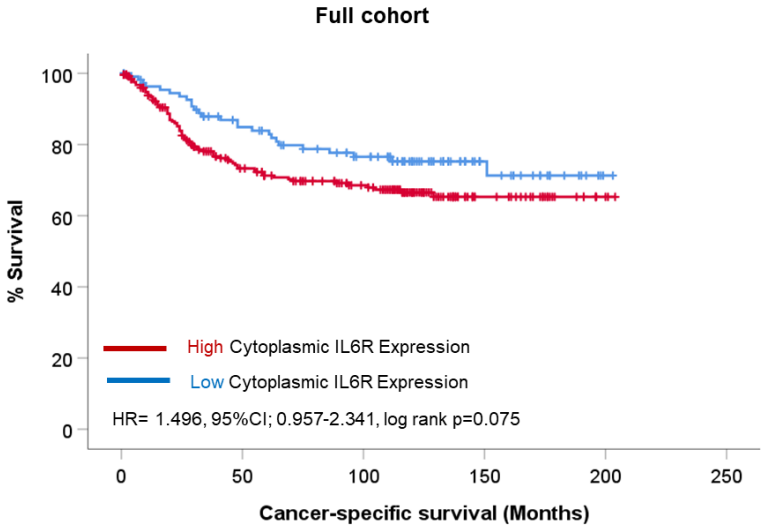


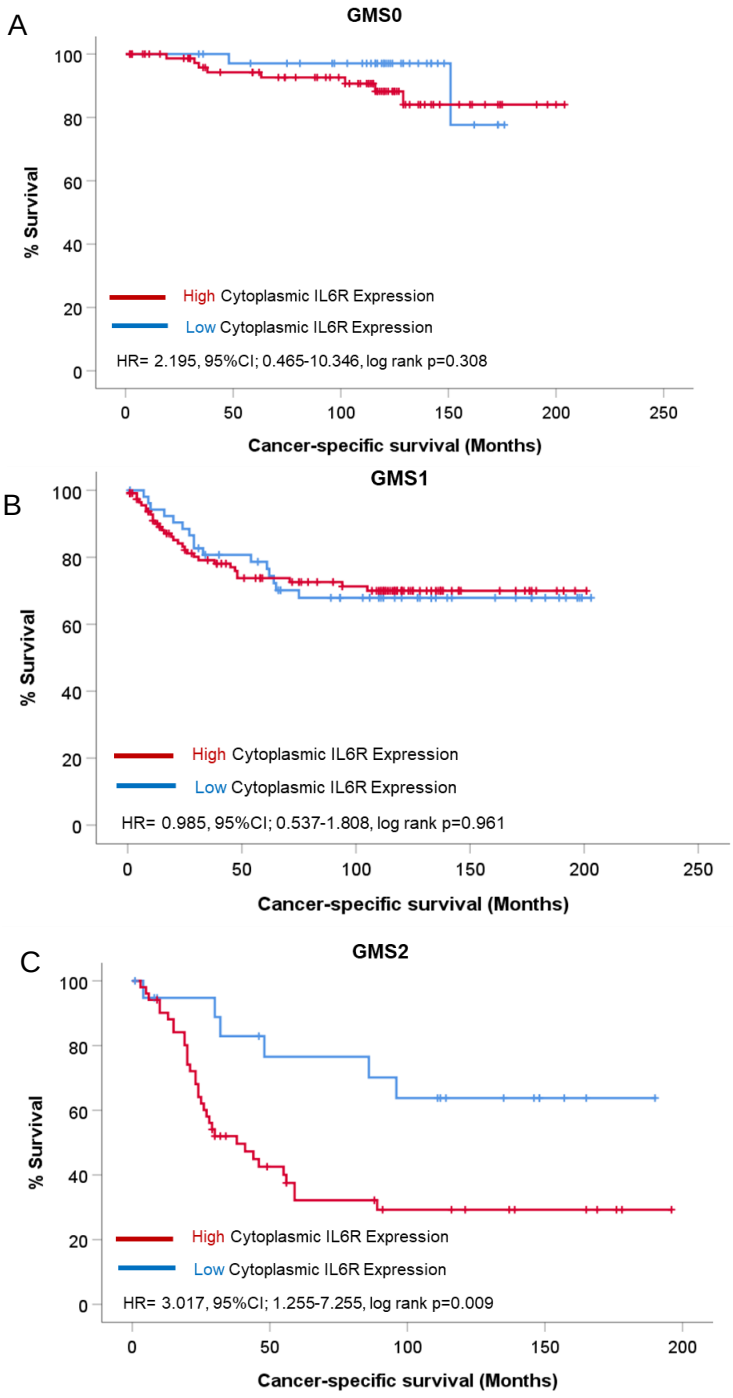
Figure 4.32 Defining cut off point for IL6R expression high and low groups. *Density plot and scatter plot for visualisation of optimal cut off point for high and low expression of tumour cytoplasmic IL6R. The optimal cut point determined was 44.9, therefore patients with a weighted histoscore of ≥44.9 were classified as high and patients with a weighted histoscore of ≤44.9were classified as low for cytoplasmic IL6R expression.*

Kaplan Meier survival analysis was performed to determine association of IL6R expression with cancer-specific survival. Overall, there was no significant association between IL6R expression and CSS, however there was a trend towards high expression associating with poorer outcome (HR= 1.496, 95%CI; 0.957-2.341, log rank p=0.075) (Figure 4.33).



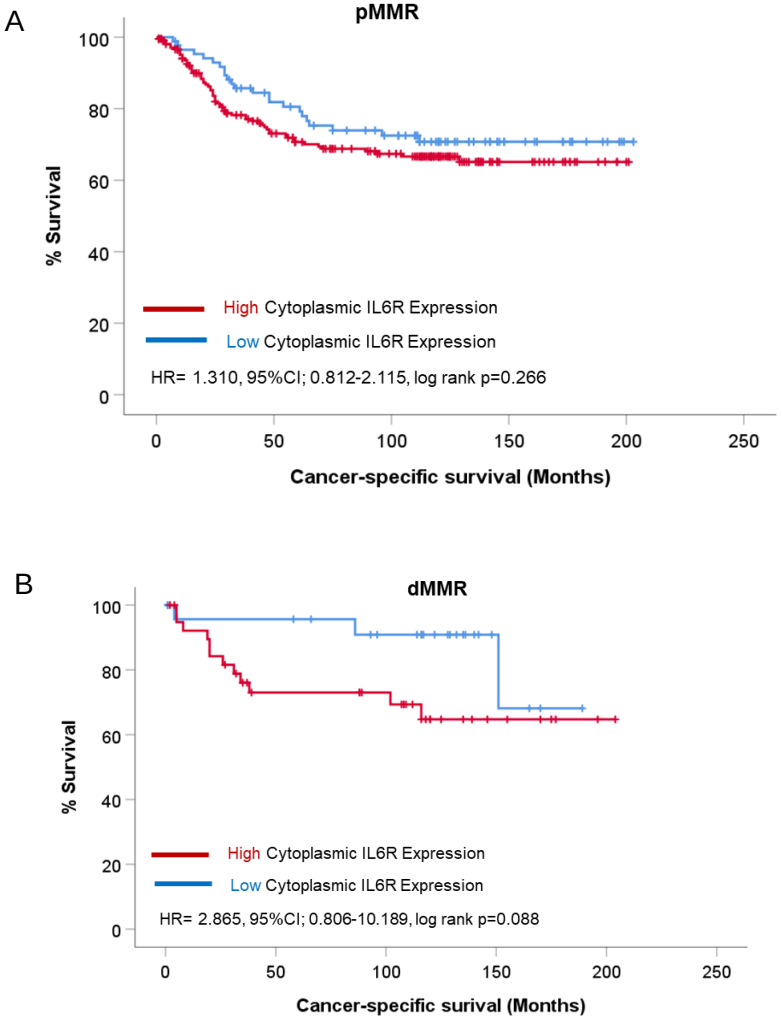
4.33 Association between IL6R expression and Cancer-specific survival. *Kaplan Meier curve showing association between IL6R tumour cytoplasmic expression and CSS in the full Glasgow combined cohort. The hazard ratio for IL6R expression in the full cohort was 1.496, (95%CI; 0.957-2.341), log rank p=0.075.*

When stratified by GMS, IL6R expression in patients classified as GMS0 and GMS1 was not associated with CSS (Figure 4.34). However, there was a profound difference in patients with stromally-dense GMS2 tumours. High IL6R expression in these patients was associated with significantly reduced CSS (HR= 3.017, 95%CI; 1.255-7.255, log rank p=0.009) (Figure 4.34). Patients with low IL6R expression observed a median survival time of 139.701 months compared to 78.987 months for patients with high expression.



4.34 Association between IL6R expression and Cancer-specific survival in each GMS classification. *Kaplan Meier curve showing association between IL6R tumour cytoplasmic expression and CSS in GMS0 patients (A), GMS1 patients (B) and GMS2 patients (C). In patients with GMS0 classified tumours the hazard ratio for IL6R expression was 2.195, (95%CI; 0.465-10.346), log rank p=0.308 and for GMS1 cases the hazard ratio was 0.985, (95%CI; 0.537-1.808), log rank p=0.961. In GMS2 stromal-rich cases the hazard ratio for IL6R expression was 3.017, 95%CI; 1.255-7.255), log rank p=0.009.*

In terms of MMR status there was no significant association between CSS and cytoplasmic IL6R expression in patients with proficient or deficient MMR tumours (Figure 4.345.



4.35 Association between IL6R expression and Cancer-specific survival stratified by MMR status. *Kaplan Meier curve showing association between IL6R tumour cytoplasmic expression and CSS in pMMR patients (A) and dMMR patients (B). In patients with MMR proficient disease the hazard ratio for IL6R expression was 1.310, 95%CI; 0.812-2.115), log rank p=0.266 and MMR deficient cases the hazard ratio was 2.86, (95%CI; 0.806-10.189), log rank p=0.088.*

Patients with right-sided disease exhibited worse prognosis if their tumour cells were highly expressing IL6R (HR= 2.225, 95%CI; 1.033-4.792, log rank p= 0.035) (Figure 4.35). In right-sided cases, patients with low expression of IL6R had a mean survival time of 171.843 months compared to 140.110 months for patients with high expression. IL6R levels were not associated with CSS in patients with left-sided disease or with rectal tumours (Figure 4.36).

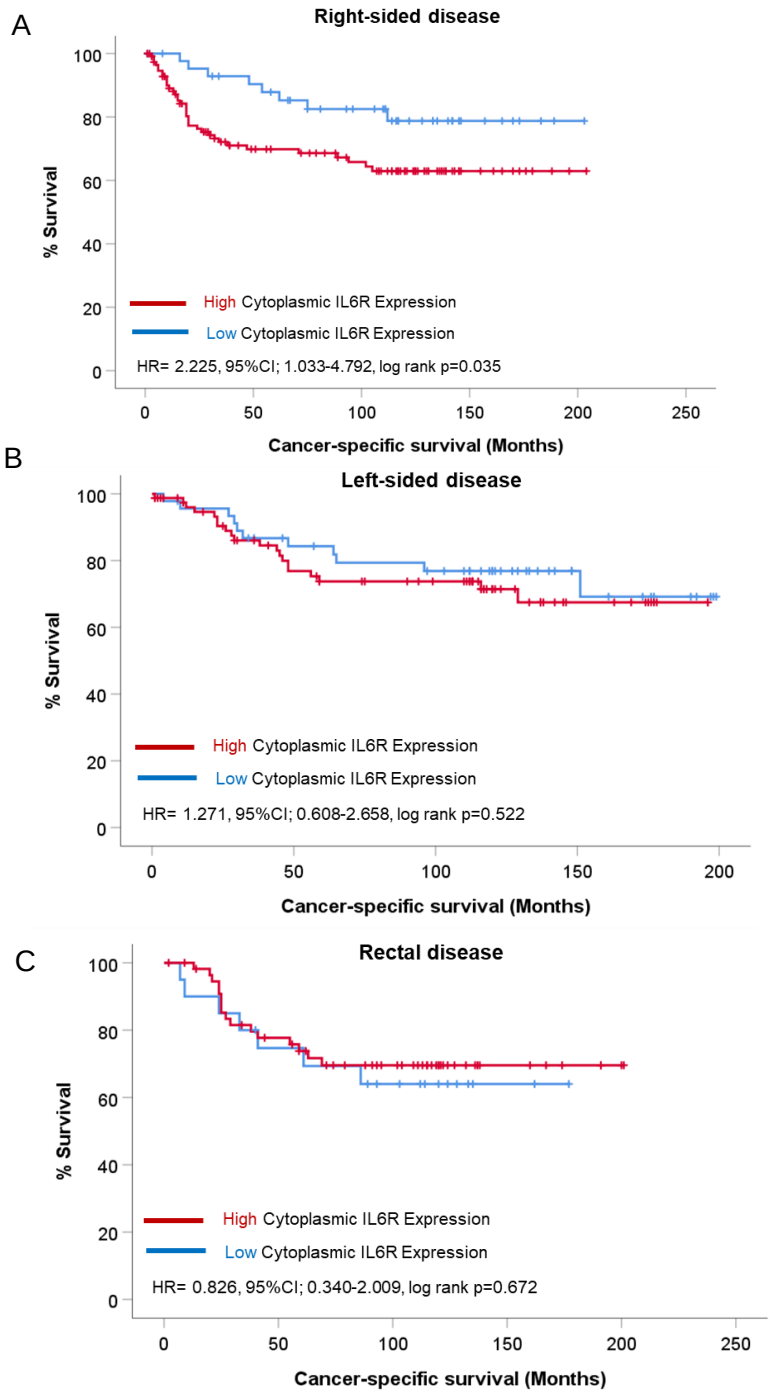


Figure 4.36 Association between IL6R expression and Cancer-specific survival stratified by disease subsite. *Kaplan Meier curve showing association between IL6R tumour cytoplasmic expression and CSS in right-sided cancer (A), left-sided cancer (B) and rectal cancer patients (C). In patients with right-sided tumour the hazard ratio for IL6R expression was 2.225, (95%CI; 1.033-4.792), log rank p=0.03 and in left-sided cases the hazard ratio was 1.271, (95%CI; 0.608-2.009), log rank =0.522. In patients with rectal tumours, the hazard ratio for IL6R expression was0.826, (95%CI;0.340-2.009,), log rank p=0.672.*

Next, IL6R expression was assessed for association with clinicopathological factors via chi-squared analysis. Low expression of IL6R was associated with female sex (p=0.001). High expression was more common in rectal cases and right-sidedness than left-sided tumours (p=0.015) and more commonly observed in MMR proficient tumours (p=0.023) (Table 4.5).

Clinical Factor	IL6R Expression		p
	Low (n=425)	High (n=197)	
Age	18 (34.6)	109 (34.4)	
<65	34 (65.4)	208 (65.6)	0.545
>65			
Sex			
Female	36 (69.2)	46 (46.1)	**0.001**
Male	16 (30.8)	171 (53.9)	
T Stage			
I	2 (3.8)	13 (4.1)	0.422
II	4 (7.7)	47 (14.8)	
III	29 (55.8)	177 (55.8)	
IV	17 (32.7)	80 (25.2)	
N Stage			
0	34 (65.4)	196 (61.8)	0.582
I	11 (21.2)	87 (27.4)	
II	7 (13.5)	34 (10.7)	
Tumour subsite			
Right-sided colon	21 (40.4)	141 (44.9)	**0.015**
Left-sided colon	26 (50.0)	101 (32.2)	
Rectum	5 (9.6)	72 (22.9)	
GMS			
0	16 (30.8)	103 (33.2)	0.893
1	26 (50.0)	144 (46.5)	
2	10 (19.2)	63 (20.3)	
Phenotypic Subtype			
1	16 (30.8)	104 (33.7)	0.966
2	15 (28.8)	81 (26.2)	
3	11 (21.2)	62 (20.1)	
4	10 (19.2)	62 (20.1)	
mGPS			
0	18 (66.7)	128 (53.3)	0.325
1	4 (14.8)	63 (26.3)	
2	5 (18.5)	49 (20.4)	
MMR status			
pMMR	37 (71.2)	265 (84.1)	**0.023**
dMMR	15 (28.8)	50 (15.9)	
Tumour differentiation			
0	43 (82.7)	208 (88.3)	0.178
1	9 (17.3)	37 (11.7)	
Marginal involvement			
0	49 (94.2)	303 (95.6)	0.439
1	3 (5.8)	14 (4.4)	
Vascular invasion			
0	39 (75)	211 (66.6)	0.147
1	13 (25)	106 (33.4)	

Table 4.5 IL6R Expression and association with clinical factors. *Chi-squared test for associations between IL6R expression and clinicopathological factors including age, sex, T stage, N stage, tumour subsite, GMS, phenotypic subtype, mGPS, MMR status, tumour differentiation, marginal involvement and vascular invasion.*

Scores for IL6R and stromal IL6 were combined to form 3 groups: low for both markers, high for one, or high for both markers. This resulted in 109 patients classed as low for both, 235 for one high and 9 patients classed as high for both stromal IL6 and IL6R. Kaplan Meier survival analysis showed no

significant association between the combined stromal IL6 and IL6R score (p=0.184) (Figure 4.37). However, patients with high for both markers observed the worst prognosis and the study appears underpowered.

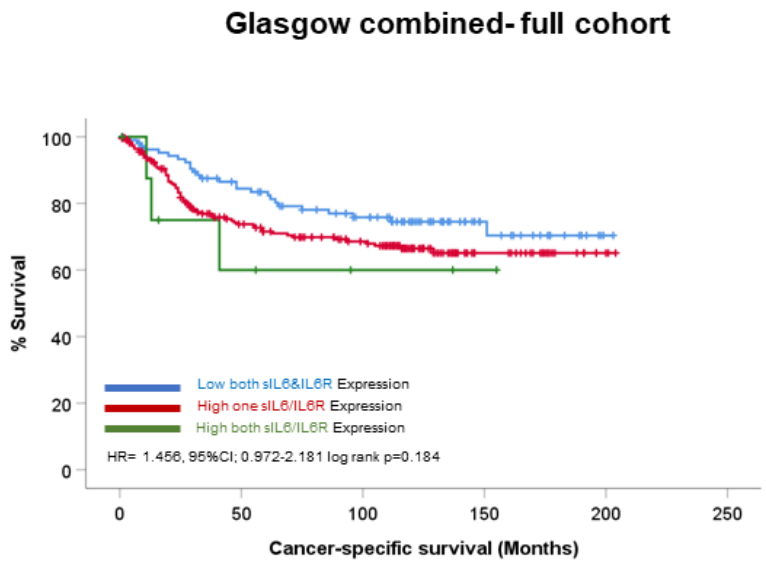


Figure 4.37 Association between combined stromal IL6/IL6R expression and Cancer-specific survival. *Kaplan Meier curve showing association between a combined score of stromal IL6 and cytoplasmic tumour IL6R and CSS in the full Glasgow combined cohort (n=353). The hazard ratio for a combined score of stromal IL6 and cytoplasmic tumoural IL6R in the full cohort was 1.456, (95%CI; 0.972-2.181), log rank p=0.184.*

Stromal IL8 and IL6R scores were combined to form a similar set of groups: low for both, high for on, or high for both. This resulted in 31 patients classed as low for both, 140 as high for one and 94 as high for both IL8 and IL6R. There was no significant association between combined IL8 and IL6R scores and CSS when a Kaplan Meier survival plot was constructed (p=0.099) (Figure 4.38). Patients low for both markers showed the best prognosis and those with one of both high observed similarly worse outcomes.

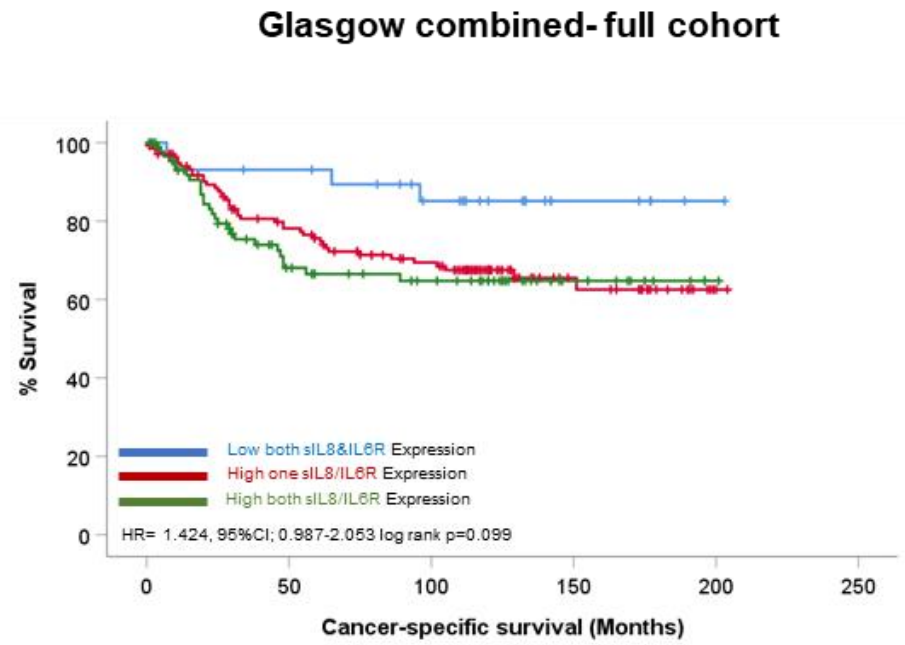


Figure 4.38 Association between combined stromal IL8/IL6R expression and Cancer-specific survival. *Kaplan Meier curve showing association between a combined score of stromal IL8 and cytoplasmic tumour IL6R and CSS in the full Glasgow combined cohort. The hazard ratio for the combined score of IL6R and stromal IL6 in GMS2 cases was 1.424, (95%CI; 0.987-2.053), log rank p=0.099.*

4.3 Discussion

The mechanisms underlying colorectal cancer development and progression are complex and heterogeneous. Dysregulation of cellular signalling pathways plays a role in tumorigenesis. IL6 is a key inflammatory cytokine involved in activating the JAK/STAT3 signalling pathway. Classical signal transduction occurs when IL6 binds membrane bound IL6R. Once bound, this triggers a cascade of signals leading to cancer growth, invasion, and metastases (75, 111, 117). IL6R can be found membrane-bound or in a soluble form in the cytoplasm. Signalling via soluble IL6R is known as trans-signalling, is associated with increased expression of adhesion markers promoting a pro-tumour microenvironment, independently of classical signalling (118). Cancer-associated fibroblasts (CAFs) are a major source of IL6 in the tumour microenvironment (119). In oesophageal cancer, stroma derived IL6 is associated with driving epithelial to mesenchymal transition (120). The data from this chapter showed a strong association between IL6R expression and poor outcomes, particularly in patients with stroma rich GMS2 tumours. The proximity of IL6 producing CAFs and tumour cells in these GMS2 cases could explain this observation (81). In terms of genetic subtypes of CRC, IL6 is also upregulated in the stromal group, CMS4 (121). Further studies using full tissue

sections could be employed to investigate the importance of the spatial orientation of IL6R+ tumour cells and CAFs producing IL6.

Tumour-associated macrophages (TAMs) are another source of IL6 in the TME (122). CAFs recruit TAMs to the TME via production of IL8 and increased expression of VCAM1 (119). IL8 is a proinflammatory chemokine which can act synergistically with IL6 to promote tumourigenesis. In prostate cancer IL8 exerts pro-tumour effects via STAT3, NFκß and AKT signalling (123). IL8 plays a role in driving CRC metastases (124). The crosstalk between IL8 and IL6 signalling has been observed in head and neck cancer (115). A feedback loop resulting in continual over-production of both IL6 and IL8 forms aiding tumour progression. Evidence for IL8 and IL6 pathway interaction is not well-characterised in colorectal cancer specifically and further studies are required to understand the importance of any crosstalk in this setting.

The results from this chapter indicate that IL6 within the stromal compartment of the TMA is of greater prognostic significance than IL6 located within tumour epithelial areas. High tumoural IL6 was associated reduced CSS ($p=0.045$), and this was potentiated in MMR deficient cases ($p=0.022$). The data suggested that IL8 was not associated with outcome in this cohort, however IL8 levels in the stroma were significantly higher in GMS2 tumours ($p=0.048$). Due to the soluble nature of IL6 and IL8, it was not possible to quantify protein expression, only mRNA copies, therefore further studies to investigate both cytokines in vitro and in vivo would be beneficial to corroborate findings. IL6R expression was assessed at the protein level via IHC. In patients with stromal-rich GMS2 tumours, high expression of IL6R was associated with reduced CSS ($p=0.009$) with a survival difference of 60.714 months between groups. High expression of IL6R was also associated with poorer outcomes in right-sided cases ($p=0.035$), with a mean survival difference of 31.733 months. IL6R expression within the tumour was stronger amongst rectal cases ($p=0.015$) and in males ($p=0.001$). These data, taken with evidence from the literature, suggest a link between stromal-tumour interactions and IL6/IL6R signal transduction. These results provide a first step towards the hypothesis that CMS4/GMS2 patients may be the target population for inhibition of STAT3 signalling.

Future work should include quantifying staining of IL6R within the stromal and immune cell compartments of the TME, given that a range of intensities was observed not only within the tumour cells. The level of IL6R expression in all areas could potentially associate with prognosis and be important for assessing pathway inhibition given that inhibitors would not necessarily be specific for tumour cells.

Chapter 5: Expression of JAK1, JAK2, JAK3 and TYK2 in colorectal cancer clinical specimens

5.1 Introduction

The IL6/JAK/STAT3 pathway has been implicated in CRC development and progression. Canonical signalling involves IL6 binding to membrane-bound IL6R resulting in activation of one or more Janus kinase tyrosine (JAK) family proteins (69) (Figure 5.1). This ultimately causes phosphorylation, dimerization and translocation to the nucleus of STAT3, where it can influence transcription of pro-tumour genes (69). There are four members of the JAK family of proteins; JAK1, JAK2, JAK3 and TYK2 (125). Studies investigating expression of JAK family oncokinases in large cancer patient cohorts are lacking. The aim of this chapter was to assess levels of JAK1, JAK2, JAK3 and TYK2 expression in the Glasgow combined colorectal tissue microarray by immunohistochemistry. Weighted histoscores for each protein were then analysed for association with clinical outcomes and characteristics. Prior to staining the patient cohort, antibodies were validated for specificity and IHC staining conditions were optimised on work-up tissue. Data were analysed with respect to histological phenotypic group GMS, MMR status and tumour subsite as a step towards identifying a specific subgroup of patients who may benefit from therapeutic targeting of the IL6/JAK/STAT3 pathway. It was hypothesised that high expression of JAK1, JAK2, TYK2 and JAK3 would be associated with reduced CSS and unfavourable clinical characteristics.

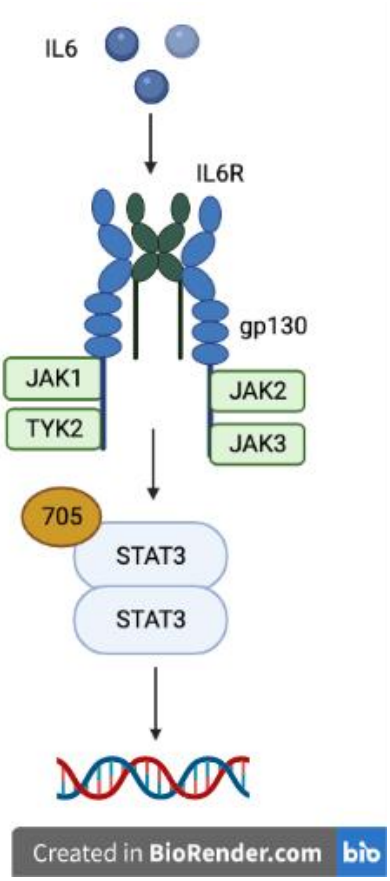


Figure 5.1. Diagram showing IL6/JAK/STAT3 signal transduction. *Representation of IL6 binding IL6R, initiating activation of JAK1, JAK2, TYK2 or JAK3 leading to activation, dimerization and translocation to the nucleus of STAT3. IL6 can activate all 4 JAK family members, however mist evidence to date suggests preferential activation of JAK1 and JAK2.*

5.2 Expression of JAK family members and association with clinical outcomes and features

The Glasgow combined cohort was stained for intermediate members of the IL6/JAK/STAT3 pathway, JAK1, JAK2, JAK3 and TYK2 via IHC. Patients were excluded due to missing or damaged TMA cores, mortality within 30 days of surgery and/or receiving neoadjuvant chemotherapy, as shown in consort diagram (Figure 5.1).

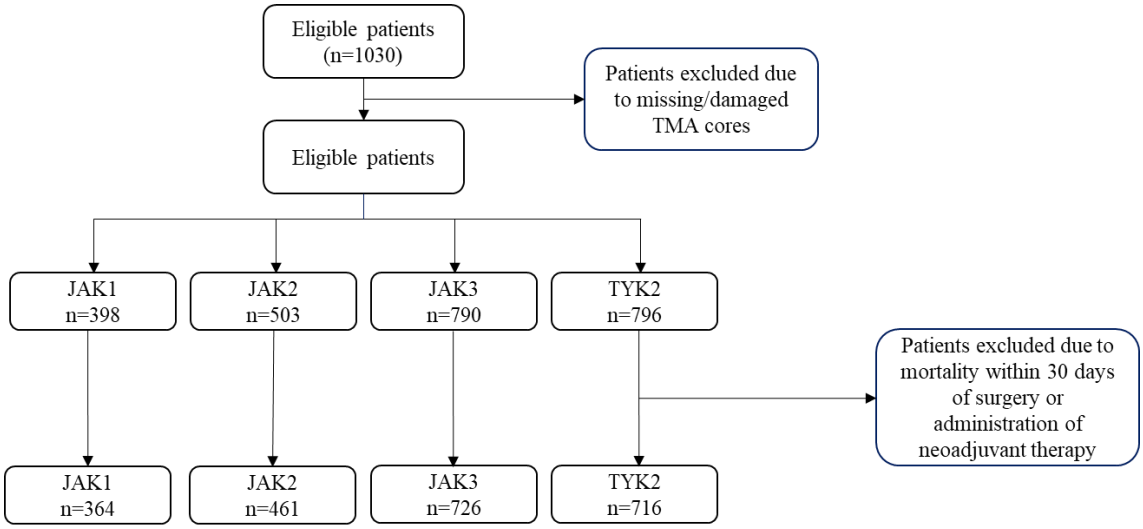


Figure 5.2 Consort diagram showing exclusion criteria. *Consort/flow diagram showing the number of patients included in analysis for each marker based on exclusions of missing/damaged cores, mortality within 30 days of surgery and/or administration of neoadjuvant chemotherapy. Following removal of patients meeting exclusion criteria, this left 364 with JAK1 data, 461 with JAK2, 726 for JAK3 and 716 stained for TYK2.*

5.2.1 Expression of JAK1 and clinical outcome

Cytoplasmic expression of Janus kinase-1 (JAK1) was assessed in the Glasgow combined cohort via immunohistochemistry. Representative images showing examples of weak, moderate, and strong cytoplasmic staining is shown in Figure 5.2. A true positive control, lung tissue, and a true negative control, prostate tissue, were stained alongside the TMA cohort (Figure 5.2). A positive and negative control colorectal section identified during the antibody optimisation process were also stained for JAK1 during the same run as the TMA staining (Figure 5.2). Antibody specificity was validated by western blotting. A single band in triplicate at 130KDa molecular weight was observed in HT29 colorectal cell lines indicating the presence of JAK1. No bands were present in adjacent wells which were loaded with PC3 prostate lysates, known to not express JAK1 (Figure 5.2). Tubulin was detected at similar intensity across all samples at 52KDa (Figure 5.2).

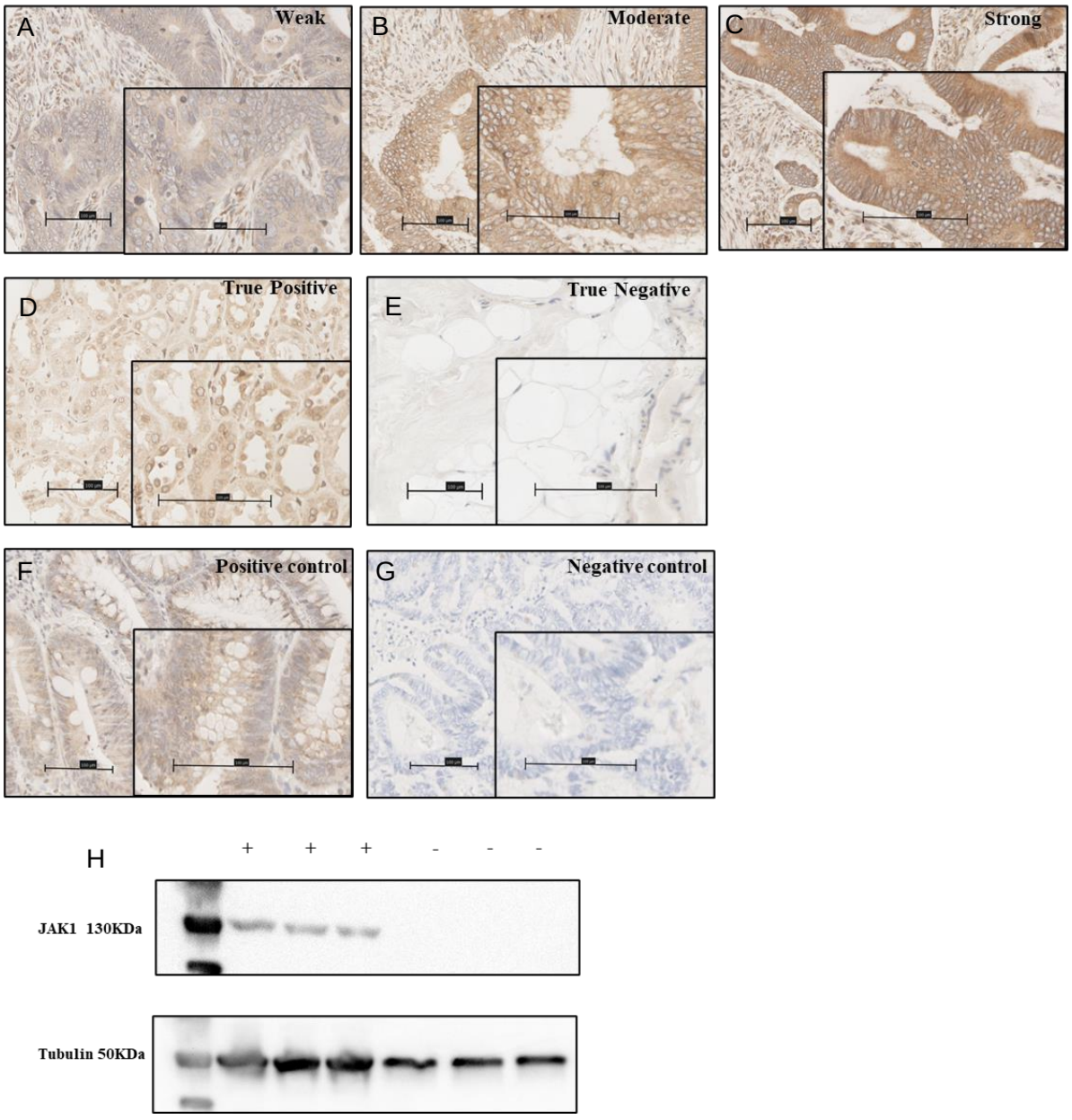


Figure 5.3 Representative images and antibody specificity for JAK1. *Images showing representative images of weak (A), moderate (B) and strong (C) cytoplasmic staining of JAK1 within tumour cells. Representative images of true positive and true negative tissue stained for JAK1 (D-E). Representative images of positive (D) and negative (E) colorectal tissue. Image of western blot showing single band in triplicate at 130KDa in HT29 cell lysates (20μL loaded per well) probed for JAK1 and no visible bands in triplicate for true negative PC3 lysates (20μL loaded per well), with bands visible at 50KDa for all samples probed for tubulin (H).*

Manual weighted histoscoring of tumour cytoplasmic expression of JAK1 was performed by a single observer (KP). Data were available for 398 patients due to missing and damaged cores. Of these

patients 34 were excluded due to administration of neoadjuvant chemotherapy or mortality within 30 days of surgery leaving 364 included in analysis. Scores ranged from 40 to 300 and data were normally distributed (Figure 5.3). To validate scoring, QuPath digital software was utilised to measure cytoplasmic JAK1 in 10% of TMA scores by Sara Al-Badran. A scatter plot was constructed to visualise correlation between manual and digital scores (Figure 5.4), with a correlation coefficient of 0.733 obtained showing strong positive association. A Bland Altman was plotted to ensure scores were sufficiently similar and not biased (Figure 5.5). The intra-class correlation coefficient of was 0.810 indicating the scores were strongly aligned. Cut off points for high and low expression were determined in R Studio using survminer and maxstat packages. The optimal cut point based on CSS was 137.70 as shown in histogram and scatterplot (Figure 5.6). This resulted in 151 patients classified as low for cytoplasmic JAK1 expression and 212 patients for high cytoplasmic JAK1 expression.

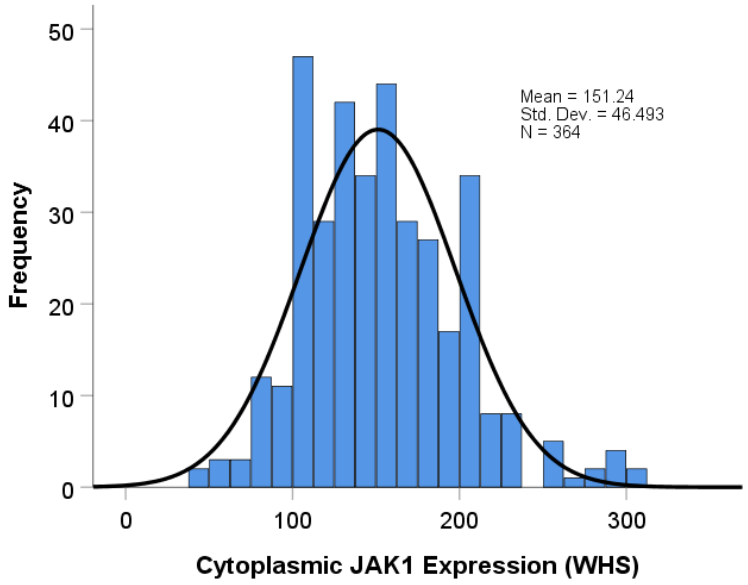


Figure 5.4 Distribution of JAK1 weighted histoscores. *Histogram showing JAK1 tumour cytoplasmic scores with normal distribution observed (n=364). The mean score for JAK1 was 151.24 and scores ranged from 0-300.*

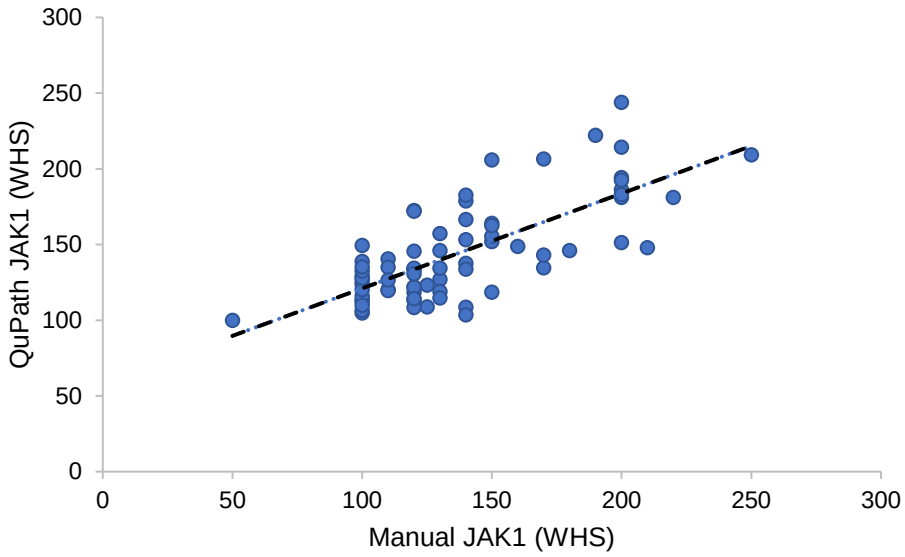


Figure 5.5 Correlation between Manual and QuPath cytoplasmic JAK1 scores. *Scatter plot showing the correlation between manual scores and digital scores for cytoplasmic JAK1 with a correlation of 0.733 obtained for the 10% of co-scored cores.*

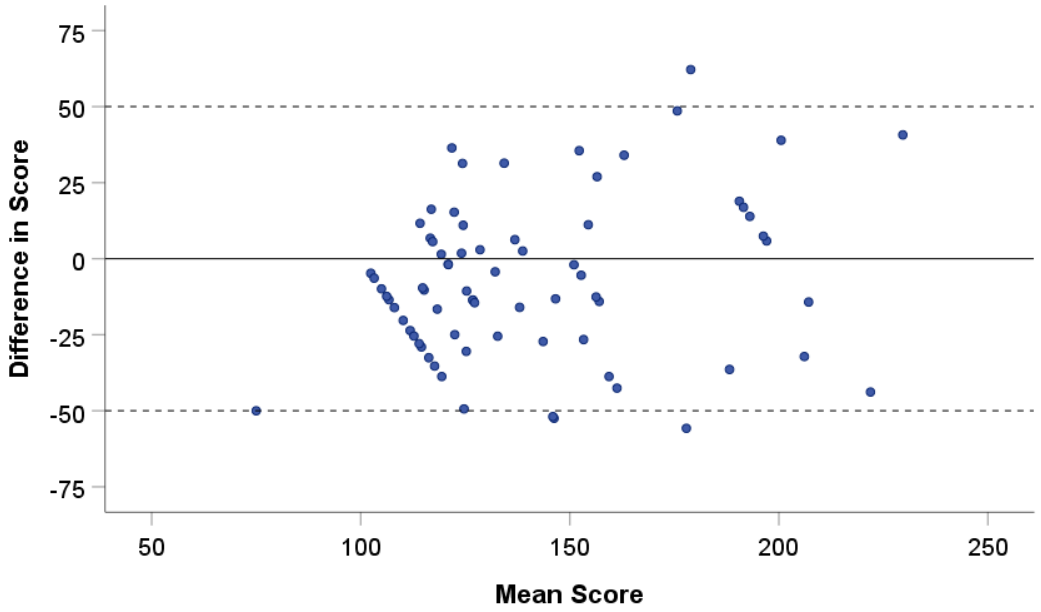


Figure 5.6 Difference between manual and digital JAK1 scores. *Bland Altman plot showing the similarity between manual and QuPath derived scores for cytoplasmic JAK1 in the 10% of Glasgow combined array co-scored.*

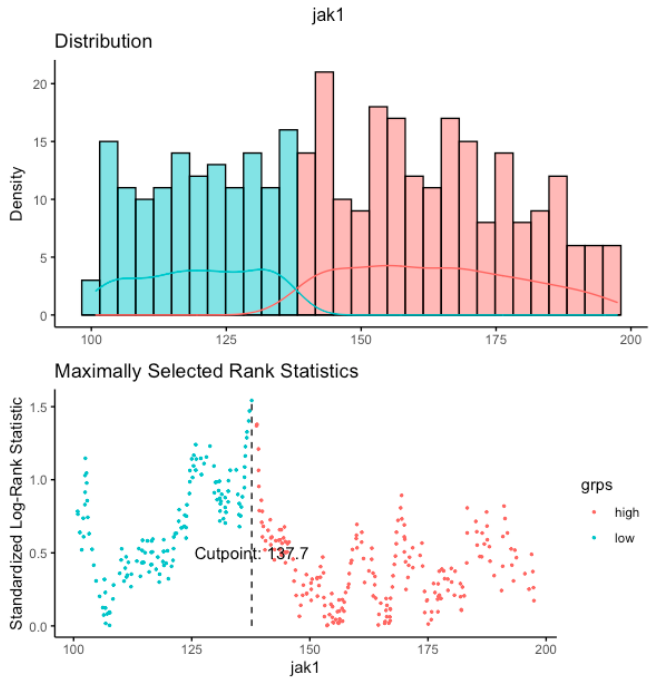


Figure 5.7 Defining optimal cut point for high and low expression of cytoplasmic JAK1 expression. *Density and scatter plot showing optimal cut off point for high and low expression of JAK1 groups based on cancer-specific survival. The optimal cut off point determined was 137.7, therefore patients with a weighted histoscore for JAK1 of ≥ 137.7 were considered high and patients with a score of ≤137.7 were classed as low for cytoplasmic JAK1.*

Kaplan Meier survival analysis was performed to determine any association between cytoplasmic JAK1 expression and CSS. Low expression of JAK1 was associated with reduced survival time (HR= 0.628, 95% CI; 0.414-0.951, log rank p= 0.026) (Figure 5.7). Patients with low expression of JAK1 had a mean survival time of 139.737 months compared to patients with high expression at 161.853 months. When patients were stratified by GMS subtype, this relationship was potentiated in GMS1 patients with low JAK1 expressing patients observing the best prognosis (HR=0.444, 95%CI; 0.235-0.839, log rank p=0.010) (Figure 5.8). The mean survival time of GMS1 patients with low JAK1 expression was 129.094 months versus GMS1 patients with high JAK1 expression at 167.897 months. GMS0 and GMS2 patients observed no survival benefit relative to JAK1 expression (Figure 5.8). Patients with MMR proficient tumours exhibited better outcomes with high JAK1 expression (HR=0.587, 95%CI; 0.376-0.9117, log rank p=0.018) (Figure 5.9). The mean survival time of pMMR patients with low JAK1 expression was 136.191 months compared to 161.113 months for the pMMR high JAK1 expression group. This was not observed in patients with MMR deficient tumours (Figure 5.9). In terms of sidedness, a trend towards survival benefit with high JAK1 expression was only observed in right-sided disease, however this did not reach significance (p=0.054; Figure 5.10). No association or trend between JAK1 expression and CSS was observed in patients with left-sided colonic tumours or rectal cases (Figure 5.10).

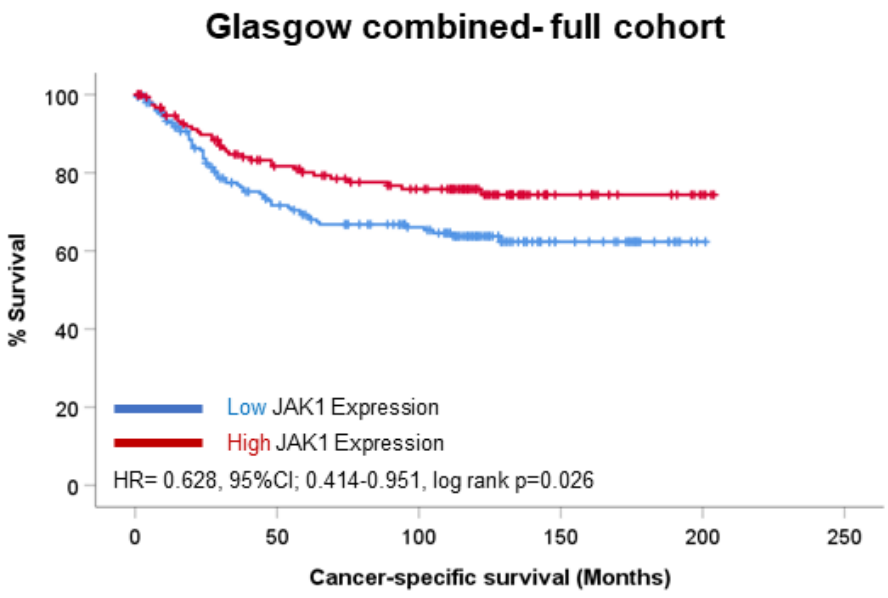


Figure 5.8 Association between tumoural cytoplasmic JAK1 expression and cancer-specific survival. *Kaplan Meier curve showing association between cytoplasmic JAK1 expression and CSS in patients from the Glasgow combined cohort. The hazard ratio for JAK1 expression in the full cohort was 0.628, (95%CI;0.414-0.951), log rank p=0.026.*

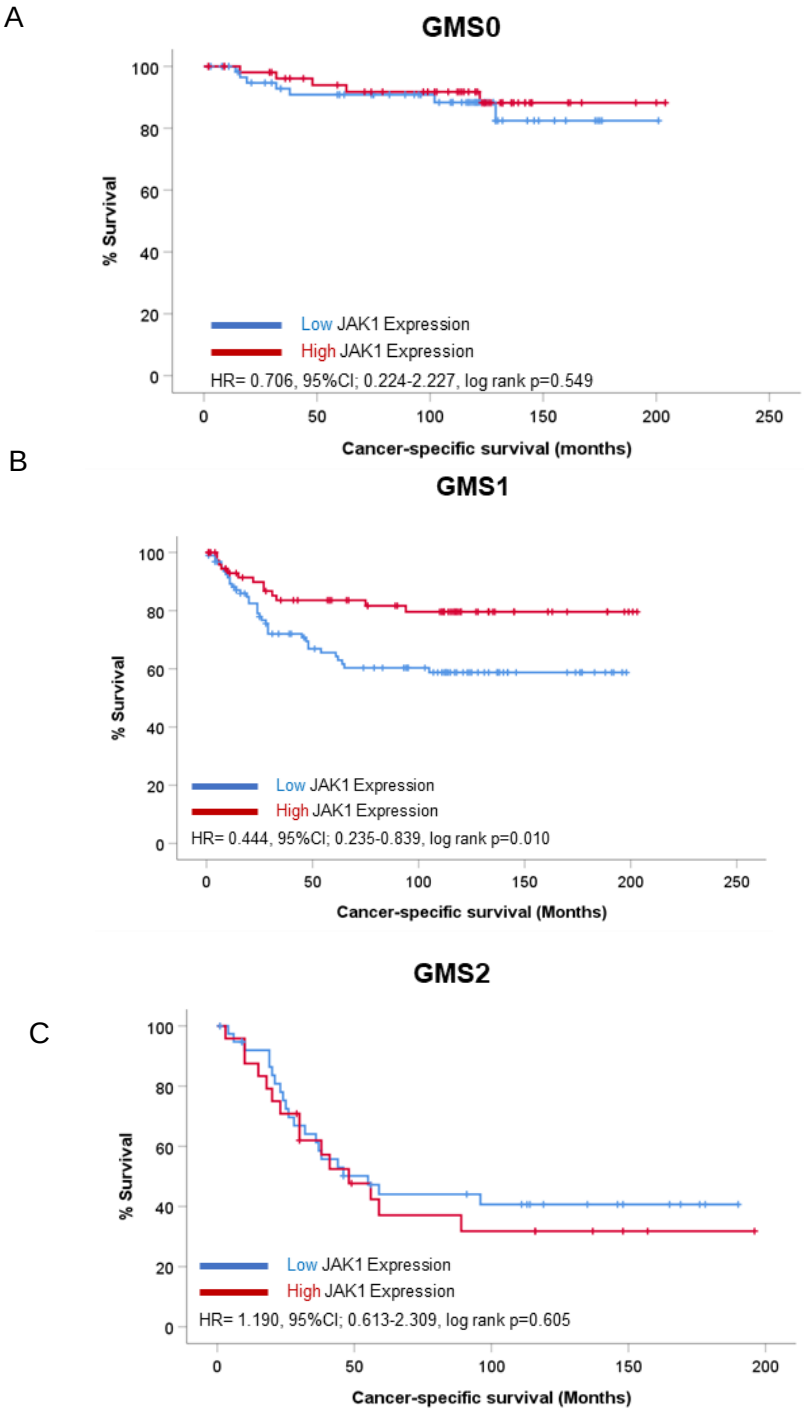


Figure 5.9 Association between JAK1 expression and CSS in each GMS classification. *Kaplan Meier survival analysis of JAK1 cytoplasmic expression and CSS in patients classified as GMS0 (A), GMS1 (B) and GMS2 (C). The hazard ratio for JAK1 expression in patients classed as GMS0 immune was 0.706, (95%CI;0.224-2.227), log rank p=0.549 and for GMS1 patients the hazard ratio was 0.444, (95%CI; 0.235-0.839), log rank p=0.010). In GMS2 stromal patients the hazard ratio for JAK1 expression was 1.190, (95%CI; 0.623-2.309), log rank p=0.605.*

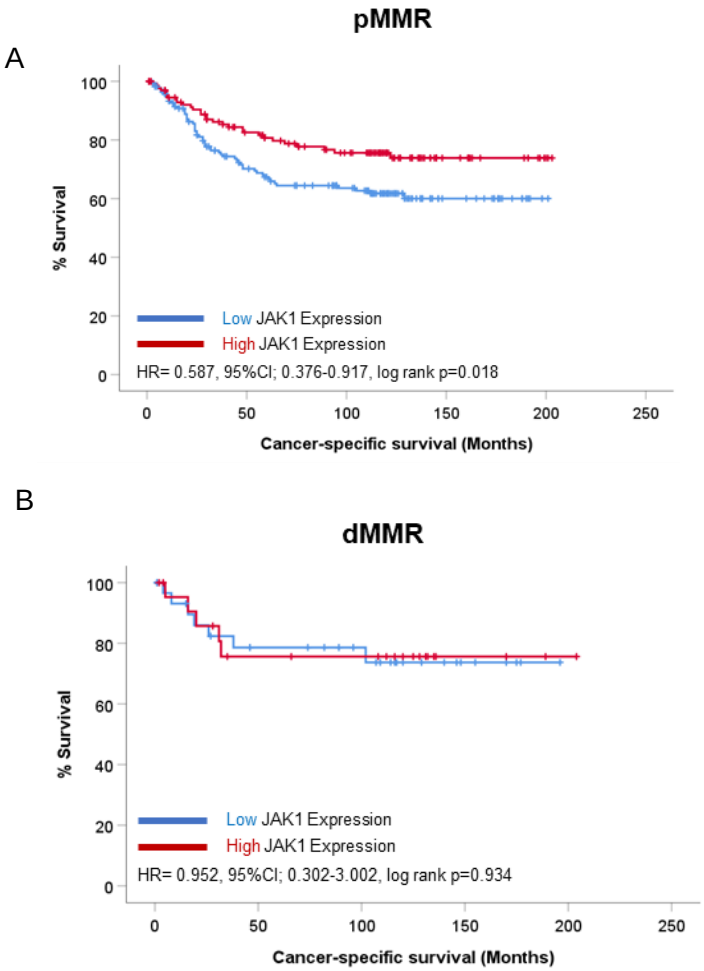


Figure 5.10 Association between JAK1 expression and CSS in pMMR and dMMR cases. *Kaplan Meier survival analysis of JAK1 cytoplasmic expression and CSS in MMR proficient patients (A) and MMR deficient patients (B). In MMR proficient patients the hazard ratio for JAK1 expression was 0.587, (95%CI; 0.376-0.917), log rank p= 0.018 and for MMR deficient cases the hazard ratio was 0.952, (95%CI; 0.302-3.002), log rank p=0.934.*

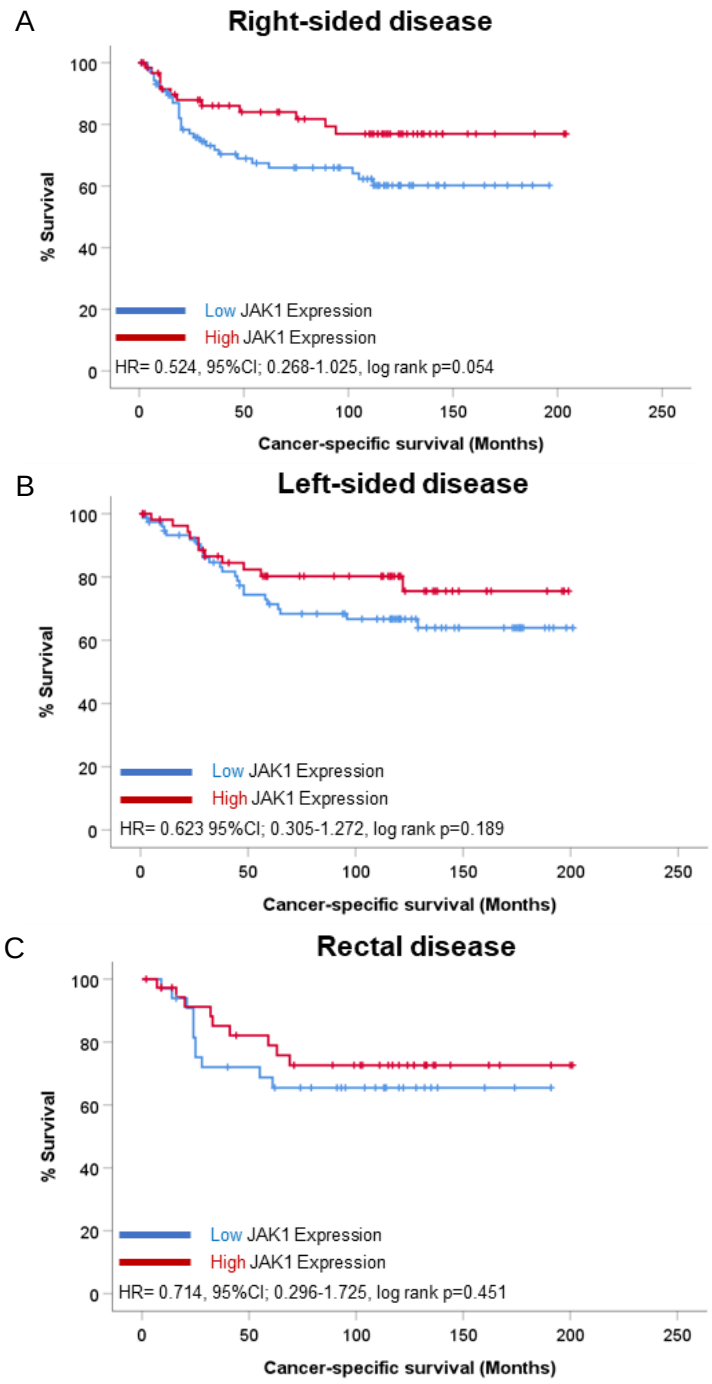


Figure 5.11 JAK1 expression, CSS and tumour subsite. *Kaplan Meier survival analysis showing association between JAK1 expression and CSS in right-sided disease (A), left-sided disease (B) and rectal disease (C). For patients with right-sided disease the hazard ratio for JAK1 expression was 0.524, (95%CI; 0.268-1.025), log rank p=0.054 and for patients with left-sided colonic tumours the hazard ratio was 0.623, (95%CI; 0.305-1.272), log rank p=0.189. For patients with rectal disease the hazard ratio relative to JAK1 expression was 0.714, (95%CI; 0.296-1.725), log rank p=0.451.*

JAK1 expression was assessed for association with clinical features by chi-squared analyses. No significant relationships were identified. However, there was a trend towards an increase in tumour vascularisation in patients with low cytoplasmic tumoural JAK1 expression ($p=0.066$) (Table 5.1).

Clinical Factor	Cytoplasmic JAK1 Expression		p
	Low (n=151)	High (n=212)	
Age			
<65	50 (32.9)	70 (33.0)	0.536
>65	102 (67.1)	142 (67.0)	
Sex			
Female	76 (50.0)	107 (50.5)	0.507
Male	76 (50.0)	105 (49.5)	
T Stage			
I	9 (5.9)	7 (3.3)	0.526
II	19 (12.5)	32 (15.1)	
III	80 (52.6)	118 (55.7)	
IV	44 (28.9)	55 (25.9)	
N Stage			
0	82 (53.9)	133 (62.7)	0.197
I	48 (31.6)	58 (27.4)	
II	22 (14.5)	21 (9.9)	
Tumour subsite			
Right-sided colon	68 (44.7)	87 (41.6)	0.170
Left-sided colon	61 (40.1)	74 (35.4)	
Rectum	23 (15.1)	48 (23.0)	
GMS			
0	45 (30.4)	74 (35.4)	0.265
1	70 (47.3)	102 (48.8)	
2	33 (22.3)	33 (15.8)	
Phenotypic Subtype			
1	44 (29.7)	75 (35.7)	0.412
2	44 (29.7)	59 (28.1)	
3	27 (18.2)	42 (20.0)	
4	33 (22.3)	34 (16.2)	
mGPS			
0	63 (56.8)	87 (56.9)	0.690
1	27 (24.3)	29 (19.9)	
2	21 (18.9)	30 (20.5)	
MMR status			
pMMR	127 (84.1)	181 (85.8)	0.383
dMMR	24 (15.9)	30 (14.2)	
Tumour differentiation	137 (90.1)	186 (87.7)	0.295
0	15 (9.9)	26 (12.3)	
1			
Marginal involvement	142 (93.4)	206 (97.2)	0.073
0	10 (6.6)	6 (2.8)	
1			
Vascular invasion			0.066
0	93 (61.2)	No 147 (69.3)	
1	59 (38.8)	65 (30.7)	

Table 5.1 Association between JAK1 expression and clinical features. *Chi-squared table of associations for JAK1 expression and clinical prognostic factors including age, sex, T stage, N stage, tumour subsite, GMS, phenotypic subtype, mGPS, MMR status, tumour differentiation, marginal involvement and vascular invasion.*

Positive staining for JAK1 was also observed within tumour membranes of some patients. Modified histoscoring was performed by a single observer (KP) to determine intensity of membranous JAK1, with 0 being absent, 1 being present. This resulted in 71 patients classed as absent for membranous JAK1 and 210 patients classed as present for membrane JAK1 staining. Kaplan Meier survival analysis showed no association between membranous staining of JAK1 and CSS in the full Glasgow combined cohort (p=0.199) (Figure 5.11).

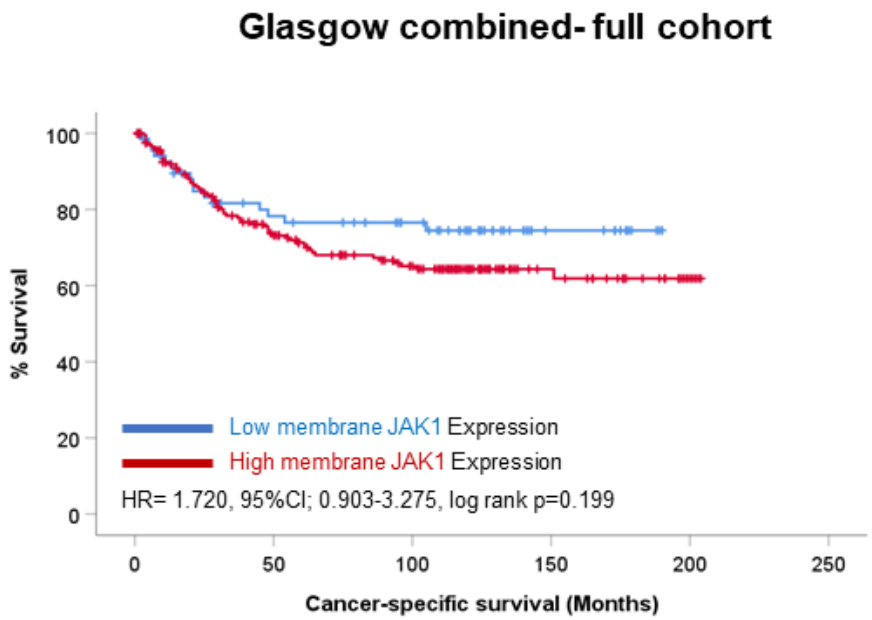


Figure 5.12 Membrane JAK1 expression and CSS with patients stratified by tumour subsite. *Kaplan Meier survival analysis of membrane JAK1 expression, which is a surrogate of activation, in patients from the Glasgow combined cohort. The hazard ratio for membranous expression of JAK1 in the full cohort was 1.720, (95%CI; 0.903-3.275), log rank p=0.199.*

5.2.2 Expression of JAK2 and clinical outcome

Cytoplasmic janus kinase-2 (JAK2) expression was assessed in the Glasgow combined cohort TMA via immunohistochemistry. Representative images of weak, moderate, and high cytoplasmic staining expression patterns are shown in Figure 5.12 (A-C). True positive tissue (lung) shows positive staining for JAK2, and the true negative prostate tissue known to not express JAK2 at high levels is shown in figure 5.12. Positive and negative control colorectal tissue was used during the antibody optimisation process (Figure 5.12). To test antibody specificity SW620 cell lines were treated with either 0.01% DMSO or AT9283 JAK2 inhibitor. Cell pellets were then stained using the JAK2 antibody at the same concentration and using identical conditions as the TMA staining protocol. Cell

pellets with JAK2 inhibited showed reduced expression of JAK2 when compared to the vehicle control (Figure 5.12, H).

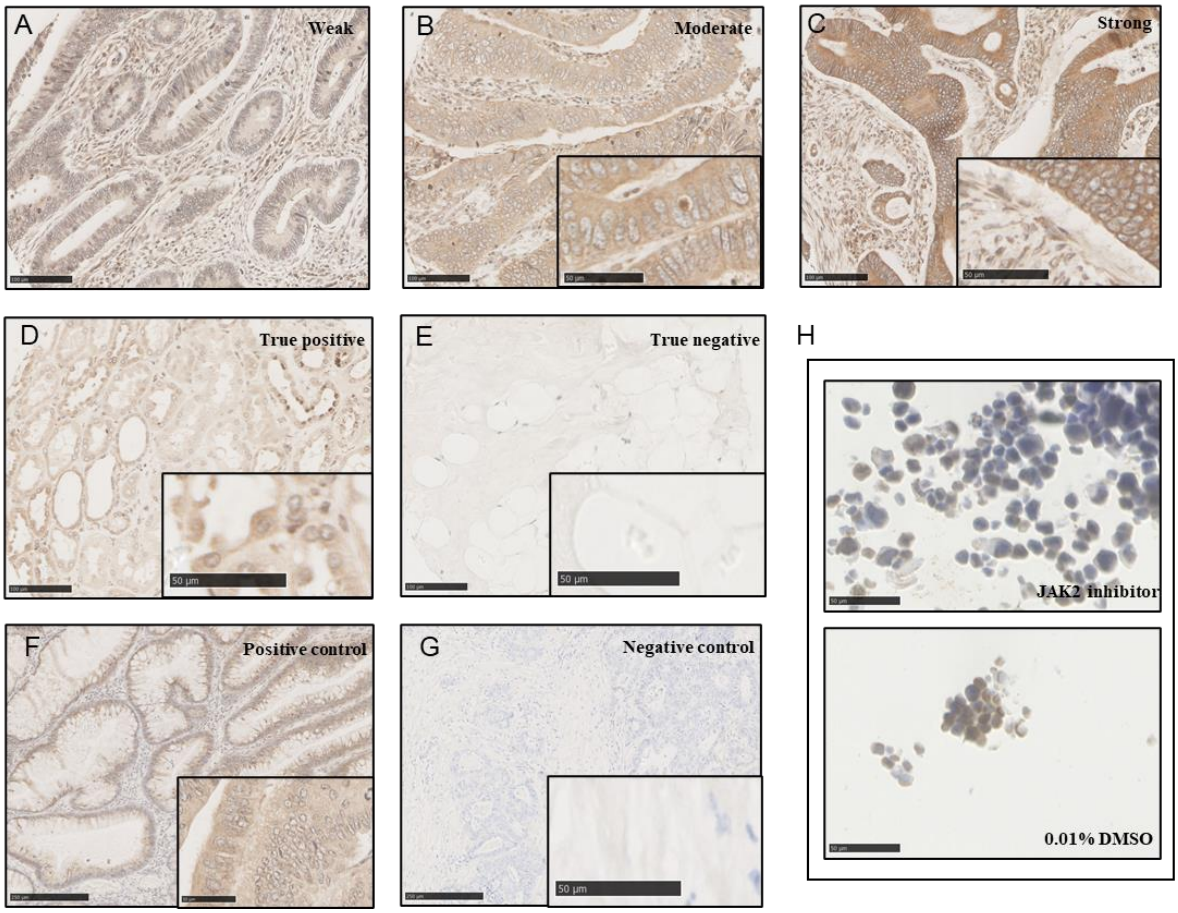


Figure 5.13 Representative images and antibody specificity for JAK2. *Images showing representative images of weak (A), moderate (B) and strong (C) cytoplasmic staining of JAK2 within tumour cells. Representative images of positive (D) and negative (E) colorectal tissue, and true positive (F) and true negative tissue (G). Representative images of cell pellets treated with a JAK2 inhibitor and vehicle control (DMSO) and subsequently stained for JAK2 via IHC to observe lower intensity of staining in inhibitor-treated cells relative to vehicle control treated cells (H).*

Expression of cytoplasmic Janus kinase-2 was scored manually by a single observer (KP) in 503 patients from the 1030 Glasgow combined patient cohort due to missing or damaged cores. Of these patients 42 were subsequently excluded due to mortality within 30 days of surgery or administration of neoadjuvant chemotherapy leaving 461 included in downstream analysis. Manual scores were validated by co-scoring of 10% of the cohort with QuPath digital pathology software by Sara Al-Badran. Manual weighted histoscores ranged from 0 to 240 with a relatively normal distribution pattern from histogram (Figure 5.13). Manual scores were strongly positively correlated with the 10% QuPath scores with a correlation coefficient of 0.815 indicating a strong positive relationship. A

scatter plot was constructed to visualise the correlation between scores (Figure 5.14). A Bland Altman was plotted to show minimal bias between scores and an intraclass correlation coefficient of 0.898 was obtained indicating the scores were significantly similar (Figure 5.15). To determine optimal cut points for high and low expression based on cancer-specific survival log rank statistics were performed using R Studio packages survminer and maxstat. A density and scatter plot were constructed to visualise the optimal cut point (Figure 5.16). The optimal cut point determined was 67.37, therefore patients with a weighted histoscore ≥67.37 were considered to highly express JAK2 (n=158) and those with a weighted histoscore ≤67.37 were classified as low for cytoplasmic JAK2 expression (n=219).

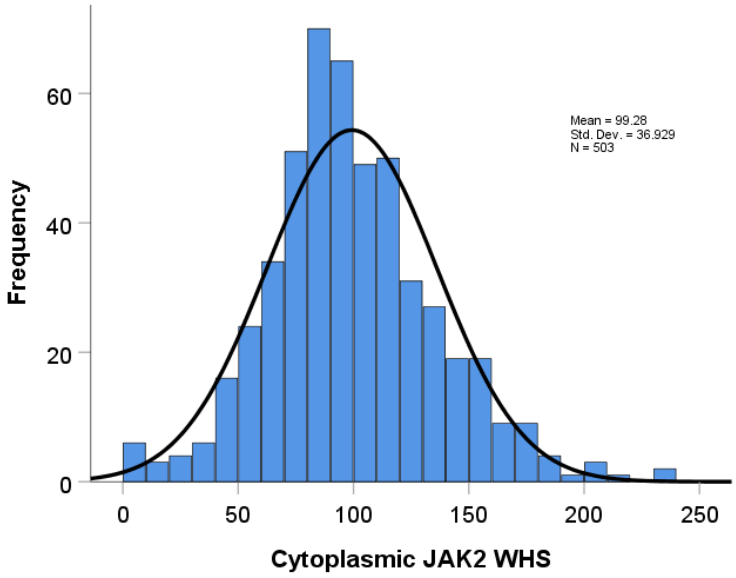


Figure 5.14 Distribution of JAK2 weighted histoscores. *Histogram showing JAK2 tumour cytoplasmic scores with a normal distribution (n=503). The mean score for JAK2 was 99.28 with scores ranging from 0-240.*

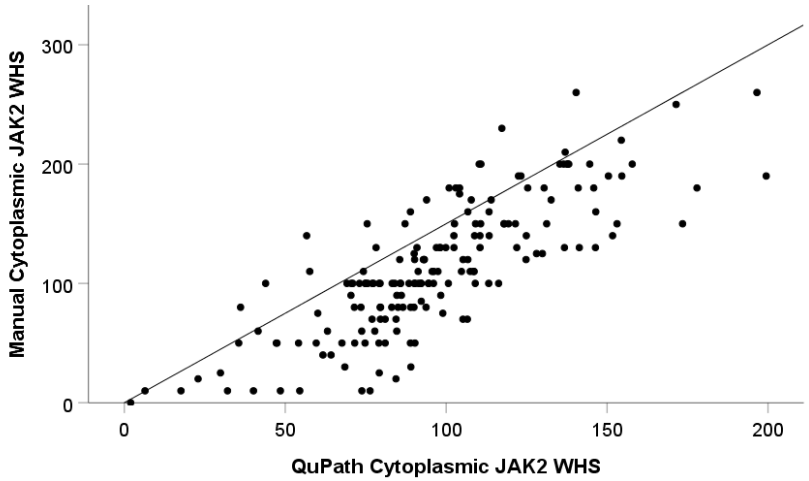


Figure 5.15. Correlation between Manual and QuPath cytoplasmic JAK2 scores. *Scatter plot showing the correlation between manual scores and digital scores for cytoplasmic JAK2 for the 10% of patients co-scored from the cohort. A correlation coefficient of 0.815 was obtained.*

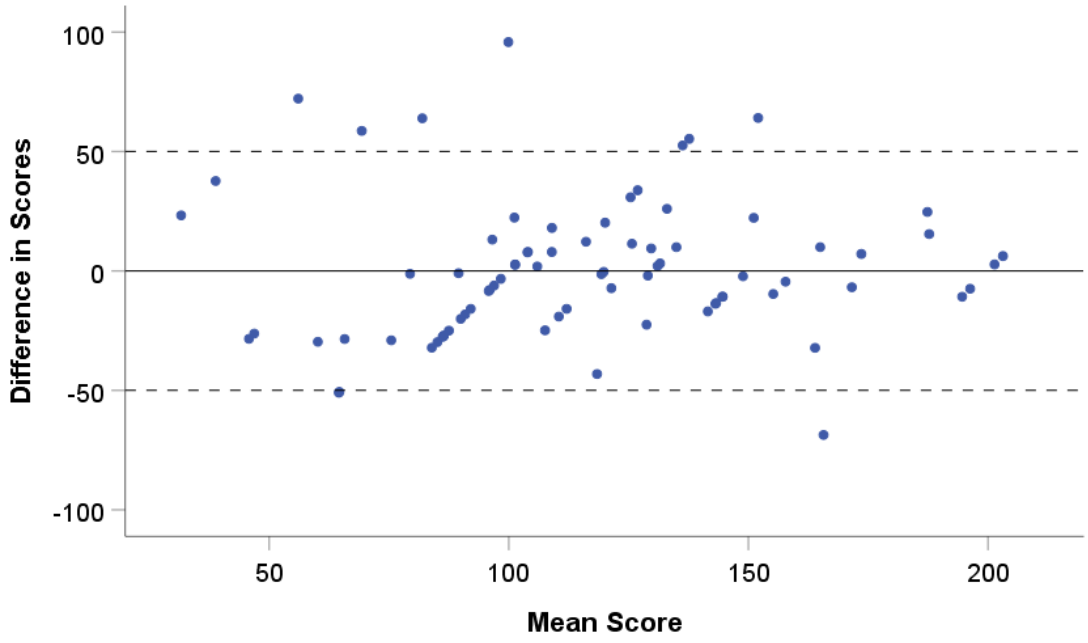


Figure 5.16. Difference between manual and digital JAK2 scores. *Bland Altman plot showing the similarity between manual and QuPath derived scores for cytoplasmic JAK2 in the 10% of patients Glasgow combined array co-scored.*

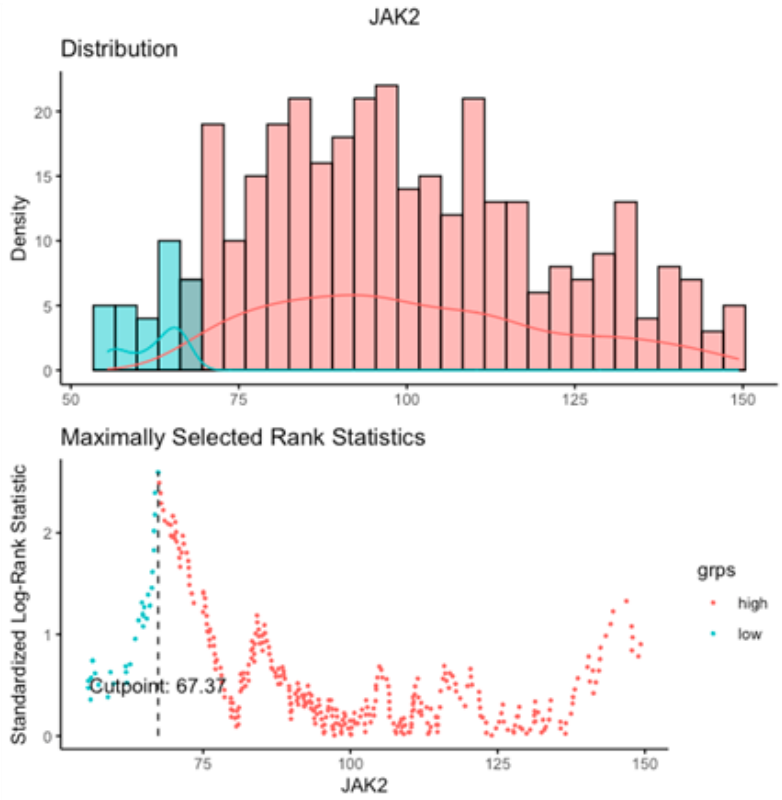


Figure 5.17. Defining optimal cut point for high and low expression of cytoplasmic JAK2 expression. *Density and scatter plot showing optimal cut off point for high and low expression of JAK2 groups based on cancer-specific survival. The optimal cut off point identified was 67.37, therefore patients with a weighted histoscore of ≤67.37 for JAK2 were classed as low and patients with a weighted histoscore of ≥67.37 were classified as high for JAK2 cytoplasmic expression.*

Kaplan Meier survival analysis was performed to determine association between cytoplasmic JAK2 expression and CSS. Cytoplasmic JAK2 expression was not associated with CSS in the full cohort (HR=1.162, 95%CI; 0.790-1.710, log rank p=0.444; Figure 5.17). When patients were segregated based on GMS, JAK2 expression was significantly associated with CSS in the stromal GMS2 subtype (HR=1.881, 95%CI;0.992-3.565, p=0.048; Figure 5.18). Patients with GMS2 classified tumours and high JAK2 expression had a mean survival time of 66.893 months compared to 108.169 months for patients with low cytoplasmic JAK2 expression. In immune GMS0 and intermediate GMS1 patients cytoplasmic JAK2 expression in tumour cells showed no association with CSS (Figure 5.18). There was no association between JAK2 expression and CSS when patients were grouped by MMR status (Figure 5.19). In terms of sidedness, no significant association between JAK2 expression and CSS was observed at any tumour subsite (Figure 5.20).

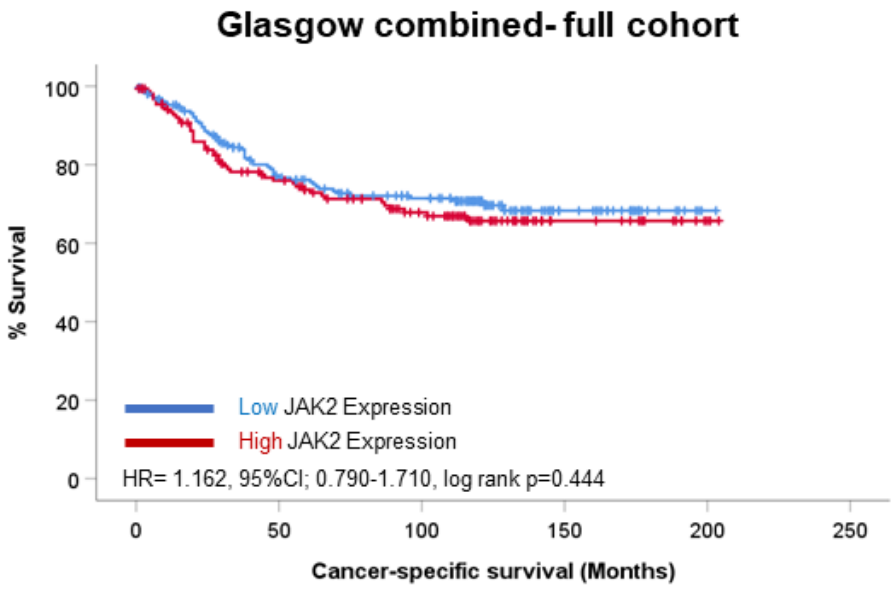


Figure 5.18 Cytoplasmic JAK2 expression and CSS in the full cohort. *Kaplan Meier survival analysis of cytoplasmic JAK2 expression in the Glasgow combined cohort. The hazard ratio for JAK2 expression in the full cohort was 1.162, (95%CI; 0.790-1.710), log rank p=0.444.*

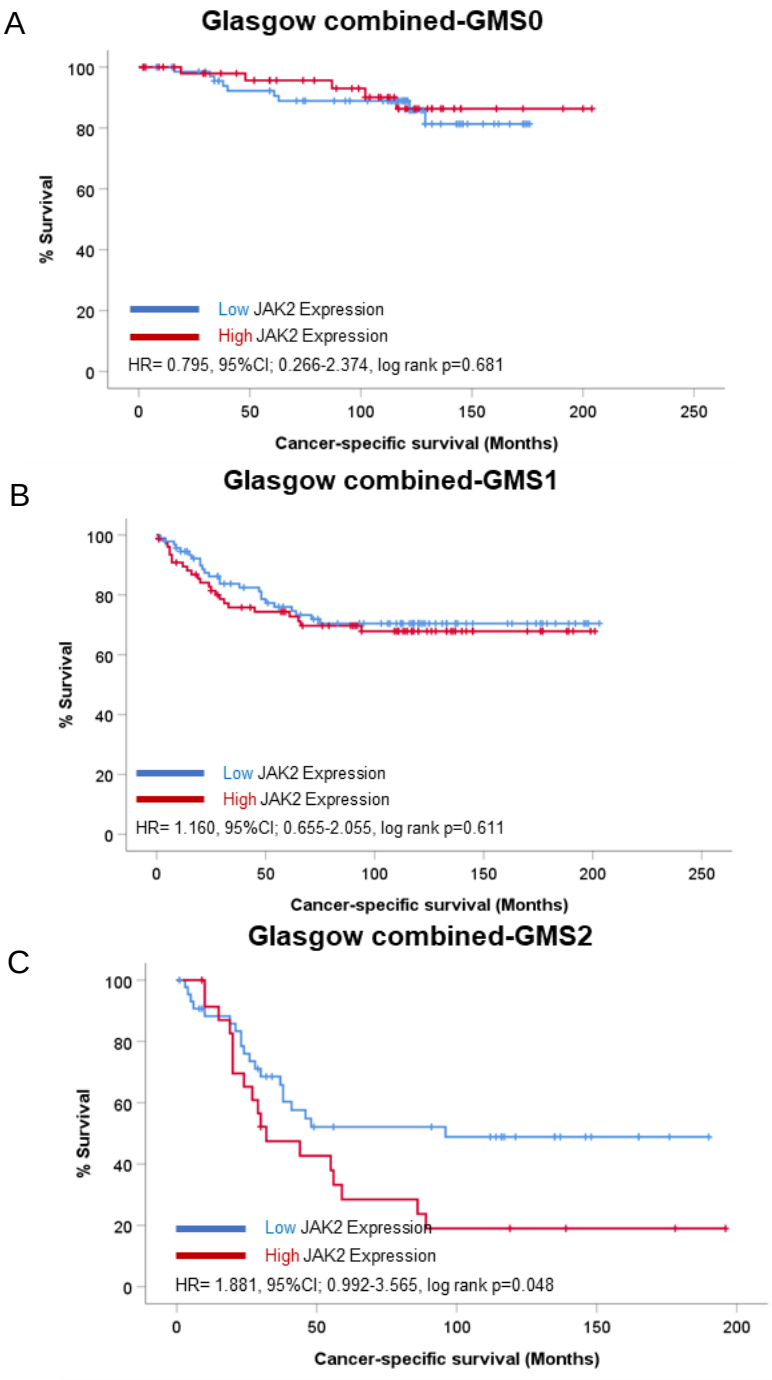


Figure 5.19 Cytoplasmic JAK2 expression and CSS with patients stratified by GMS. *Kaplan Meier survival analysis of cytoplasmic JAK2 expression in GMS0 patients (A), GMS1 patients (B) and GMS2 patients (C). In patients with GMS0 immune graded tumours the hazard ratio for JAK2 expression was 0.795, (95%CI; 0.266-2.374), log rank p=0.681 and for patients with GMS1 tumours the hazard ratio was 1.160, (95%CI; 0.655-2.055), log rank p=0.611. Patients with stromal-rich GMS2 tumours observed a hazard ratio of 1.881, (95%CI; 0.992-3.565, log rank p=0.048 for cytoplasmic JAK2 expression.*

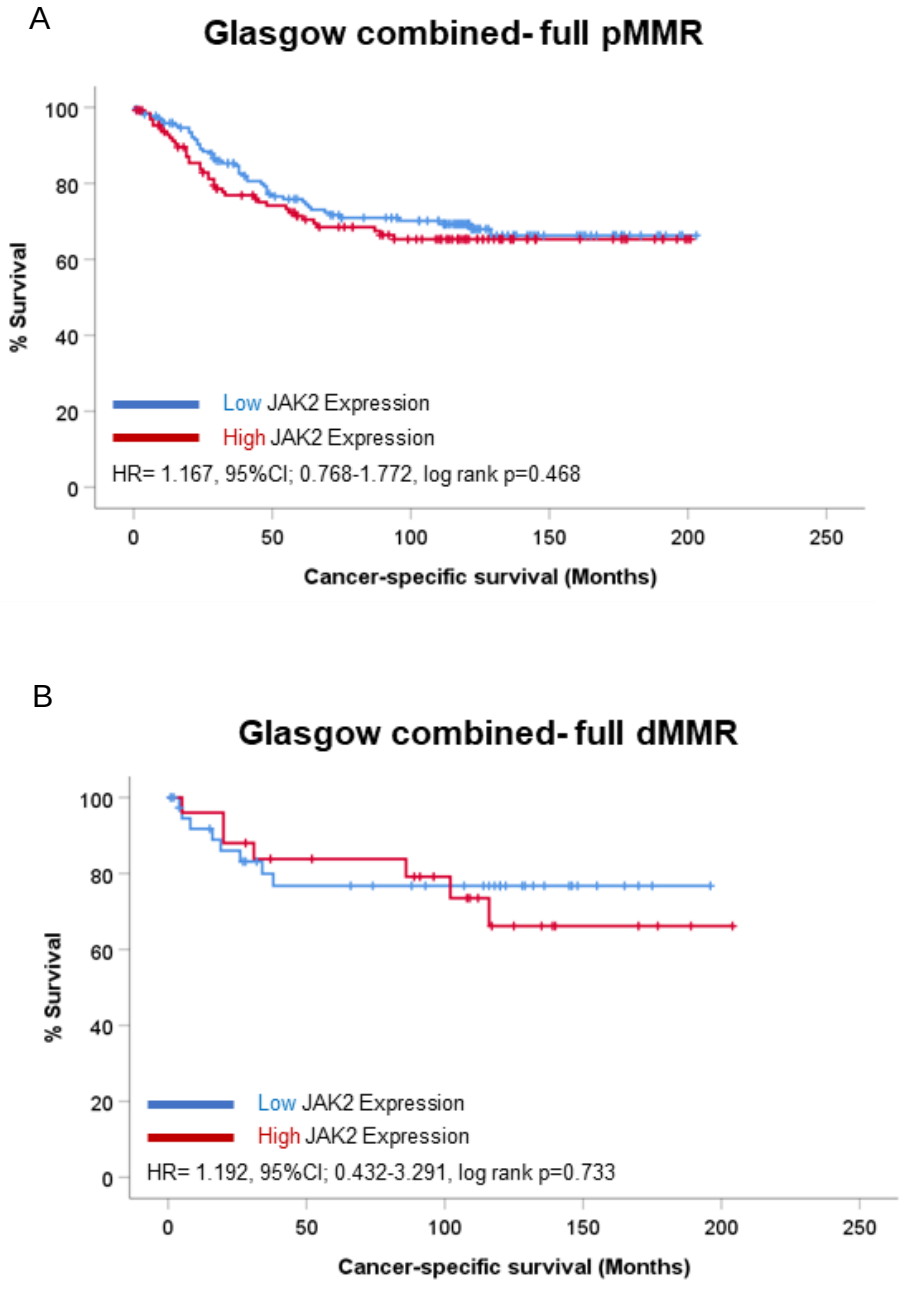


Figure 5.20 Cytoplasmic JAK2 expression and CSS with patients stratified by MMR status. *Kaplan Meier survival analysis of cytoplasmic JAK2 expression in patients with MMR proficient tumours (A) and MMR deficient tumours (B). In MMR proficient cases, the hazard ratio for JAK2 expression was 1.167, 95%CI; 0.768-1.772), log rank p=0.468 and for patients with MMR deficient disease the hazard ratio for JAK2 was 1.192, (95%CI; 0.432-3.291), log rank p=0.733.*

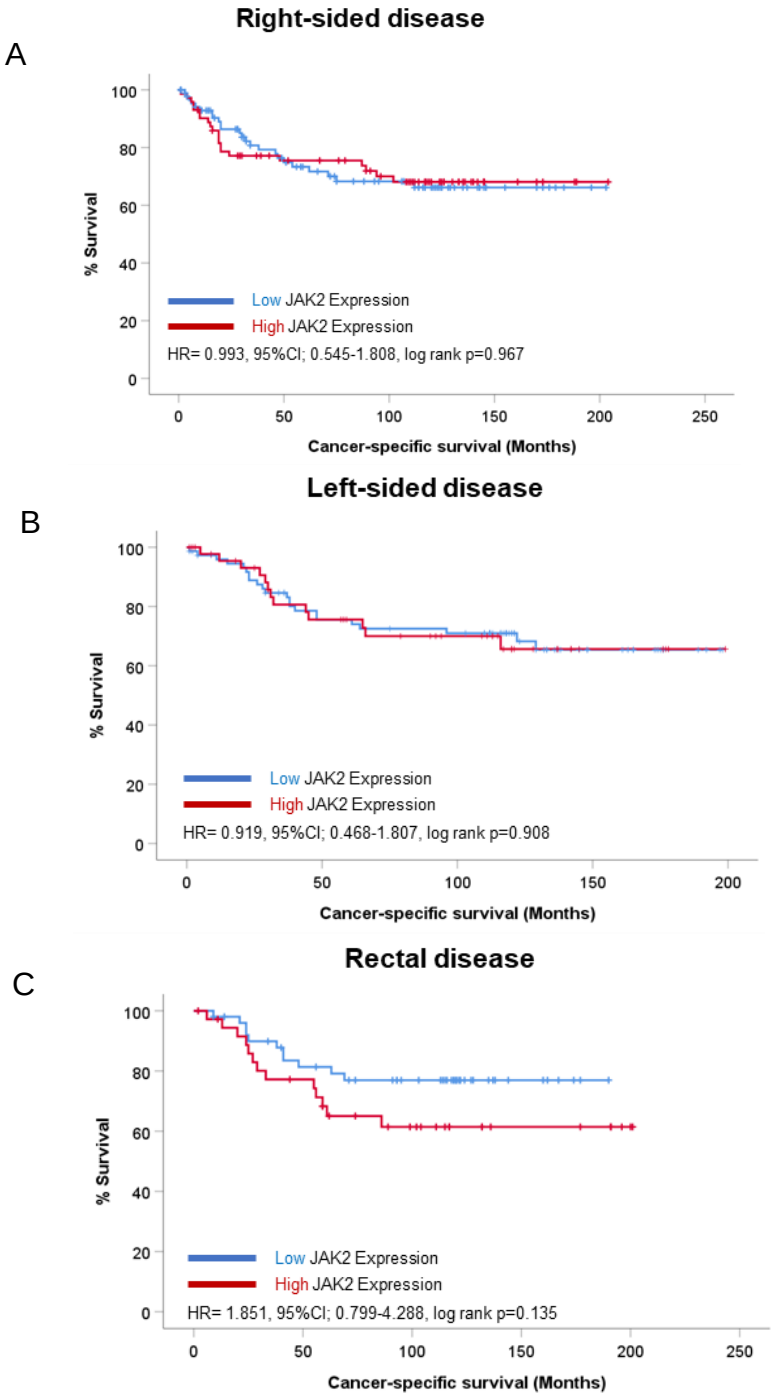


Figure 5.21 Cytoplasmic JAK2 expression and CSS with patients stratified by tumour subsite. *Kaplan Meier survival analysis of cytoplasmic JAK2 expression in patients with right-sided tumours (A), left-sided tumours (B) and rectal tumours (C). In patients with right-sided tumours the hazard ratio associated with JAK2 expression was 0.993, (95%CI:0.545-1.808), log rank p=0.967 and for patients with left sided tumours the hazard ratio was 0.919, (95%CI; 0.468-1.807), log rank p=0.908. The hazard ratio for JAK2 expression in patients with rectal tumours was 1.851, (95%CI; 0.799-4.288), log rank p=0.135.*

Next, JAK2 expression was assessed for association with clinicopathological characteristics via Chi-squared analyses (Table 5.2). Cytoplasmic JAK2 was associated with T stage ($p<0.001$), N stage ($p=0.035$), modified Glasgow prognostic score (systemic inflammation) ($p=0.029$), tumour budding ($p=0.027$), phenotypic subtype ($p=0.047$) and marginal involvement ($p=0.041$).

Clinical feature	Cytoplasmic JAK2 Expression		*p* value
	Low (n=219)	High (n=158)	
Age >65 <65	 24 (32.2) 50 (67.6)	 132 (34.1) 255 (65.9)	0.446
Sex Female Male	 30 (40.5) 44(59.5)	 196 (50.6) 191 (49.4)	0.071
T Stage I II III IV	 3 (4.1) 0(0) 42 (56.8) 29 (39.2)	 15 (3.9) 61 (15.8) 212 (54.8) 99 (25.6)	**<0.001****
N Stage I II III	 36 (48.6) 25 (33.8) 13 (17.6)	 249 (64.3) 98 (25.3) 40 (10.3)	**0.035***
Tumour subsite Right Left Rectum	 33 (44.6) 26 (35.1) 15 (20.3)	 169 (43.7) 129 (33.3) 89 (23.0)	0.868
mGPS 0 1 2	 24 (42.9) 23 (41.1) 9 (16.1)	 171 (58.4) 69 (23.5) 53 (18.1)	**0.029***
Tumour budding Absent-Low Moderate-High	 37 (59.7) 25 (40.3)	 253 (72.9) 94 (27.1)	**0.027***
Phenotypic subtype Immune Canonical Latent Stromal	 18 (25.7) 16 (22.9) 21 (30.0) 15 (21.4)	 145 (38.2) 101 (26.6) 65 (17.1) 69 (18.2)	**0.047***
GMS 0 1 2	18 (25.7) 36 (51.4) 16 (22.9)	143 (37.5) 169 (44.4) 69 (18.1)	0.148
Marginal Involvement Absent Present	66 (89.2) 8 (10.8)	369 (95.3) 18 (4.7)	**0.041***

Table 5.2 JAK2 Expression and Clinical Features. *Table of Chi-squared associations between JAK2 expression and clinicopathological prognostic features including age, sex, T stage, N stage, tumour subsite, GMS, phenotypic subtype, MMR status, mGPS, tumour budding and marginal involvement.*

JAK2 staining was also present in the tumour membrane. Modified histoscoring was performed by a single observer (KP) to determine intensity of membranous JAK2, with 0 being absent, 1 being present. This resulted in 64 patients classified as present for membrane JAK2, and 234 absent for

JAK2 membranous expression. There was no significant association between membranous JAK2 and CSS when Kaplan Meier survival curves were plotted (p=0.406) (Figure 5.21).

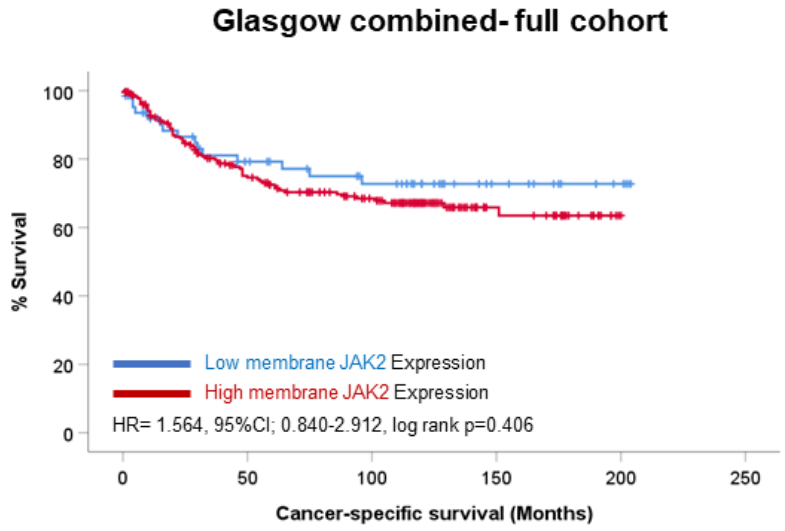


Figure 5.22 Membrane JAK2 expression and CSS with patients stratified by tumour subsite. *Kaplan Meier survival analysis of membrane JAK2 expression in patients from the Glasgow combined cohort. In the full cohort the hazard ratio for membranous JAK2 expression was 1.564, (95%CI; 0.840-2.912) log rank p=0.406.*

Membrane scores for JAK1 and JAK2 were combined to form a new score of absent for both markers, one marker present, and both markers present. There were 147 patients with JAK1 and JAK2 membrane expression present, 78 with presence of JAK1 or JAK2 membranous staining and 17 exhibiting no membranous staining for JAK1 or JAK2. In terms of prognosis, patients with high JAK1 and JAK2 membrane staining had poorer outcomes in terms of CSS when compared to patients low for both (HR= 1.638, 95%CI; 1.059-2.534, log rank p= 0.039) (Figure 5.22). The mean survival time of patients with high membranous JAK1/2 was 77.35 months compared to patients with low expression at 93.44 months. Interestingly, this observation was potentiated in GMS2 stromally dense tumours (HR=3.227, 95%CI; 1.410-7.387, log rank p=0.013) (Figure 5.23).

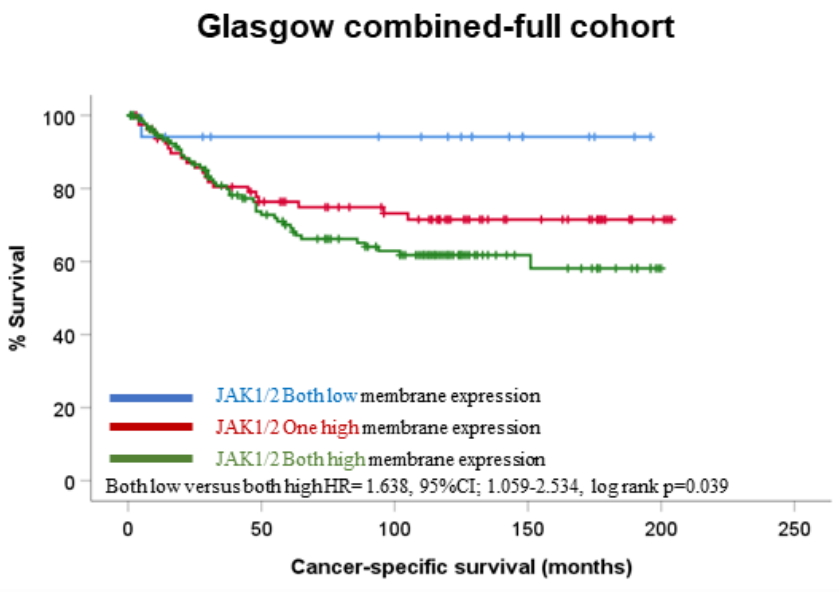


Figure 5.23 Membranous expression of JAK1 and/or JAK2 expression and CSS. *Kaplan Meier survival analysis of JAK1 and JAK2 membranous expression in patients from the Glasgow combined cohort. In the full cohort, the hazard ratio for combined JAK1 and JAK2 membranous expression score both high versus both low was 1.638, (95%CI; 1.059-2.534), log rank p=0.039.*

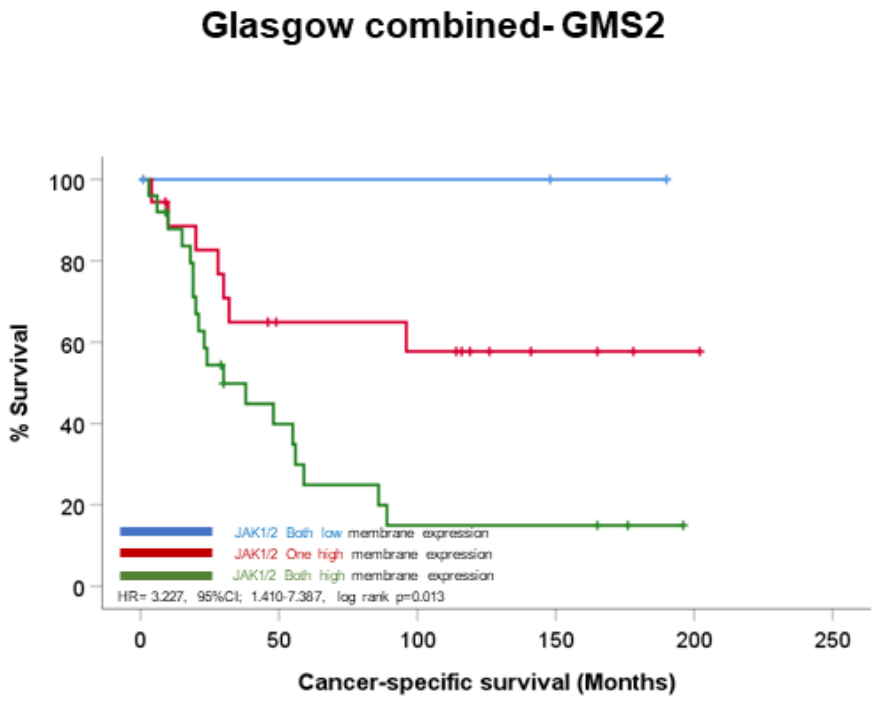


Figure 5.24 Membranous expression of JAK1 and/or JAK2 expression and CSS in GMS2 patients. *Kaplan Meier survival analysis of JAK1 and JAK2 membranous expression in patients classified as GMS2 from the Glasgow combined cohort. In GMS2 stromal-rich cases the hazard ratio for combined JAK1 and JAK2 membranous score was 3.227, (95%CI; 1.410-7.387), log rank p=0.013 for both high versus both low.*

5.2.4 Expression of JAK3 and clinical outcome

Cytoplasmic expression of Janus kinase-3 (JAK3) was assessed in patients from the Glasgow combined cohort by staining TMAs via IHC. Representative images of weak, moderate, and strong staining are shown in Figure 5.24 (A-C). Negative and positive controls colorectal tissue was stained at the same time as the cohort with representative images in Figure 5.24 (D-E). True negative, normal colon tissue, and true positive, liver tissue, are shown in Figure 5.24 (F-G). Antibody specificity was performed by staining sw620 cell pellets treated with 0.1%DMSO (H) or AT2983 JAK3 inhibitor (I) for JAK3 via IHC (Figure 5.24). Further validation was performed by western blotting to probe for JAK3 in HT29 cell lysates and PC3 cell lysates. The bands detected from HT29 cells were of a higher intensity than PC3 cells and tubulin remained constant for each sample at 52KDa (Figure 5.24) (J).

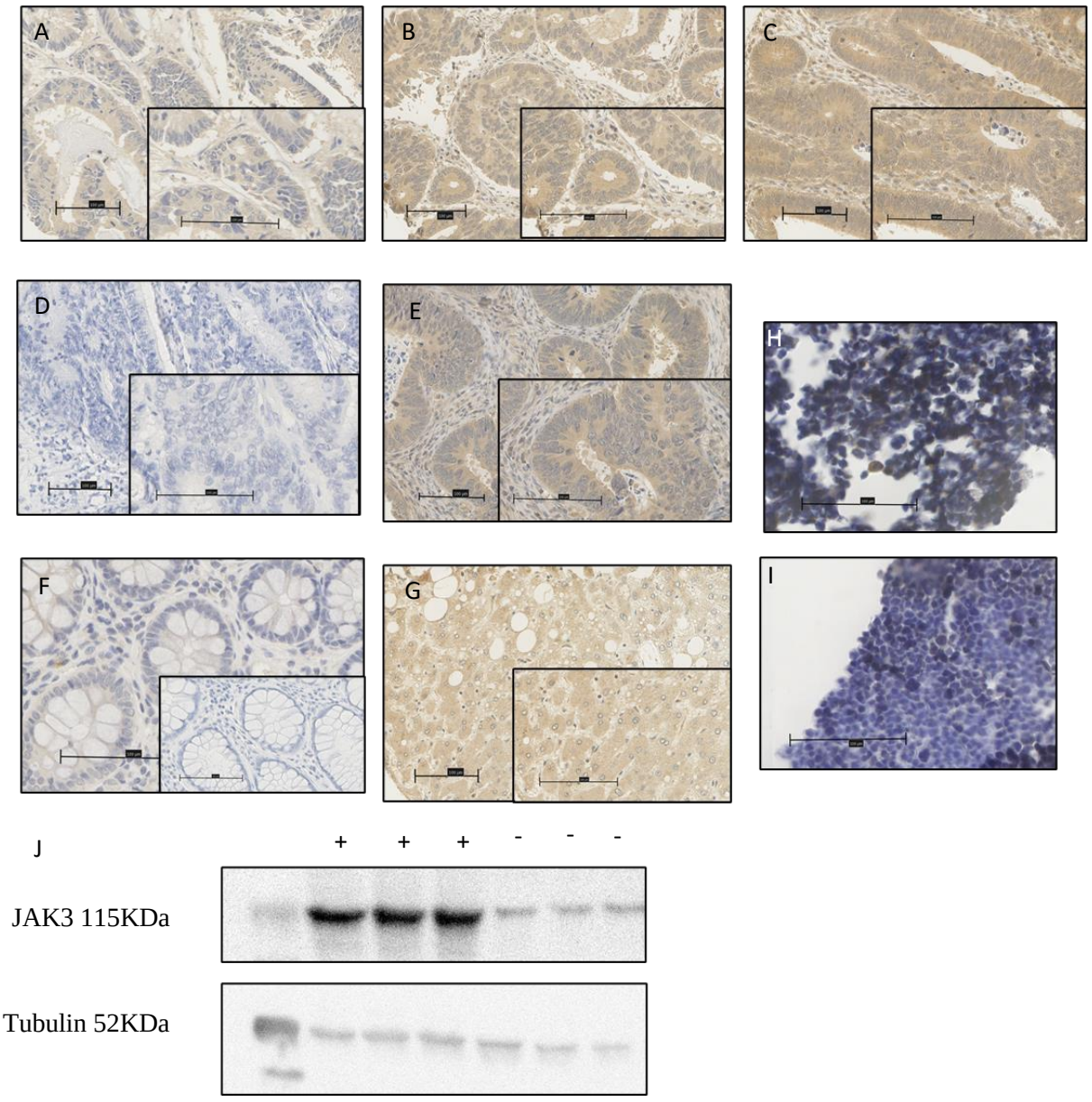


Figure 5.25 Representative images and antibody specificity for JAK3. *Images showing representative images of weak (A), moderate (B) and strong (C) cytoplasmic staining of JAK3 within tumour cells. Representative images of positive (D) and negative (E) colorectal tissue, and true positive (F) and true negative tissue (G). Image of western blot showing single band in triplicate for HT29 cell lysates (20μL per lane loaded) probed for JAK3 and weaker intensity of bands in triplicate for true negative PC3 lysates (20μL per lane loaded), with tubulin detected at 52KDa at a similar intensity across all 6 samples.*

Weighted histoscoring was performed manually by a single observer (KP) in 790 patients from the Glasgow combined array. When patients who received neoadjuvant therapy or those who died within 30 days of surgery were excluded, this left 726 patients included in analysis. Scores ranged from 0-260 with a mean of 154.58. A histogram was plotted to determine the distribution pattern of scores which was normal (Figure 5.25). Manual scores were validated using QuPath for 10% of cores by KP.

A correlation coefficient of 0.876 indicating a strong positive correlation between scores. A scatter plot was constructed to visualise this correlation (Figure 5.26). An intra-class correlation coefficient of 0.934 was determined indicating the scores were strongly related and not significantly different. A Bland Altman plot was constructed to visualise the bias between manual and digital JAK3 scores (Figure 5.27). Survminer and maxstat packages in R Studio were utilised to determine the optimal cut point for high and low expression based on CSS by log rank statistics. The optimum cut off determined was 208.33. Scatter and density plots were constricted to visualise this cut point (Figure 5.28). This resulted in 632 patients classified as low for expression of JAK3 and 94 patients as high for cytoplasmic JAK3 expression.

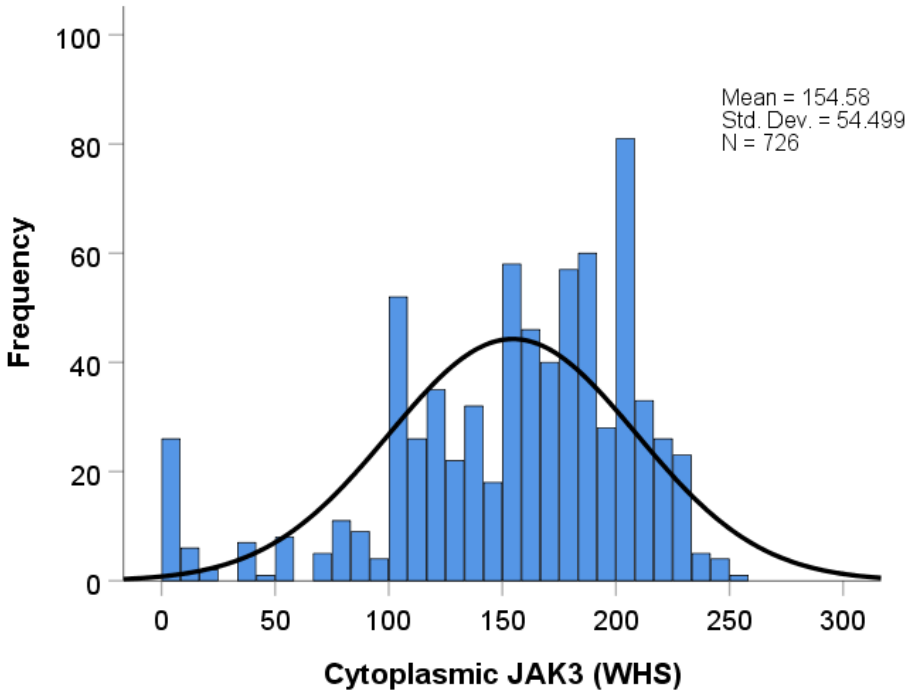


Figure 5.26 Distribution of JAK3 weighted histoscores. *Histogram showing JAK3 tumour cytoplasmic scores with normal distribution pattern (n=726). The mean score for JAK3 was 154.58 and scores ranged from 0-260.*

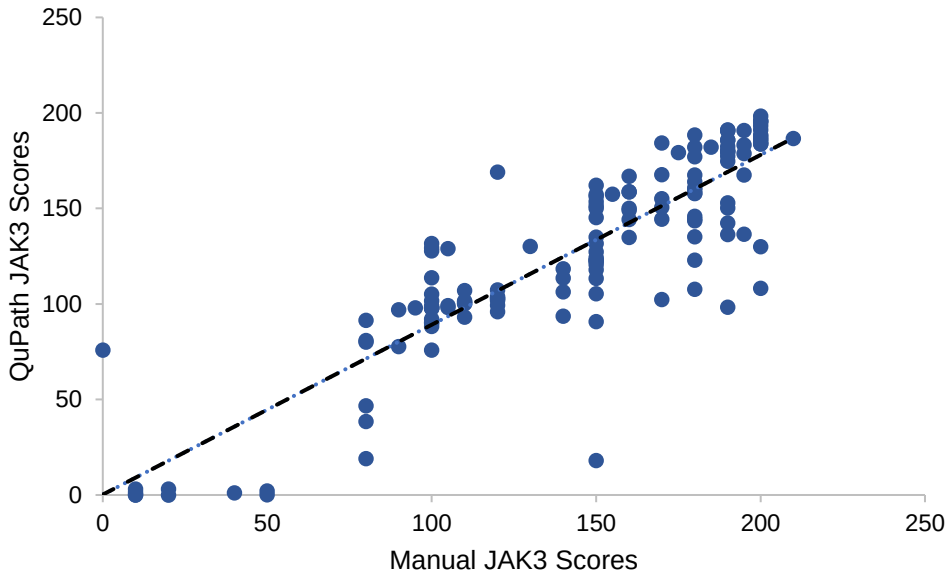


Figure 5.27 Correlation between Manual and QuPath cytoplasmic JAK3 scores. *Scatter plot showing the correlation between manual scores and digital scores for cytoplasmic JAK3 in 10% of patients from the Glasgow combined cohort that were co-scored, with a correlation coefficient of 0.876 obtained.*

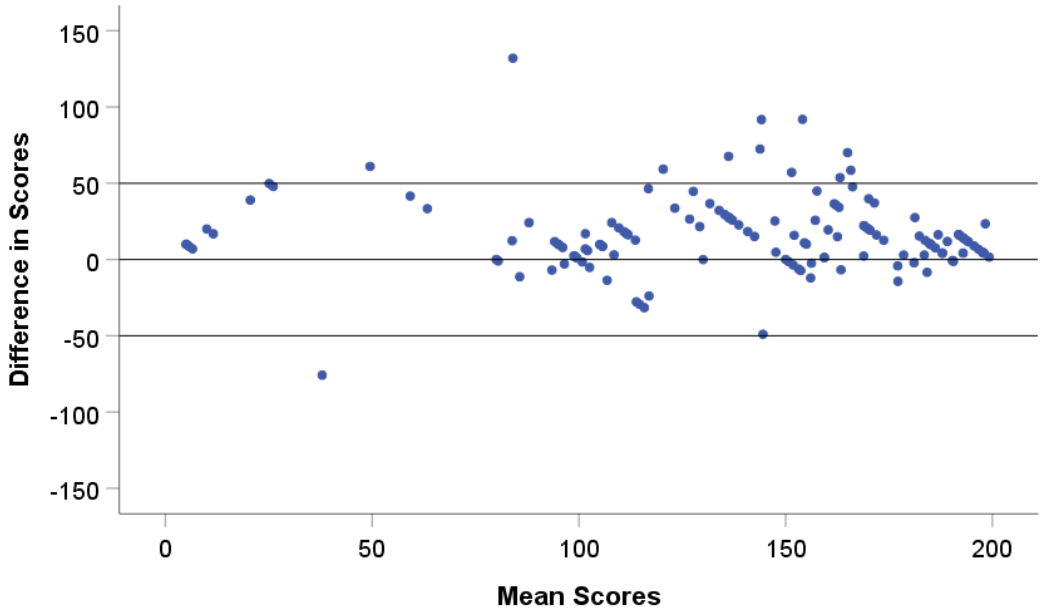


Figure 5.28 Difference between manual and digital JAK3 scores. *Bland Altman plot showing the similarity between manual and QuPath derived scores for cytoplasmic JAK3 in the 10% of patients from the Glasgow combined array co-scored for validation purposes.*

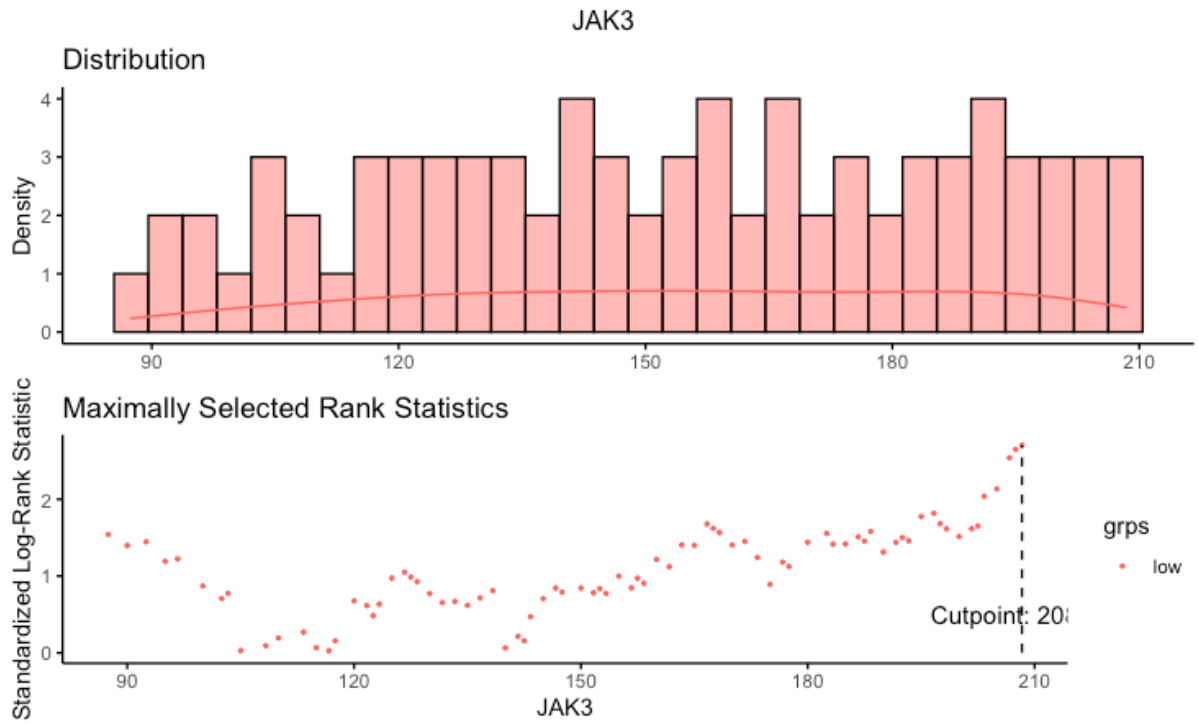


Figure 5.29 Defining optimal cut point for high and low expression of cytoplasmic JAK3 expression. *Density and scatter plot showing optimal cut off point for high and low expression of JAK3 groups based on cancer-specific survival. The optimal cut off point determined was 208.33, therefore patients with a weighted histoscore of ≤208.33 2 were considered as low expression and patients with a weighted histoscore of ≥208.33 were classed as high for JAK3 expression.*

Kaplan Meier survival analysis was performed to determine any association between cytoplasmic tumoural expression of JAK3 and CSS. In the full cohort low JAK3 expression was associated with reduced survival time (HR=0.503, 95%CI;0.302-0.837, log rank p=0.007) (Figure 5.29). Patients with low expression of JAK3 had a mean survival of 142.535 months versus 167.563 months for patients with high JAK3 expression. When patients were stratified by GMS histological subtypes, patients classed as intermediate GMS1 observed a survival benefit with high cytoplasmic JAK3 expression (HR=0.449, 95%CI; 0.227-0.888, log rank p=0.018) (Figure 5.30). GMS1 patients with low JAK3 expression had a mean survival time of 137.608 months compared to patients with high tumoural JAK3 expression at 166.930 months. In GMS0 immune patients and GMS2 stromal patients there was no association between JAK3 expression and CSS (Figure 5.30). JAK3 expression was associated with CSS in patients with MMR proficient but not patients with MMR deficient tumours (Figure 5.31). In pMMR cases, the mean survival time of patients with low JAK3 expression was 139.809 months versus 173.354 months in patients with high JAK3 expression. When patients were segregated by primary subsite, JAK3 expression was associated with CSS in left-sided colon cases (HR=0.451, 95%CI; 0.207-0.982, log rank p=0.039) (Figure 5.31). Patients with left-sided colon tumours and low JAK3 expression had a mean survival time of 142.225 months compared to high JAK3 expression at

171.912 months. There was no association between cytoplasmic JAK3 expression and outcome in patients with right-sided colonic or rectal tumours (Figure 5.32).

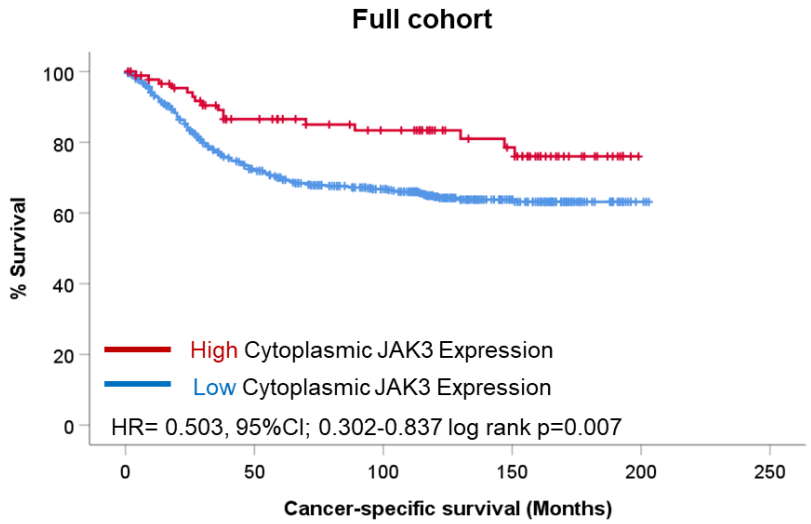


Figure 5.30 Figure 5.29 Cytoplasmic JAK3 expression and CSS in the full cohort. *Kaplan Meier survival analysis of cytoplasmic JAK3 expression in the Glasgow combined cohort. The hazard ratio for JAK3 expression in the full cohort was 0.503, (95%CI; 0.302-0.837), log rank p=0.007.*

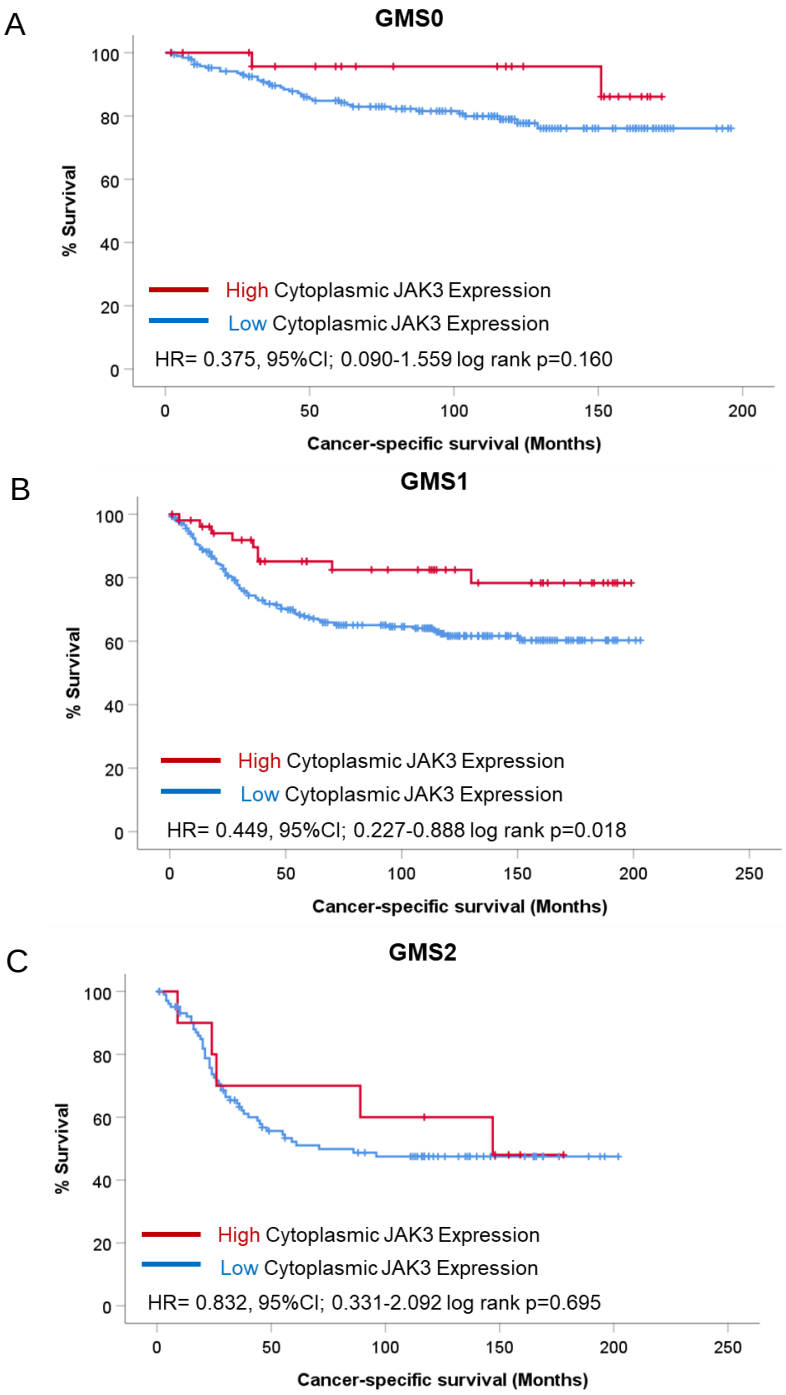


Figure 5.31 Cytoplasmic JAK3 expression and CSS with patients stratified by GMS. *Kaplan Meier survival analysis of cytoplasmic JAK3 expression in GMS0 patients (A), GMS1 patients (B) and GMS2 patients (C). The hazard ratio for patients with GMS0 graded tumours was 0.375, (95%CI; 0.0900-1.559), log rank p=0.160 and patients with GMS1 tumours saw a hazard ratio of 0.449, (95%CI; 0.227-0.888), log rank =0.018. In patients with GMS2 stromally dense tumours the hazard ratio for JAK3 was 0.832, (95%CI; 0.331-2.092), log rank p=0.695.*

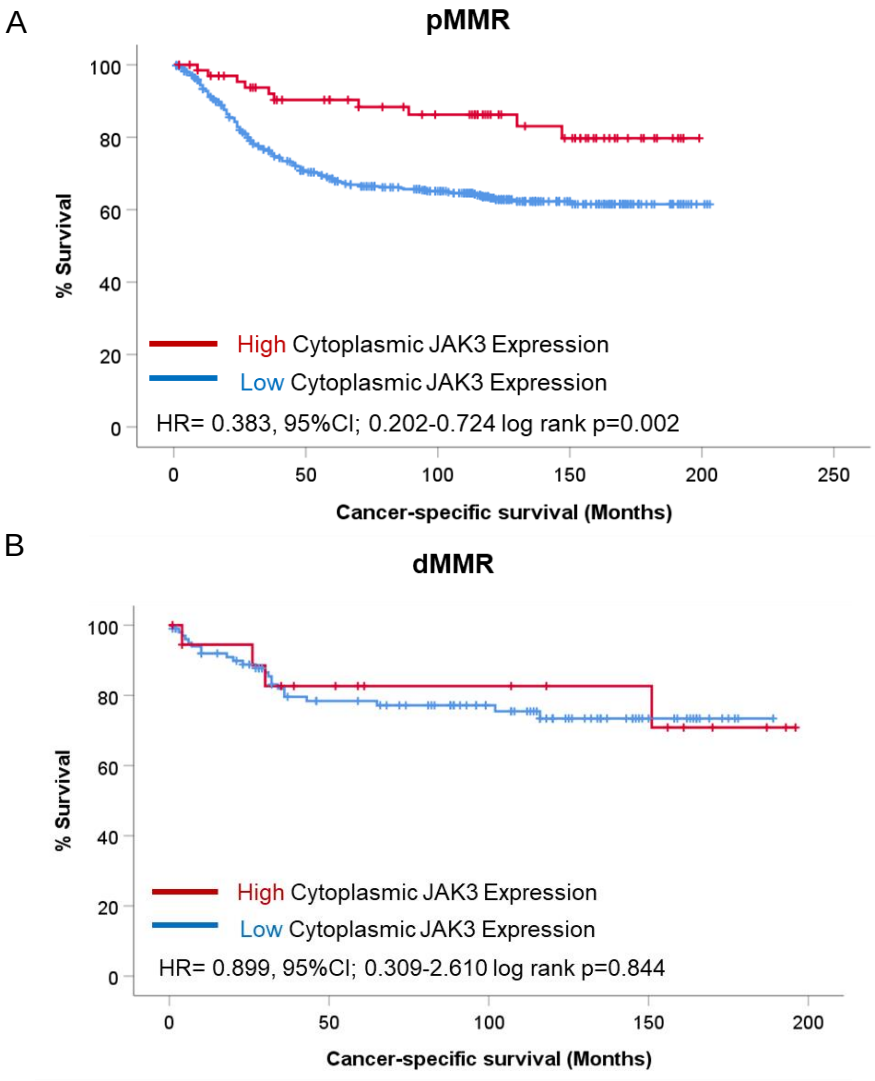


Figure 5.32 Cytoplasmic JAK3 expression and CSS with patients stratified by MMR status. *Kaplan Meier survival analysis of cytoplasmic JAK3 expression in patients with MMR proficient tumours (A) and MMR deficient tumours (B). In MMR proficient cases the hazard ratio for JAK3 expression was0.383, (95%CI; 0.202-0.724), log rank p=0.002 and in MMR deficient cases the hazard ratio was 0.899, (95%CI; 0.309-2.610), log rank p=0.844.*

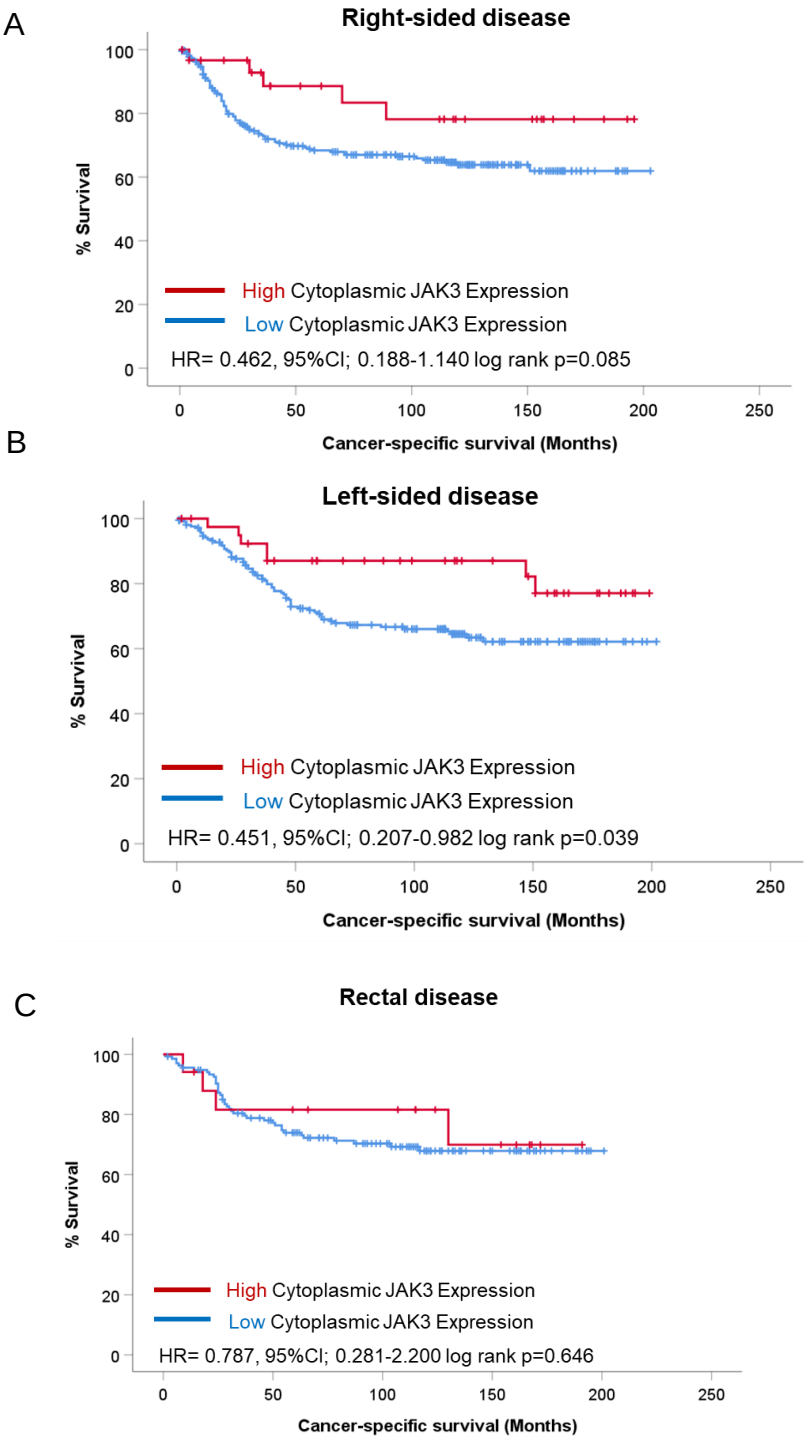


Figure 5.33 Cytoplasmic JAK3 expression and CSS with patients stratified by tumour subsite. *Kaplan Meier survival analysis of cytoplasmic JAK3 expression in patients with right-sided tumours (A), left-sided tumours (B) and rectal tumours (C). In patients with right-sided tumours the hazard ratio for JAK3 expression was 0.462, (95%CI; 0.188-1.140), log rank p=0.085 and in left-sided colonic tumours the hazard ratio was 0.451, (95%CI; 0.207-0.982), log rank p=0.039. In rectal cases the hazard ratio for JAK3 expression was 0.787, (95%CI; 0.281-2.200), log rank p=0.646.*

Chi-squared analyses were performed to determine association between JAK3 expression and clinicopathological characteristics. High cytoplasmic JAK3 was significantly associated with increased vascular invasion (p=0.009) and lower N stage (p=0.008) (Table 5.3).

Clinical Factor	JAK3 Expression		p
	Low (n=632)	High (n=94)	
Age			
<65	195 (30.9)	26 (27.7)	0.309
>65	437 (69.1)	68 (72.3)	
Sex			
Female	305 (48.3)	51 (54.3)	0.165
Male	327 (51.7)	43 (45.7)	
T Stage			
I	23 (3.6)	4 (4.3)	0.051
II	70 (11.1)	20 (21.3)	
III	343 (54.3)	48 (51.1)	
IV	196 (31.0)	22 (23.4)	
N Stage			
0	370 (58.8)	70 (75.3)	**0.008**
I	185 (29.4)	16 (17.2)	
II	74 (11.8)	7 (7.5)	
Tumour subsite			
Right-sided colon	279 (44.5)	33 (35.5)	0.102
Left-sided colon	212 (33.8)	42 (45.2)	
Rectum	136 (21.7)	18 (19.4)	
GMS			
0	198 (32.7)	29 (31.2)	0.141
1	299 (49.4)	54 (58.1)	
2	108 (17.9)	10 (10.8)	
Phenotypic Subtype			
1	197 (32.6)	29 (31.9)	0.323
2	170 (28.1)	28 (30.8)	
3	130 (21.5)	24 (26.4)	
4	108 (17.9)	10 (11.0)	
mGPS			
0	263 (50.9)	34 (57.6)	0.129
1	160 (30.9)	20 (33.9)	
2	94 (18.2)	5 (8.5)	
MMR status			
pMMR	520 (83.3)	72 (79.1)	0.197
dMMR	104 (16.7)	19 (20.9)	
Tumour differentiation			
0	565 (89.4)	89 (90.1)	0.072
1	67 (10.6)	5 (5.3)	
Marginal involvement			
0	597 (94.5)	420 (92.9)	0.414
1	11 (4.2)	32 (7.1)	
Vascular invasion			
0	418 (66.1)	74 (78.7)	**0.009**
1	214 (33.9)	20 (21.3)	

Table 5.3 JAK3 Expression and Clinical Features. *Table of Chi-squared associations between JAK3 expression and clinicopathological prognostic features including age, sex, T stage, N stage, tumour subsite, GMS, phenotypic subtype, MMR status, tumour differentiation mGPS, tumour differentiation, marginal involvement and venous invasion.*

5.2.5 Expression of TYK2 and clinical outcome.

Expression of TYK2 was assessed in tumour cytoplasm of patients from the Glasgow combined cohort tissue microarray. Representative images of weak, moderate, and strong staining are shown in Figure 5.33 (A-C). Positive and negative control colorectal tissue were stained alongside the cohort and representative images are shown in Figure 5.33 (D-E). True positive tissue (liver and true negative tissue (breast) were stained during the same run as the TMA cores and representative images are shown in Figure 5.33. Antibody specificity was performed via staining of cell pellets treated with 0.01% DMSO vehicle control or TYK2 inhibitor AT9283, representative images are shown in Figure 5.33 (H). The inhibitor-treated pellets showed significantly reduced TYK2 expression compared to the vehicle control-treated SW620 cells.

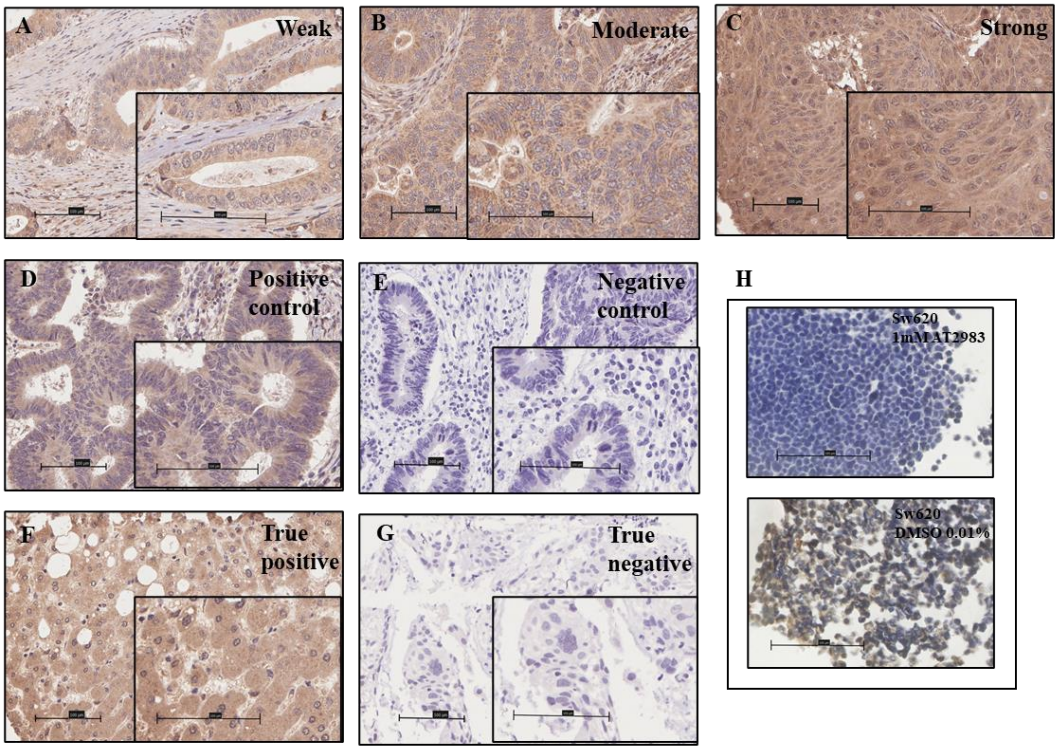


Figure 5.34 Representative images and antibody specificity for TYK2. *Images showing representative images of weak (A), moderate (B) and strong (C) cytoplasmic staining of TYK2 within tumour cells. Representative images of positive (D) and negative (E) colorectal tissue, and true positive (F) and true negative tissue (G). Cell pellets made from SW620 cell lines treated with either AT2983 TYK2 inhibitor or 0.01% DMSO vehicle control, stained via IHC for TYK2 to show higher expression in vehicle treated cells compared to TYK2 inhibitor treated cells (H).*

Expression of TYK2 was semi-quantitively measured using manual weighted histoscoring by a single observer (KP). Scores were obtained for 796 patients from the 1030 patient cohort due to missing and/or damaged TMA cores. After excluding patients who died within 30 days of surgery or those

who received neoadjuvant chemotherapy 716 patients were included in further analysis. Weighted histoscores ranged from 0 to 300 with a mean score of 188.70. A histogram was plotted to visualise the range of scores and data were relatively normally distributed (Figure 5.34). Due to scores being devised by a single observer (KP), scoring was validated in 10% of tumour cores utilising QuPath digital pathology software by KP. A correlation coefficient of 0.890 indicating a strong positive correlation was obtained between scores and data were visualised in a scatter plot (Figure 5.35). A Bland Altman plot was constructed to visualise the bias between scores and to prove scores were not statistically different (Figure 5.36). An intra-class correlation coefficient of 0.946 indicated the scores were strongly positively correlated and not different from each other. Cut offs were determined using maxstat and survminer in R studio via log rank statistics based on CSS. This method yielded an optimum cut off score for high and low expression groups of 182.5. The cut point was visualised by constructing scatter and density plots (Figure 5.37). Subsequently there were 264 patients classified as low for TYK2 expression and 452 patients as high for expression of cytoplasmic TYK2.

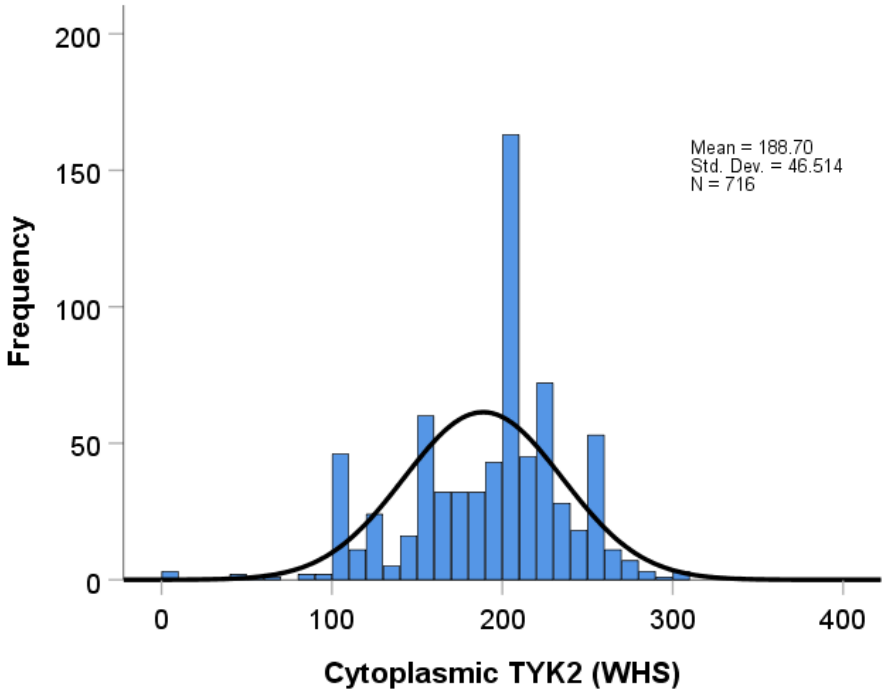


Figure 5.35 Distribution of weighted histoscores for TYK2. *Histogram showing the range of scores obtained for cytoplasmic tumour TYK2 expression and distribution pattern of the data (n=716). The mean score for TYK2 expression was 188.70 and the scores ranged from 0-300.*

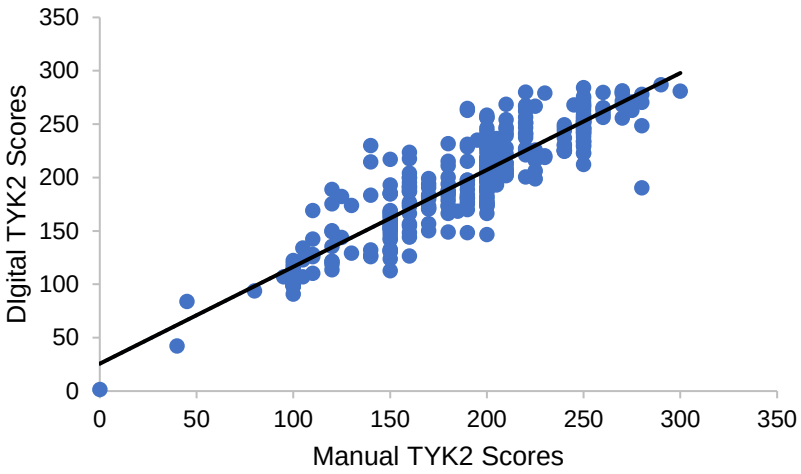


Figure 5.36 Correlation between manual and digital weighted histoscoring of TYK2. *Scatter plot showing correlation between tumour cytoplasm weighted histoscores for TYK2 for the 10% of cases co-scored using QuPath digital pathology. A correlation coefficient of 0.890 was obtained.*

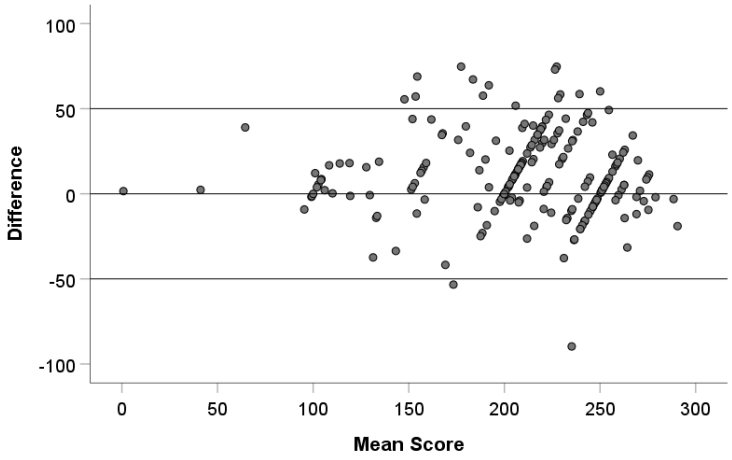


Figure 5.37 Validation of TYK2 manual scoring. *Bland Altmann plot showing difference between manual and digital QuPath scores for TYK2 in tumour cytoplasm for the 10% of co-scored cases in the Glasgow combined array.*

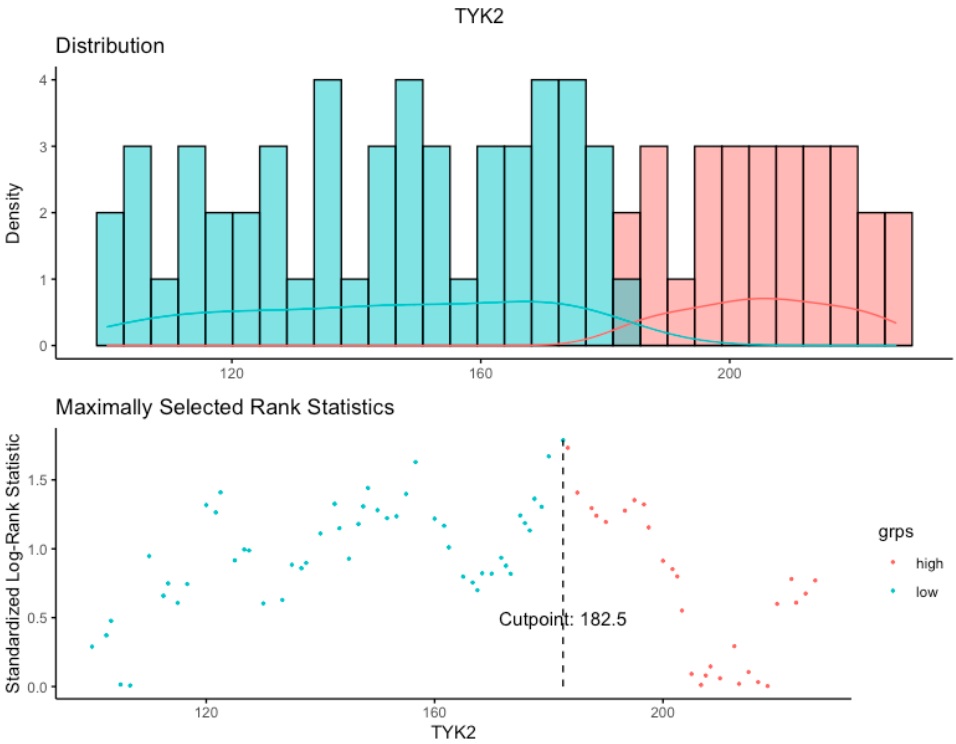


Figure 5.38 Defining cut off point for TYK2 expression high and low groups. *Density plot and scatter plot for visualisation of optimal cut off point for high and low expression of tumour cytoplasmic TYK2. The optimal cut point determined was 182.5, therefore patients with a weighted histoscore of ≥182.5 were considered high for TYK2 expression and patients with a weighted histoscore of ≤182.5 were classed as low for cytoplasmic tumoural TYK2 expression.*

Kaplan Meier plot survival analysis was used to determine any association between cytoplasmic tumour TYK2 expression and CSS. In the full cohort TYK2 expression was not significantly associated with CSS (p=0.094) (Figure 5.38). Patients were subdivided into GMS classification groups and Kaplan Meier survival analysis was performed again. No significant association between cytoplasmic expression of TYK2 and CSS was observed in any subtype (Figure 5.39). In the GMS2, although not a significant relationship, patients with high cytoplasmic TYK2 trended towards worse prognosis (Figure 5.39). No significant association between cytoplasmic TYK2 expression and CSS was found in either MMR proficient or MMR deficient disease (Figure 5.40). When TYK2 was analysed in the context of tumour subsite, a significant association was observed between CSS and cytoplasmic expression in rectal tumours, with high expression associated with worse outcomes (HR= 1.344, 95%CI; 0.742-2.434, log rank p=0.050) (Figure 5.41). Patients with rectal tumours and high TYK2 expression had a mean survival of 139.512 months compared to low TYK2 expression at 169.759 months. No significant relationship between CSS and cytoplasmic TYK2 expression as seen in right-sided or left-sided colonic disease (Figure 5.41).

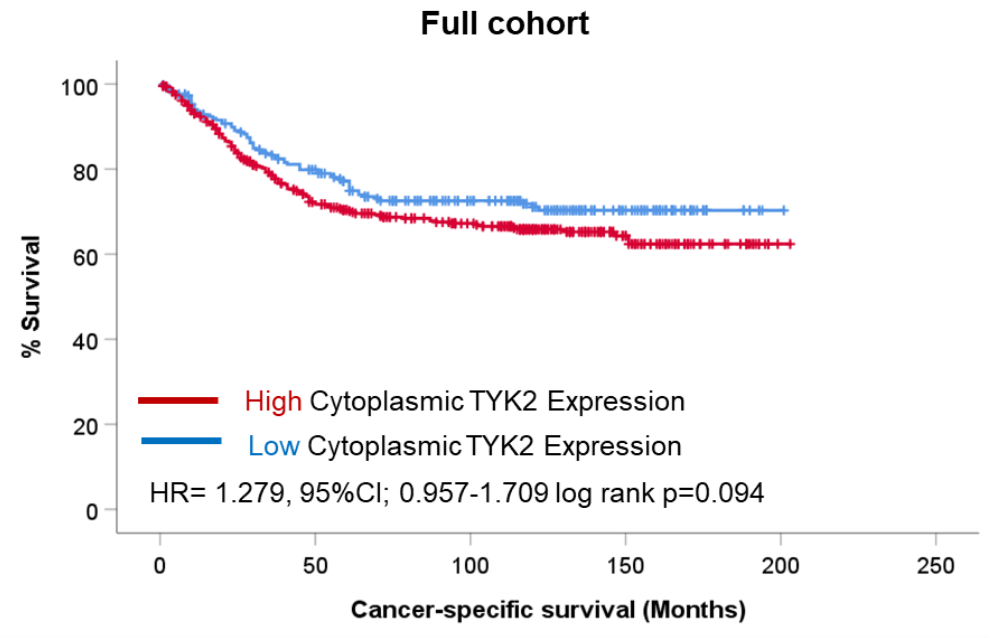


Figure 5.39 Cytoplasmic expression of TYK2 expression and CSS. *Kaplan Meier survival analysis of cytoplasmic TYK2 expression in patients from the Glasgow combined cohort. In the full cohort the hazard ratio for TYK2 expression was 1.279, (95%CI; 0.957-1.709), log rank p=0.094.*

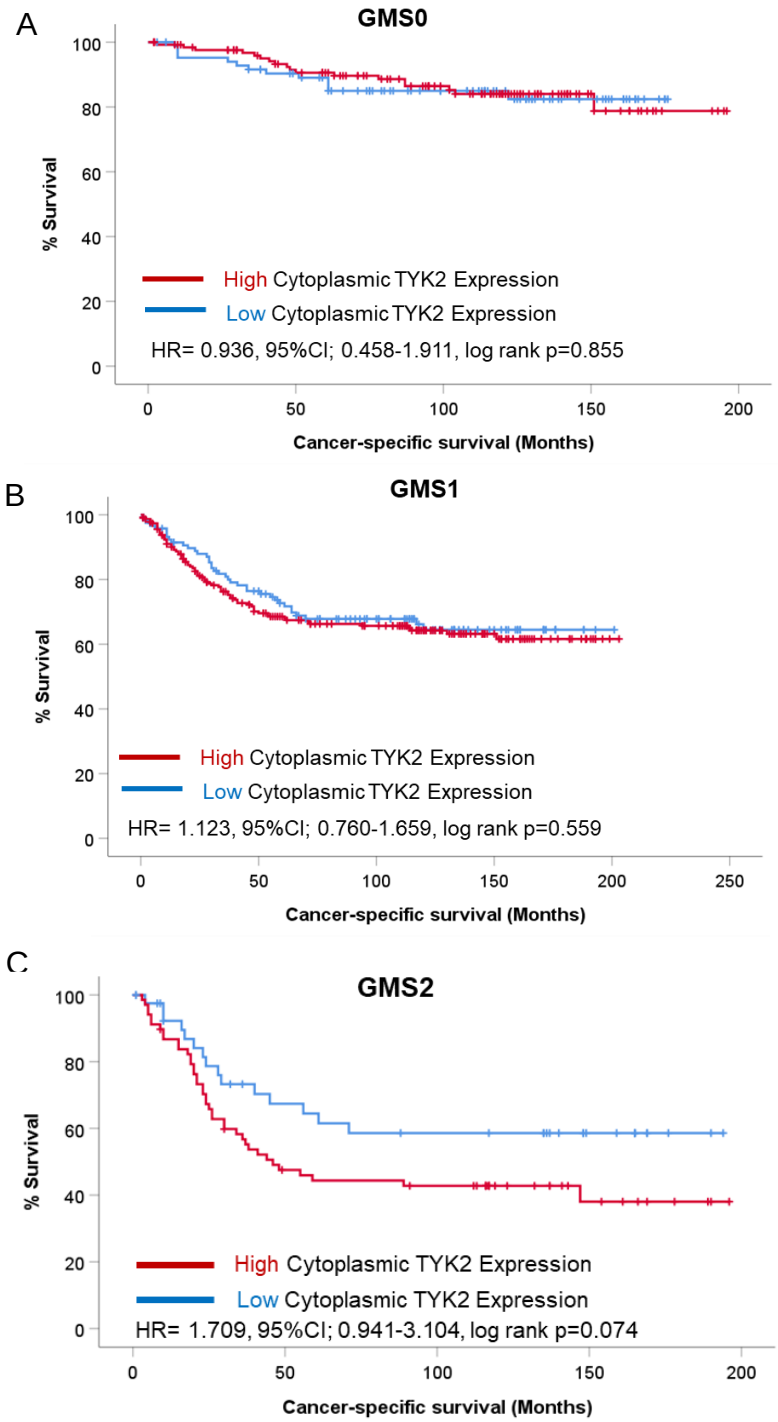


Figure 5.40 Cytoplasmic expression of TYK2 expression and CSS relative to GMS classification. *Kaplan Meier survival analysis of cytoplasmic TYK2 expression in patients from the Glasgow combined cohort classified as GM0 (A), GMS1 (B), and GMS2 (C). For patients classified as GMS0 immune the hazard ratio for TYK2 expression was 0.936, (95%CI; 0.458-1.911), log rank p=0.855 and for patients with GMS1 tumours the hazard ratio was 1.123, (95%CI; 0.760-1.659), log rank p=0.559. In GMS2 stromal-rich cases the hazard ratio for TYK2 expression was 1.709, (95%CI; 0.941-3.104), log rank p= 0.074.*

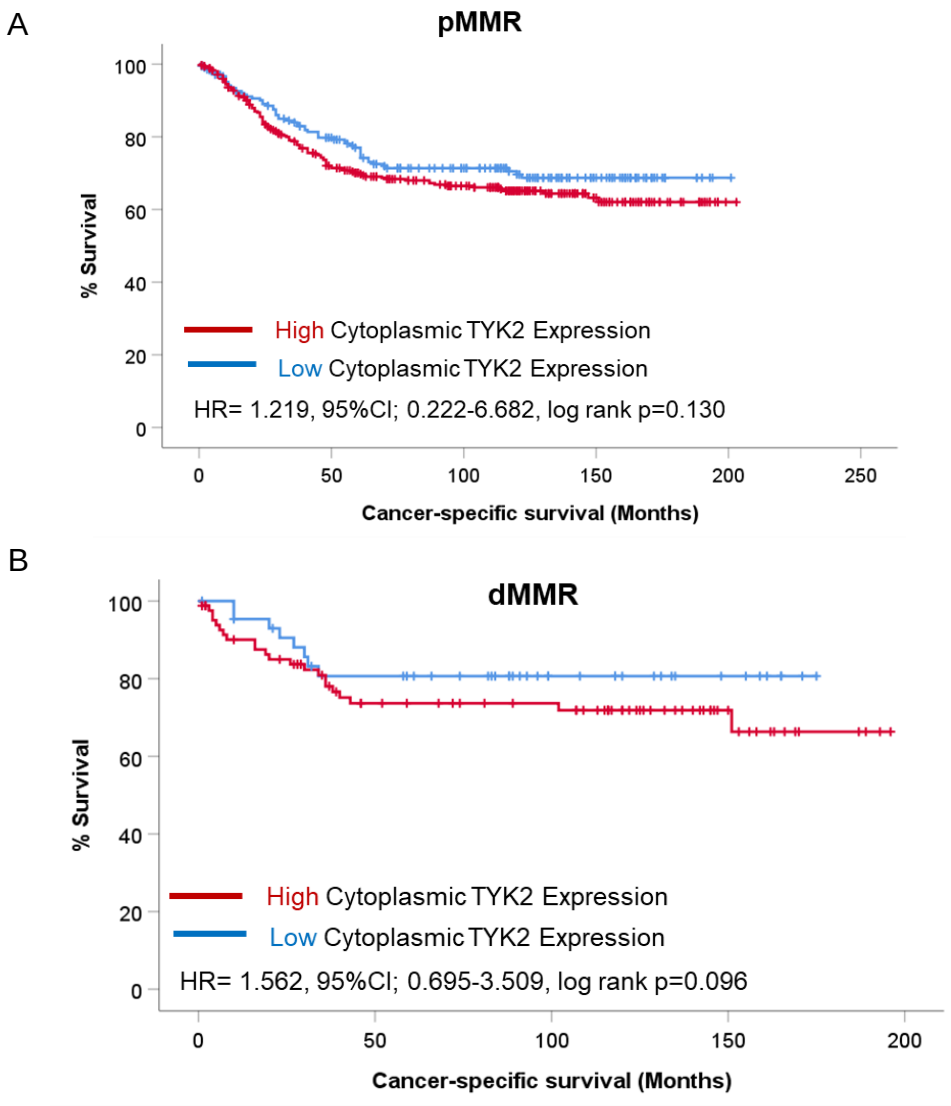


Figure 5.41 Cytoplasmic expression of TYK2 expression and CSS relative to MMR status. *Kaplan Meier survival analysis of cytoplasmic TYK2 expression in patients from the Glasgow combined cohort with MMR proficient tumours(A) and MMR deficient tumours (B). In MMR proficient cases the hazard ratio for TYK2 expression was 1.219, (95%CI; 0.222-6.682), log rank =0.130 and in MMR deficient cases the hazard ratio was 1.562, (95%CI; 0.695-3.509), log rank p=0.096.*

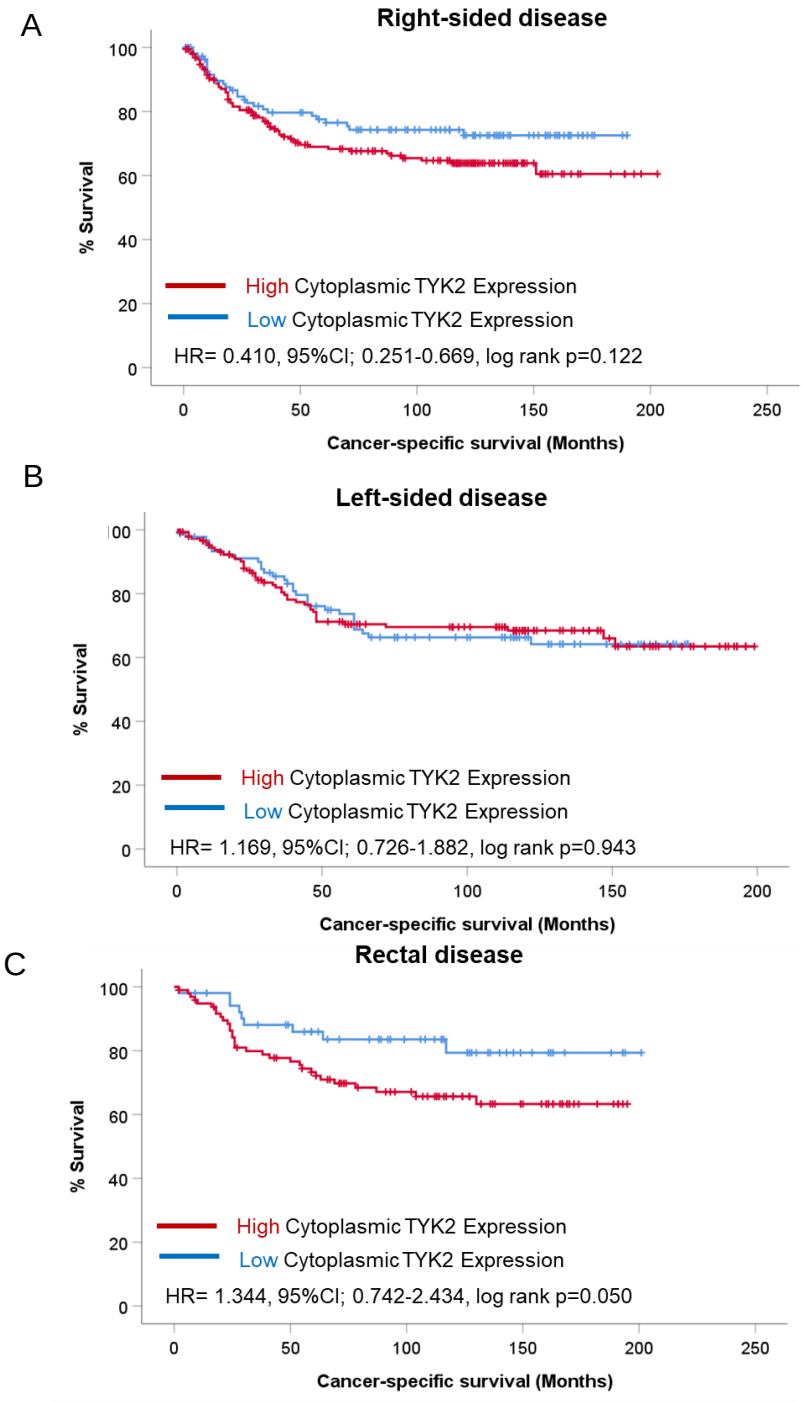


Figure 5.42 Cytoplasmic expression of TYK2 expression and CSS relative to tumour subsite. *Kaplan Meier survival analysis of cytoplasmic TYK2 expression in patients from the Glasgow combined cohort with right-sided disease (A), left-sided disease (B), and rectal disease (C). In patients with right-sided colonic tumours the hazard ratio for TYK2 expression was 0.410, (95%CI; 0.251-0.669), log rank p=0.122 and in left-sided colon tumours the hazard ratio was 1.169, (95%CI; 0.726-1.882), log rank p=0.943. In rectal cases the hazard ratio for TYK2 expression was 1.344, (95%CI; 0.742-2.434) log rank p=0.050.*

Next, TYK2 expression was assessed for associations with clinicopathological characteristics via Chi-squared analyses. High TYK2 expression was associated with the canonical phenotypic subtype ($p<0.001$) and systemic inflammation measured via modified Glasgow prognostic score (mGPS) ($p=0.034$) (Table 5.4).

.

Clinical Factor	TYK2 Expression		p
	Low (n=264)	High (n=452)	
Age			
<65	82 (31.1)	139 (30.8)	0.498
>65	182 (68.9)	313 (69.2)	
Sex			
Female	124 (47.0)	222 (49.1)	0.317
Male	140 (53.0)	230 (50.9)	
T Stage			
I	8 (3.0)	21 (4.6)	0.462
II	30 (11.4)	61 (13.5)	
III	151 (57.2)	236 (52.2)	
IV	75 (28.4)	134 (29.6)	
N Stage			
0	166 (62.9)	286 (63.7)	0.976
I	66 (25.0)	110 (24.5)	
II	32 (12.1)	53 (11.8)	
III			
Tumour subsite			
Right-sided colon	117 (44.7)	200 (44.6)	0.880
Left-sided colon	92 (35.1)	151 (33.7)	
Rectum	53 (20.2)	97 (21.7)	
GMS			
0	90 (35.4)	134 (30.7)	0.396
1	122 (48.0)	231 (52.9)	
2	42 (16.5)	72 (16.5)	
Phenotypic Subtype			
1	90 (35.6)	133 (30.6)	**<0.001**
2	55 (21.7)	155 (35.6)	
3	66 (26.1)	75 (17.2)	
4	42 (16.6)	72 (16.6)	
mGPS			
0	111 (48.3)	185 (54.1)	**0.034**
1	85 (37.0)	92 (26.9)	
2	34 (14.8)	65 (19.0)	
MMR status			
pMMR	214 (82.3)	363 (81.4)	0.422
dMMR	46 (17.7)	83 (18.6)	
Tumour differentiation			
0	235 (89.0)	405 (89.6)	0.449
1	29 (11.0)	47 (10.4)	
Marginal involvement			
0	253 (95.8)	420 (92.9)	0.075
1	11 (4.2)	32 (7.1)	
Vascular invasion			
0	182 (68.9)	290 (64.2)	0.111
1	82 (31.1)	162 (35.8)	

Table 5.4 TYK2 Expression and Clinical Features. *Table of Chi-squared associations between cytoplasmic tumoural TYK2 expression and clinicopathological prognostic features including age, sex, T stage, N stage, tumour subsite, GMS, phenotypic subtype, mGPS, MMR status, tumour differentiation, marginal involvement and vascular invasion.*

5.2.5 JAK2 and TYK2 combined score and clinical outcome

Given the trend towards association with high JAK2 and TYK2 expression and reduced CSS, scores for these proteins were combined to form 3 groups, high for both, high for 1 marker, and low for both. This resulted in 54 patients classified as low for both, 142 and high for 1 marker and 84 as high for both markers. In the full cohort the combined JAK2/TY2 score was associated with CSS (HR=1.203, 95%CI; 0.881-1.644, log rank p=0.016) (Figure 5.42). Patients with low expression of both proteins had a mean survival of 152.943 months, compared to 148.745 months for both high and 135.028 for high for 1 marker. This relationship was somewhat potentiated in patients with GMS2 stromal rich tumours (HR=1.801, 95%CI; 1.077-3.012, log rank p=0.022) (Figure 5.43). Patients with low expression of both markers had a mean survival time of 157.500 months compared to patients high for both markers at 63.107 months (p=0.010), and patients high for 1 marker at 70.406 months.

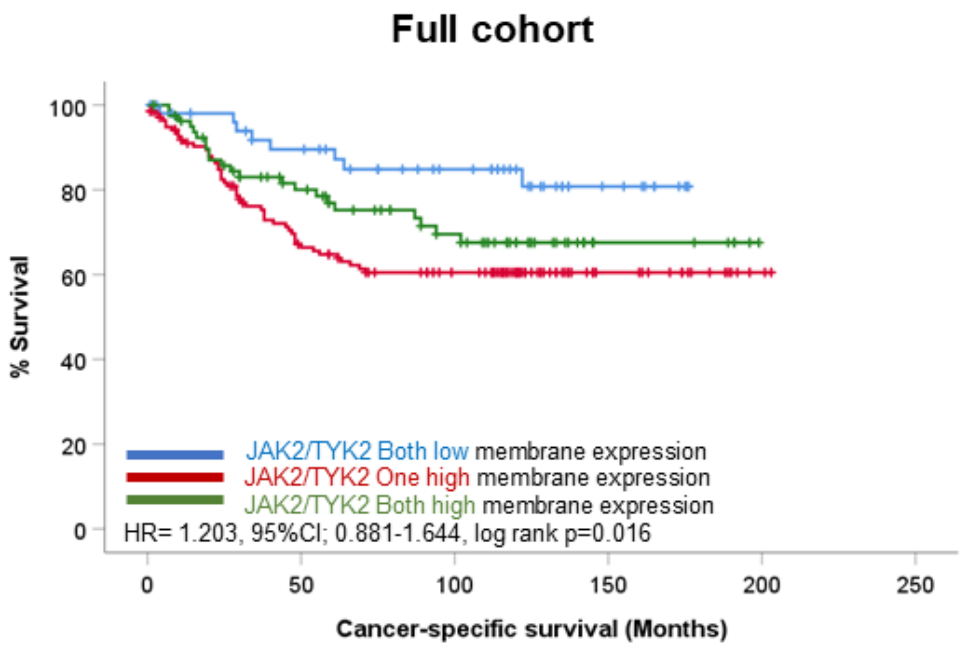


Figure 5.43 Combined cytoplasmic expression of JAK2 and TYK2 expression and CSS. *Kaplan Meier survival analysis of combined cytoplasmic JAK3 and TYK2 expression in patients from the Glasgow combined cohort. In the full cohort the hazard ratio for combined JAK2/TYK2 score was 1.203, (95%CI; 0.881-1.644), log rank p=0.016.*

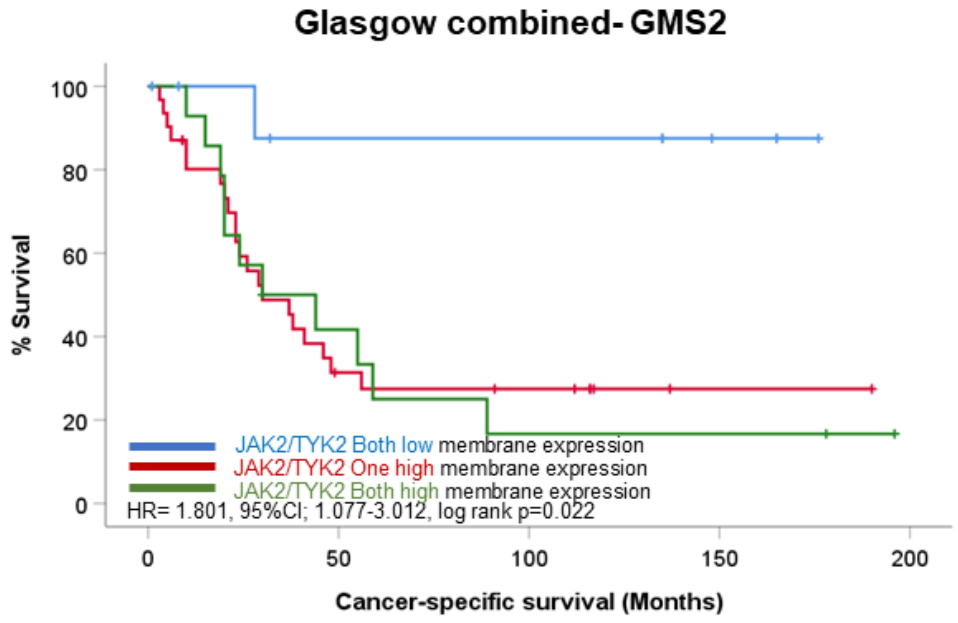


Figure 5.44 Combined cytoplasmic expression of JAK2 and TYK2 expression and CSS in GMS2 patients. *Kaplan Meier survival analysis of combined cytoplasmic JAK3 and TYK2 expression in patients with stromally-dense GMS2 tumours from the Glasgow combined cohort. The hazard ratio for combined JAK2/TYK2 score in GMS2 patients from the combined array was 1.801, (95%CI; 1.077-3.012), log rank p=0.022.*

5.3 Discussion

Over the past decade, our understanding of the underlying mechanisms leading to CRC development and progression has greatly increased. It is now widely accepted that tumourigenesis is characterised by the dysregulation of cellular signalling pathways leading to promotion of all seven hallmarks of cancer (126). One such cancer-associated pathway is IL6/JAK/STAT3 signal transduction. The prognostic roles of IL6 and STAT3 are relatively well established in many solid tumour types, with both pathway members linked to poor outcomes (127). Akin to other cytokine receptors, IL6R does not contain an intracellular tyrosine kinase domain and signalling therefore relies on activation of JAK family members to transduce the signal (128). Expression of these intermediate IL6/JAK/STAT3 pathway components; JAK1, JAK2, TYK2 and to some extent JAK3 are less well defined in the cancer setting.

Data suggests JAK2 and TYK2 are generally associated with unfavourable clinical outcomes in cancer. At the genome level, mutations in JAK2 and TYK2 genes are associated with reduced survival

in both colon and rectal cancer (129). In CRC cell lines and spheroids, JAK2 expression is associated with reduced apoptosis, enhanced clonogenic potential and promotion of resistance to radiotherapy (130). TYK2 hyperactivation is associated with tumour cell invasiveness in prostate cancer (131). TYK2 has also been shown to interact with a variety of other oncogenic signalling pathways including RAF/ERK, MAPK and PI3K/AKT/mTOR (132). Data from this chapter showed high JAK2 expression was associated with reduced survival times in GMS2 patients (HR=1.881, 95%CI;0.992-3.565, p=0.048). Although TYK2 expression alone showed no prognostic capability, when TYK2 scores were combined with JAK2, high expression was associated with poorer outcomes in the full cohort, and this relationship was potentiated in GMS2 tumour cases (HR=1.801, 95%CI; 1.077-3.012, log rank p=0.022). This indicates further evidence that targeting IL6/JAK/STAT3 may be useful in patients with highly stromal tumours.

JAK3 is generally associated with haematologic cell signalling and expression has not been frequently reported in solid tumours (133). JAK3 expression, at this time, has not been extensively investigated in large scale tissue-based studies. However, JAK3 is overexpressed in colorectal cell lines and inhibiting JAK3 in CRC cell lines does reduce STAT3 activation (134). Interestingly, the data discussed in this chapter, suggested high tumour expression of JAK3 was associated with increased survival time (HR=0.503, 95%CI;0.302-0.837, log rank p=0.007). This observation was most evident in GMS1 intermediate tumours, MMR proficient and right and left-sided colon cases. Further research is needed to confirm these findings in a validation cohort.

The role of JAK1 in CRC is also not well-defined in the literature and there are conflicting data about whether expression is favourable or unfavourable for prognosis. In breast cancer low JAK1 expression is associated with reduced survival, and high JAK1 expression associates with increased anti-tumour infiltrating immune cells (135). Similar results were observed for the Glasgow combined cohort in this chapter with low expression of JAK1 predictive of reduced CSS. However, in non-small lung cancer expression of p-JAK1 associates with poor prognosis and unfavourable clinical characteristics (136). In this chapter, only total JAK1/JAK2/JAK3/TYK2 were stained for, which might not provide a clear picture of the role each protein plays in CRC. Activation may be of more importance than total protein, therefore future work should include staining the cohort TMAs for the phosphorylation sites of each JAK family member.

Positive staining for each of the JAK family proteins was observed in the tumour cytoplasm. JAK1 and JAK2 staining was also seen as a membranous pattern on some patient's tumour cells. When modified histoscoring as 'absent' or 'present' for membrane expression was performed, patients with membrane JAK1 and JAK2 had significantly worse outcomes than those with no staining, particularly in the stromal subtype (HR=3.227, 95%CI; 1.410-7.387, log rank p=0.013). Membranous expression could be a surrogate marker of activation as this could indicate the JAKs were associating with the

cytoplasmic domains of the cytokine receptor (128). No previous studies have shown a role for membrane expression of JAK1/2 in stratifying patient survival outcomes any solid tumour type, however membranous staining patterns have been observed in pancreatic ductal adenocarcinoma (137). Multiplex IF or IHC could be utilised to confirm this and assess colocalization of IL6R and JAK1 and/or JAK2.

Future work could also include investigating which of the JAK family members if predominant in activating STAT3 in the CRC setting. Each of the proteins is capable of transducing signals from IL6R to STAT3, however there is limited evidence as to which is the most common component of the pathway in tumour cells. JAK1, JAK2, JAK3 and TYK2 are also capable of activating other STAT proteins such as STAT1 and STAT5 (138). Mechanistic studies looking specifically at STAT3 activation are needed to improve our understanding of each JAK family members role in CRC.

In conclusion, expression of oncokinases JAK2 and TYK2 associated with poor outcome in the Glasgow combined retrospective CRC cohort, particularly in patients with a low inflammatory infiltrate. Conversely, JAK1 and JAK3 expression within tumour cytoplasm was associated with good prognosis. In order to fully understand the role of these proteins in CRC, the activation sites of each marker should be stained for and assessed for associations with clinical characteristics and outcome. The majority of data from this chapter indicate that patient with stromal-rich GMS2 tumours represent the target CRC patient population for therapeutic inhibition of IL6/JAK/STAT3 signalling.

Chapter 6: Expression of STAT3, pSTAT3^{tyr705}, and pSTAT3^{ser727} in colorectal cancer clinical specimens

6.1 Introduction

Signal-transducer and activator of transcription- 3 (STAT3) is a master regulator of transcription involved in promoting many of the hallmarks of cancer. Constitutive activation of STAT3 is present in colorectal, prostate, gastric, pancreatic cancers (139-141). STAT3 is classically activated via IL6 binding IL6R, causing phosphorylation of JAK1/2/3/TYK2, leading to STAT3 dimerization and translocation to the nucleus. STAT3 can be phosphorylated at two sites, canonically at tyrosine 705, and subsequently at serine 727. Activation is associated with invasion, migration, metastases and angiogenesis (142). Previous studies have shown that expression of pSTAT3 in tumour cells is associated with poor prognosis and unfavourable characteristics in CRC including Duke's stage, invasion, and overall survival (74, 143). There is limited evidence in the literature investigating the importance of STAT3 expression relative to tumour phenotype, however upregulation is most commonly associated with the immune subtype, CMS1 (144).

The aims of this chapter were to investigate protein expression of STAT3, pSTAT3^{tyr705} and pSTAT3^{ser727} in the Glasgow combined retrospective CRC cohort and determine association with prognosis and clinical features. Data were analysed with respect to GMS histological subtypes, MMR status and tumour subsite as a step towards understanding if tumour phenotype could be utilised to predict a subgroup of patients that may benefit from therapeutic inhibition of STAT3 activation. It was hypothesised that patients with high tumoural expression of nuclear and cytoplasmic STAT3, pSTAT3^{tyr705}, pSTAT3^{ser727} would have significantly reduced CSS compared to patients with low expression.

6.2 Expression of STAT3, pSTAT3^{tyr705} and pSTAT3^{ser727} in clinical colorectal cancer specimens

The Glasgow combined cohort tissue microarrays were stained for STAT3, pSTAT3^{tyr705} and pSTAT3^{ser727} via IHC. The total number of patients included in analyses were less than the 1030 patients in the full cohort due to exclusion criteria. Patients were excluded from analysis due to missing or damaged TMA cores, mortality within 30 days of surgery and/or receiving neoadjuvant chemotherapy, as shown in consort diagram (Figure 6.1).

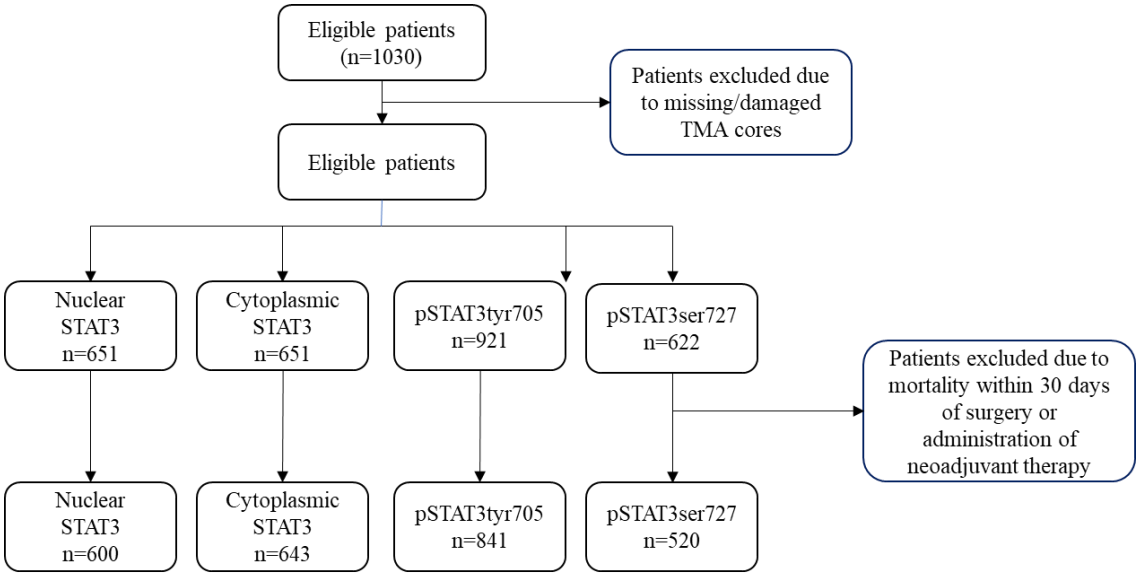


Figure 6.1 Consort diagram showing exclusion criteria. *Consort/flow diagram showing the number of patients included in analysis of each marker based on exclusion criteria of missing/damaged TMA cores, mortality within 390 days of surgery and/or administration of neoadjuvant chemotherapy.*

6.2.1 Expression of cytoplasmic STAT3 and clinical outcome

STAT3 expression was analysed in the Glasgow combined colorectal cohort via immunohistochemical tissue of the tissue microarrays. Representative images of weak, moderate and strong staining for STAT3 in the cytoplasm of tumour cells are shown in Figure 6.1A-C. Positive and negative control colorectal tissue are also shown in Figure 6.2(D-E). For STAT3, no true negative tissue was available for staining, however true positive, liver tissue, is shown in 6.2 (F). Antibody specificity testing was performed via western blotting, with single bands detected at 88KDa in HT29 and HeLa lysates and no bands detected at the correct molecular weight in true negative PC3 prostate

lysates (G) (Figure 6.2). Tubulin was detected at 52KDa in all 9 samples with similar band size indicating samples contained similar levels of protein (Figure 6.2).

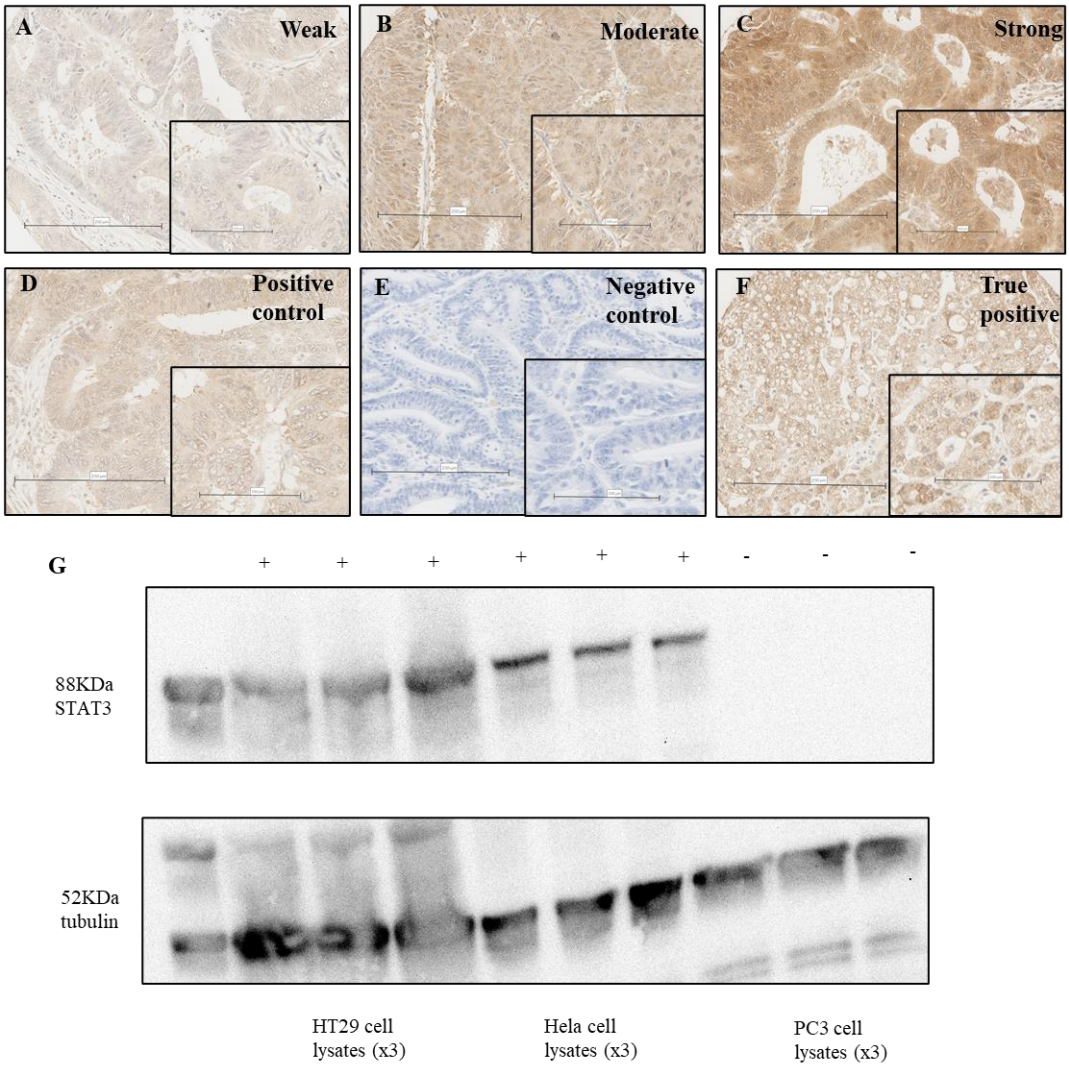


Figure 6.2 Representative images and antibody specificity for STAT3. *Images showing representative images of weak (A), moderate (B) and strong (C) cytoplasmic staining of STAT3 within tumour cells. Representative images of positive (D) and negative (E) colorectal tissue, and true positive tissue(F). Image of western blot showing single band in triplicate at 88Kda for HT29 cell lysates (20μL loaded per lane) and HeLa lysates (20μL loaded per lane) probed for STAT3 and no visible bands at 88KDa in triplicate for true negative PC3 lysates (20μL loaded per lane. For all 9 samples there was a visible band at 52KDa when probed for tubulin (G).*

Cytoplasmic STAT3 scoring was performed manually by a single observer (KP) in 651 patients from the Glasgow combined cohort. After patients who died within 30 days of surgery and/or those who received neoadjuvant therapy were excluded, data from 643 patients were available for downstream analysis. Cytoplasmic STAT3 scores ranged from 0-300 with a mean score of 62.37. A histogram was plotted to visualise the distribution of the data, which showed a positively skewed pattern (Figure

6.3). Validation of manual scoring was performed on 10% of cores using QuPath digital pathology software by Sara Al-Badran. A scatter plot was constructed to show the correlation between manual and digital scores with a correlation coefficient of 0.825 obtained (Figure 6.4). A Bland Altman plot was constructed to visualise variation and bias between manual and digital scores (Figure 6.5). The intra-class correlation coefficient was 0.863, which indicated a strong positive relationship between manual and digital validation scores. Optimal cut points for high and low expression of cytoplasmic STAT3 were determined using R Studio packages survminer and maxstat by log rank statistics based on CSS. A density and scatter plot were constructed to visualise the cut point of 63.0 (Figure 6.6). Patients with a weighted histoscore for cytoplasmic STAT3 of ≥63 were classified as high for cytoplasmic STAT3 expression (n=432) and ≤63 were low for cytoplasmic STAT3 expression (n=211).

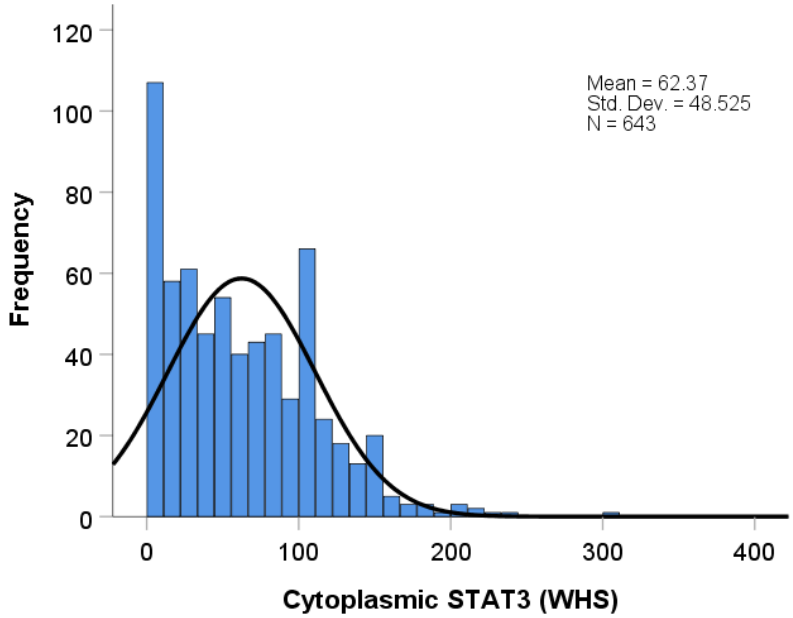


Figure 6.3 Distribution of weighted histoscores for cytoplasmic STAT3. *Histogram showing the range of scores obtained for cytoplasmic tumour STAT3 expression and positively skewed distribution pattern of the data (n=643). The mean score for cytoplasmic STAT3 was 62.37 and scores ranged from 0-300.*

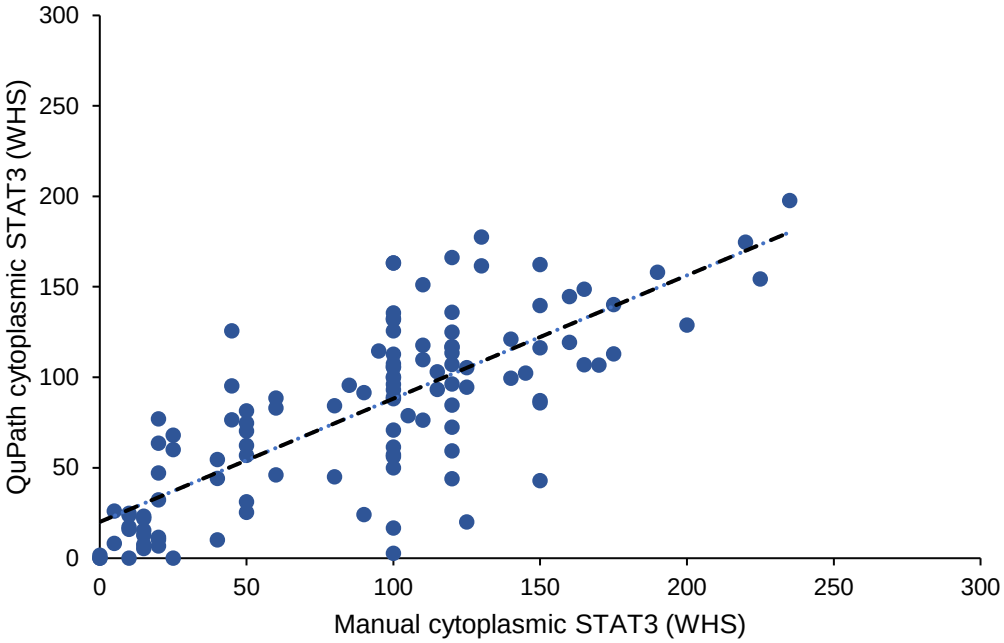


Figure 6.4. Correlation between manual and digital weighted histoscoring of cytoplasmic STAT3. *Scatter plot showing correlation between manual and digital tumour weighted histoscores for cytoplasmic STAT3 in 10% of the Glasgow combined array co-scored for validation purposes. A correlation coefficient of 0.825 was obtained between scores.*

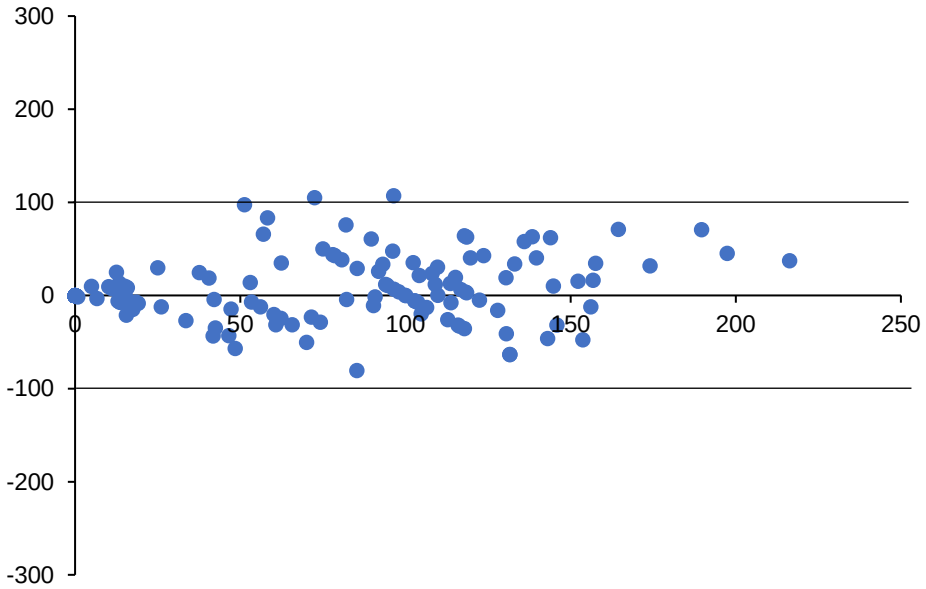


Figure 6.5 Validation of cytoplasmic STAT3 manual scoring. *Bland Altman plot showing difference between manual and digital QuPath scores for STAT3 in tumour cytoplasm of 10% of patients in the Glasgow combined array scored for validation purposes.*

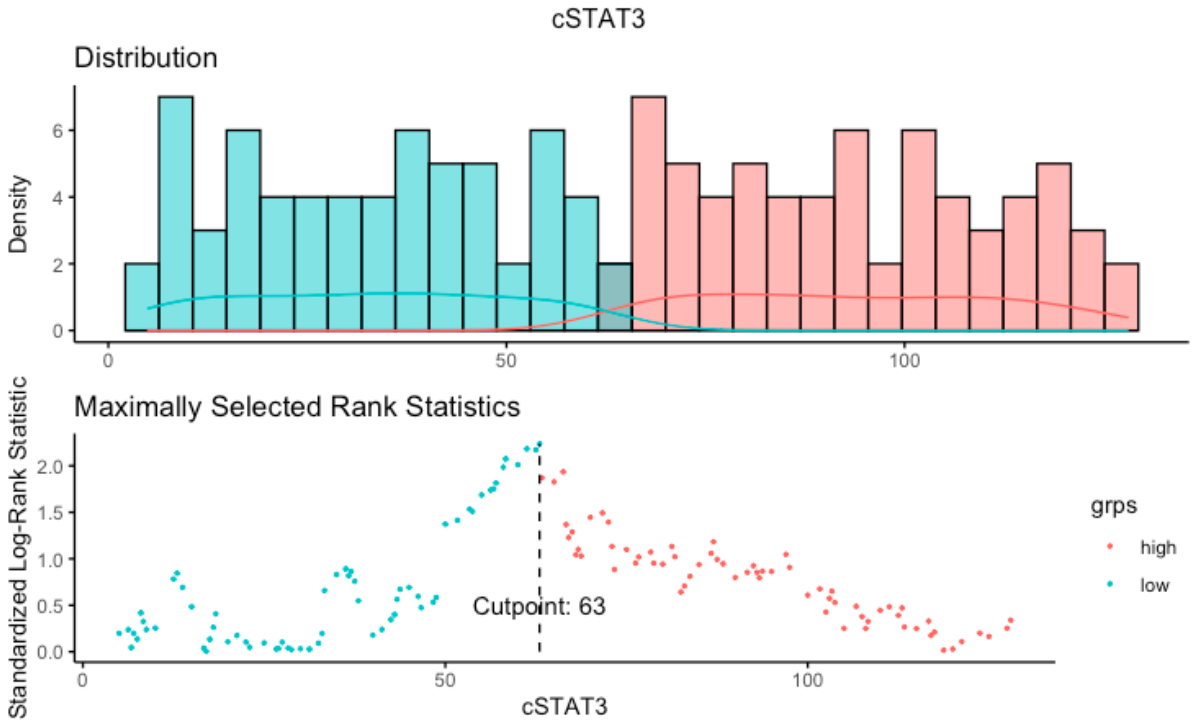


Figure 6.6 Defining cut off point for STAT3 expression high and low groups. *Density plot and scatter plot for visualisation of optimal cut off point for high and low expression of tumour cytoplasmic STAT3. The cut-off point determined was 63.0, therefore patients with a weighted histoscore of ≥63.0 were classed as high for cytoplasmic STAT3 expression and patients with a weighted histoscore of ≤63.0 were classed as low for cytoplasmic STAT3 expression.*

Kaplan Meier survival analysis was performed to determine any association with expression of STAT3 in the cytoplasm of tumour cells and CSS. Overall, high cytoplasmic expression of STAT3 was significantly associated with reduced CSS (HR=1.384, 95%CI; 1.040-1.841, log rank p=0.025) (Figure 6.6). The mean survival time of patients with high expression of STAT3 in the tumour cytoplasm was 139.370 months compared to 151.803 months for patients with low expression. When patients were stratified based on GMS histological subtype, no significant association between cytoplasmic STAT3 expression and CSS was observed in any group (Figure 6.7). There was a trend towards high expression associating with poorer outcome in GMS1 patients (p=0.080). There was no significant association between STAT3 expression and CSS when patients were stratified by MMR status, however there was a trend towards high expression associating with reduced survival time in MMR proficient cases (p=0.060) (Figure 6.8). High expression of cytoplasmic STAT3 was significantly associated with reduced CSS in patients with right-sided tumours (HR= 1.565, 95%CI; 1.030-2.378, log rank p=0.034) (Figure 6.9). The mean survival time of right-sided colon disease patients with high STAT3 expression was 131.051 months versus 147.555 months for patients with

low cytoplasmic STAT3 expression. No significant association between STAT3 expression in the cytoplasm and CSS was observed for left-sided or rectal cases (Figure 6.9).

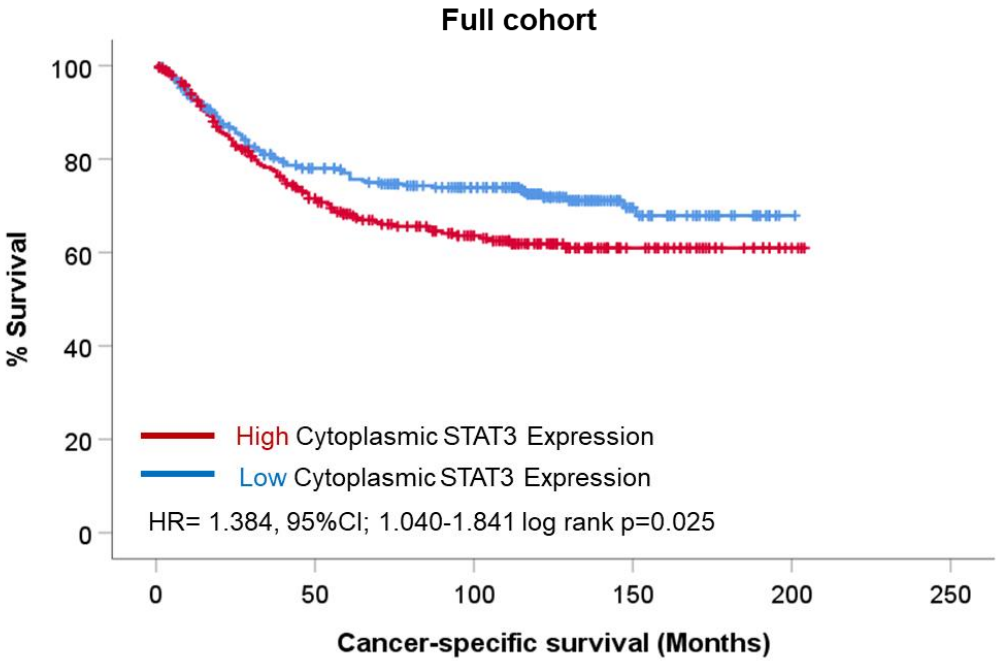


Figure 6.7 Cytoplasmic expression of STAT3 expression and CSS. *Kaplan Meier survival analysis of cytoplasmic STAT3 expression in patients from the Glasgow combined cohort. In the full cohort the hazard ratio for cytoplasmic STAT3 expression within the tumour cells was 1.384, (95%CI; 1.040-1.841) log rank p= 0.025.*

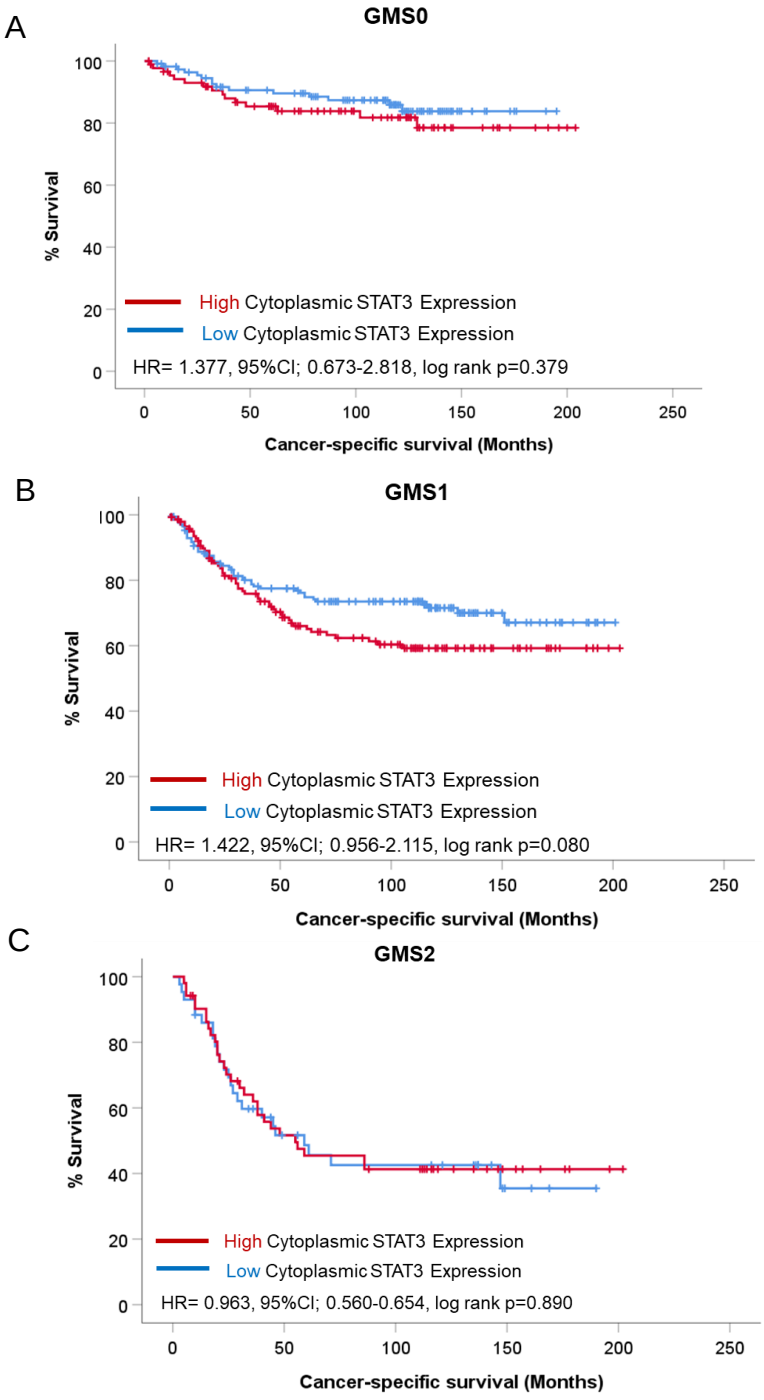


Figure 6.8 Cytoplasmic expression of STAT3 expression and CSS relative to GMS classification. *Kaplan Meier survival analysis of cytoplasmic STAT3 expression in patients from the Glasgow combined cohort classified as GMS0 (A), GMS1 (B), and GMS2 (C). In patients classified as GMS0 the hazard ratio for cytoplasmic STAT3 was 1.377, (95%CI; 0.673-2.818), log rank p=0.379 and in GMS1 patients the hazard ratio was 1.422, (95%CI; 0.560-0.654), log rank p=0.890. In GMS2 stromal-rich cases the hazard ratio for cytoplasmic STAT3 was 0.963, (95%CI; 0.560-0.654), log rank p=0.890.*

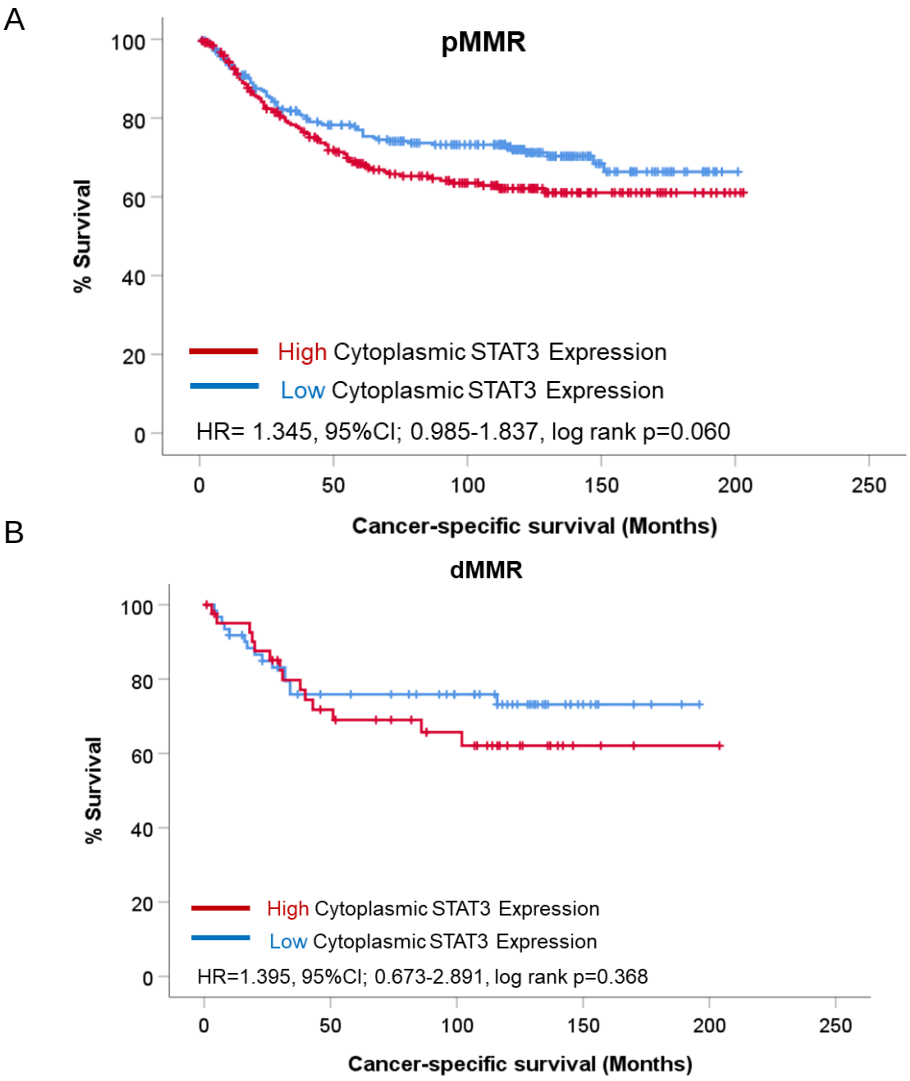


Figure 6.9 Cytoplasmic expression of STAT3 expression and CSS relative to MMR status. *Kaplan Meier survival analysis of cytoplasmic STAT3 expression in patients from the Glasgow combined cohort with MMR proficient tumours (A) and MMR deficient tumours (B). In MMR proficient cases the hazard ratio for cytoplasmic STAT3 was 1.345 (95%CI; 0.985-1.837), log rank p=0.060 and in MMR deficient cases the hazard ratio was 1.395 (95%CI; 0.673-2.891), log rank p=0.368.*

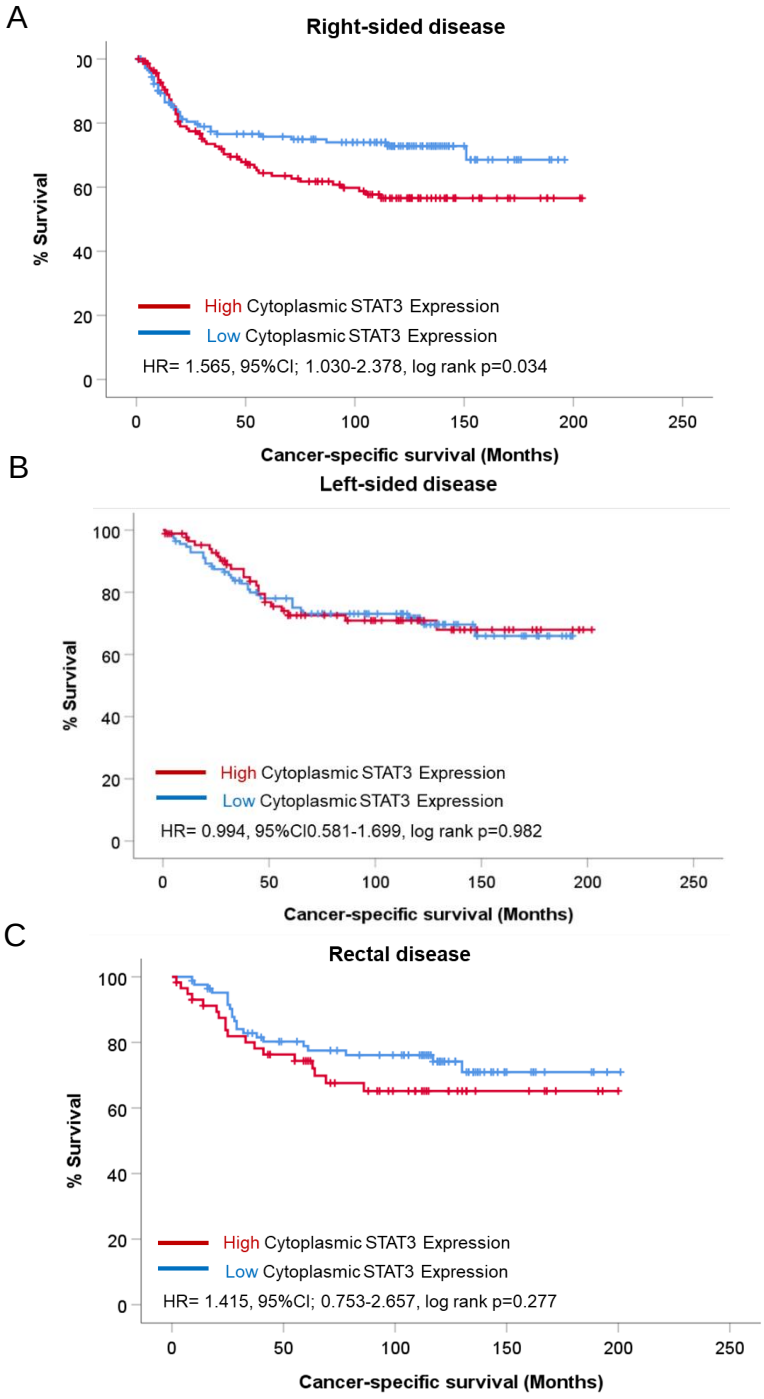


Figure 6.10 Cytoplasmic expression of STAT3 expression and CSS relative to tumour subsite. *Kaplan Meier survival analysis of cytoplasmic STAT3 expression in patients from the Glasgow combined cohort with right-sided (A), left-sided (B), and rectal disease (C). In patients with right-sided colonic tumours the hazard ratio for cytoplasmic STAT3 was 1.565, (95%CI; 1.030-2.378), log rank p=0.034 and in left-sided cases the hazard ratio was 0.994, (95%CI; 0.581-1.699, log rank p=0.982. In rectal cases the hazard ratio for cytoplasmic STAT3 was 1.415, (95%CI; 0.753-2.657), log rank p=0.277.*

Chi-squared tests were performed to determine any association between cytoplasmic expression of STAT3 and clinicopathological features. There were no significant associations between STAT3 expression, and any clinical feature included in the analysis (Table 6.1). There was a slight trend towards high cytoplasmic expression associating with increased systemic inflammation measured by mGPS ($p=0.075$).

Clinical Factor	Cytoplasmic STAT3 expression		p
	Low (n=347)	High (n=296)	
Age			
<65	109 (31.4)	95 (32.1)	0.460
>65	238 (68.6)	201 (67.9)	
Sex			
Female	169 (48.7)	149 (50.3)	0.369
Male	178 (51.3)	147 (49.7)	
T Stage			
I	16 (4.6)	10 (3.4)	0.462
II	42 (12.1)	41 (13.9)	
III	195 (56.2)	161 (54.4)	
IV	94 (27.1)	84 (28.4)	
N Stage			
0	209 (60.6)	189 (64.1)	0.591
I	100 (29.0)	75 (25.4)	
II	36 (10.4)	31 (10.5)	
Tumour subsite			
Right-sided colon	148 (42.9)	145 (49.7)	0.196
Left-sided colon	113 (32.8)	89 (30.5)	
Rectum	84 (24.3)	58 (19.9)	
GMS			
0	115 (34.6)	94 (32.3)	0.279
1	172 (51.8)	144 (49.5)	
2	45 (13.6)	53 (18.2)	
Phenotypic Subtype			
1	116 (35.0)	95 (32.8)	0.291
2	107 (32.3)	83 (28.6)	
3	64 (19.3)	58 (20.0)	
4	44 (13.3)	54 (18.6)	
mGPS			
0	178 (57.4)	104 (47.9)	0.075
1	81 (26.1)	64 (29.5)	
2	51 (16.5)	49 (22.6)	
MMR status			
pMMR	283 (82.3)	253 (85.8)	0.138
dMMR	61 (17.7)	42 (14.2)	
Tumour differentiation			
0			
1	303 (87.3)	259 (87.5)	0.521
	44 (12.7)	37 (120.5)	
Marginal involvement			
0	325 (93.7)	284 (95.9)	0.132
1	22 (6.3)	12 (4.1)	
Vascular invasion			
0	228 (65.7)	180 (60.8)	0.115
1	119 (34.3)	116 (39.2)	

Table 6.1 Cytoplasmic STAT3 Expression and Clinical Features. *Table of Chi-squared associations between cytoplasmic tumoural STAT3 expression and clinicopathological prognostic features including age, sex, T stage, N stage, subsite, GMS, phenotypic subtype, mGPS, MMR status, tumour differentiation, marginal involvement and vascular invasion.*

6.2.2 Expression of nuclear STAT3 and clinical outcome

When STAT3 becomes activated, it translocates from the cytoplasm to the nucleus. Positive staining for STAT3 was present in the nuclei of tumour cells in some of the Glasgow combined array TMA cores. Representative images of weak (A), moderate (B) and strong (C) nuclear STAT3 staining and shown in Figure 6.11.

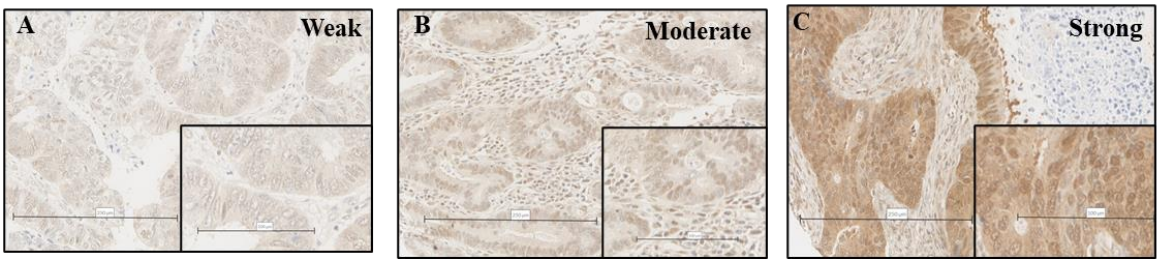


Figure 6.11 Nuclear STAT3 Expression in the Glasgow combined cohort. *Representative images of weak (A), moderate (B) and strong (C) nuclear staining of STAT3 in patients from the Glasgow combined cohort stained via conventional IHC.*

Nuclear scoring of STAT3 of 651 stage I-IV CRC patients from the Glasgow combined array was performed manually by a single observer (KP). After exclusion criteria were applied (30-day mortalities post-surgery and neoadjuvant therapy) a total of 600 patients were included in downstream analysis. Scores ranged from 0 to 245 with a mean score of 47.01. A histogram was plotted to visualise the distribution pattern of scores which was positively skewed (Figure 6.12). Validation of scoring was performed on 10% of TMA cores manually by Sara Al-Badran. A scatter plot was constructed to visualise the correlation between manual and digital scores with a correlation coefficient of 0.901 obtained, indicating a strong positive relationship (Figure 6.13). A Bland Altman plot was constructed to show visualise the difference between the sets of scores (Figure 6.14). The intra class correlation coefficient between scores was 0.910. Survminer and maxstat packages in R Studio were utilised to determine optimal cut points for high and low expression of nuclear STAT3 relative to CSS. A scatter plot and density plot were used to visualise the optimally defined cut point of 31.67 (Figure 6.15). Patients with a weighted histoscore of ≥31.67 for nuclear STAT3 were assigned to the high expression group (n=108) and those with a score of ≤31.67 were considered low expressors (n=492).

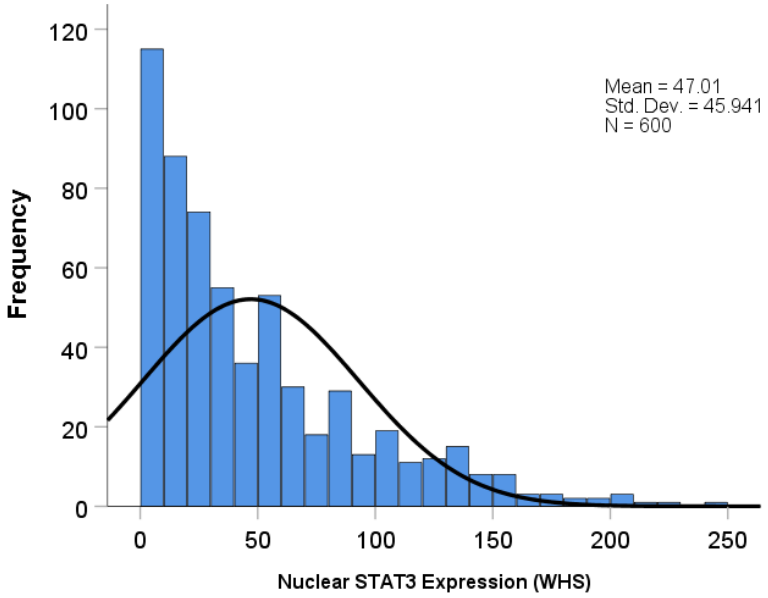


Figure 6.12 Distribution of weighted histoscores for nuclear STAT3. *Histogram showing the range of scores obtained for tumour nuclear STAT3 expression and distribution pattern of the data (n=600). The mean score for nuclear STAT3 was 47.01 and the scores ranged from 0-245.*

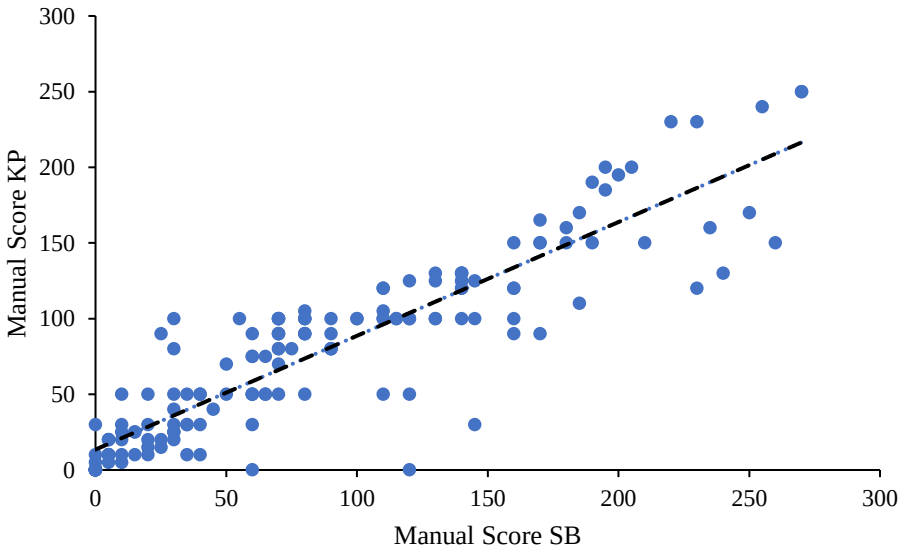


Figure 6.13 Correlation between manual and digital weighted histoscoring of nuclear STAT3. *Scatter plot showing correlation between tumour weighted histoscores for nuclear STAT3 between two observers manual scores for 10% of the Glasgow combined cohort utilised for validation purposes. The correlation coefficient obtained was 0.901.*

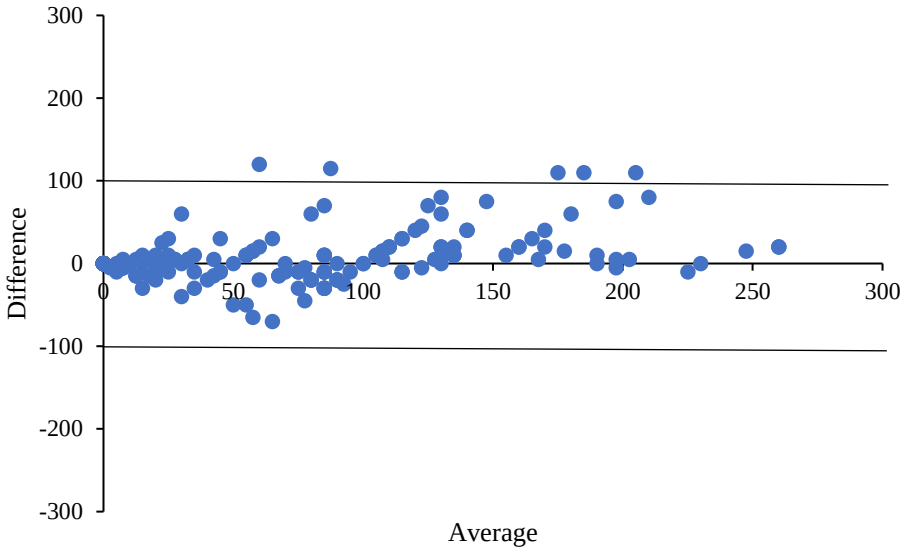


Figure 6.14 Validation of nuclear STAT3 manual scoring. *Bland Altman plot showing difference between two observers manual scores for STAT3 in tumour nuclei in the 10% of the Glasgow combined cohort utilised for validation purposes.*

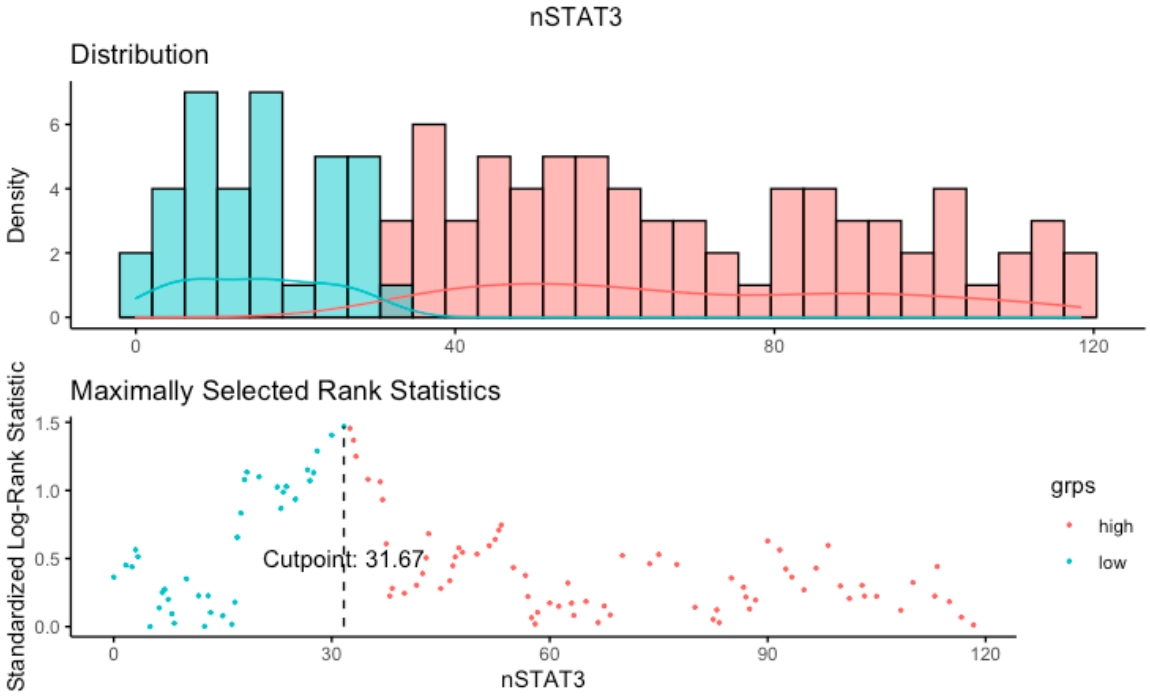


Figure 6.15 Defining cut off point for nuclear STAT3 expression high and low groups. *Density plot and scatter plot for visualisation of optimal cut off point for high and low expression of tumour nuclear STAT3. The optimal cut off point determined by survminer was 31.67, therefore patients with a weighted histoscore of ≤31.67 were classed as low expression and patients with a weighted histoscore of ≥31.67 were classified as high for nuclear expression of total STAT3.*

Kaplan Meier curves survival analysis was performed to determine any association between nuclear STAT3 expression and CSS. Overall, there was no significant association between nuclear STAT3 expression and CSS (p=0.142) (Figure 6.16). When patients were stratified by GMS histological subtypes, high nuclear STAT3 expression was significantly associated with reduced survival time in GMS1 patients (HR=1.729, 95%CI; 1.128-2.652, log rank p=0.011) (Figure 6.17). The mean survival time for GMS1 patients with high nuclear STAT3 expression was 134.648 months compared to 155.282 months for patients with low nuclear STAT3 expression. No association between nuclear STAT3 expression and CSS was observed in the GMS0 or GMS2 group of patients (Figure 6.17). When patients were segregated by MMR status and subsequently tumour subsite no significant association between nuclear STAT3 expression and CSS was observed in any group (Figure 6.18 and Figure 6.19, respectively)

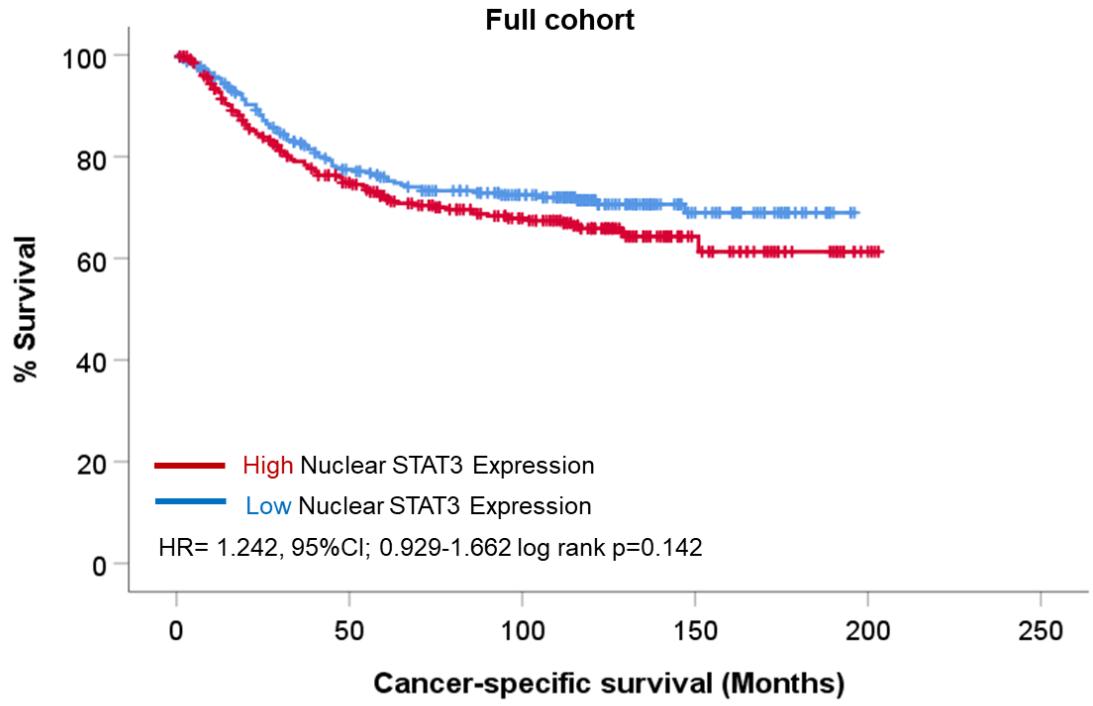


Figure 6.16 Nuclear expression of STAT3 expression and CSS. *Kaplan Meier survival analysis of nuclear STAT3 expression in patients from the Glasgow combined cohort. The hazard ratio for tumour nuclear STAT3 expression in the full cohort was 1.242, (95%CI; 0.929-1.662), log rank p=0.142.*

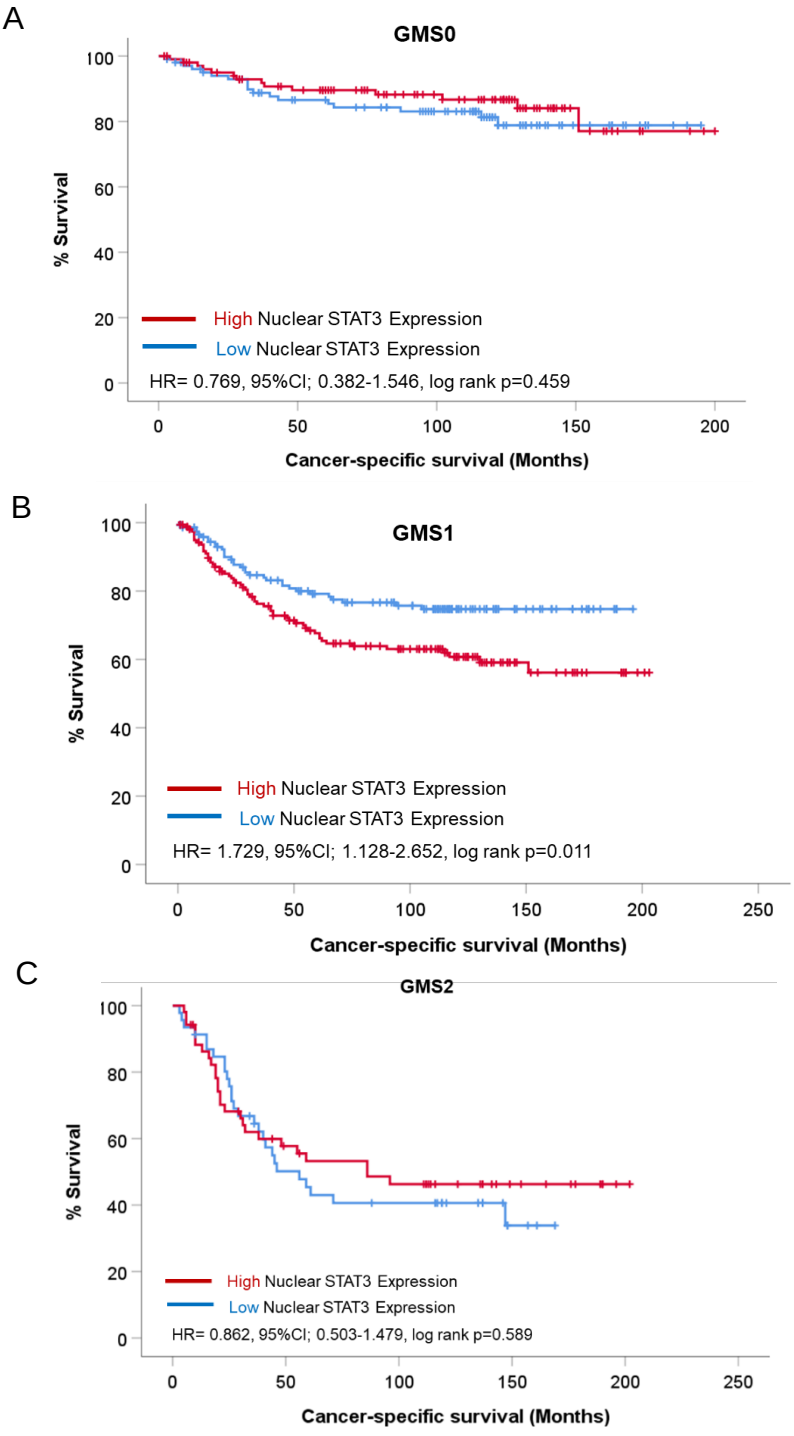


Figure 6.17. Nuclear expression of STAT3 expression and CSS relative to GMS classification. *Kaplan Meier survival analysis of nuclear STAT3 expression in patients from the Glasgow combined cohort classified as GM0 (A), GMS1 (B), and GMS2 (C). In GMS0 immune graded cases the hazard ratio for nuclear STAT3 was 0.769; (95%CI;0.382-1.546), log rank p=0.459, and the hazard ratio in patients with GMS1 tumours was 1.729, (95%CI;1.128-2.652), log rank p=0.011. In GMS2 stromal-rich cases the hazard ratio for nuclear STAT3 expression was 0.862, (95%CI; 0.503-1.479), log rank= 0.589.*

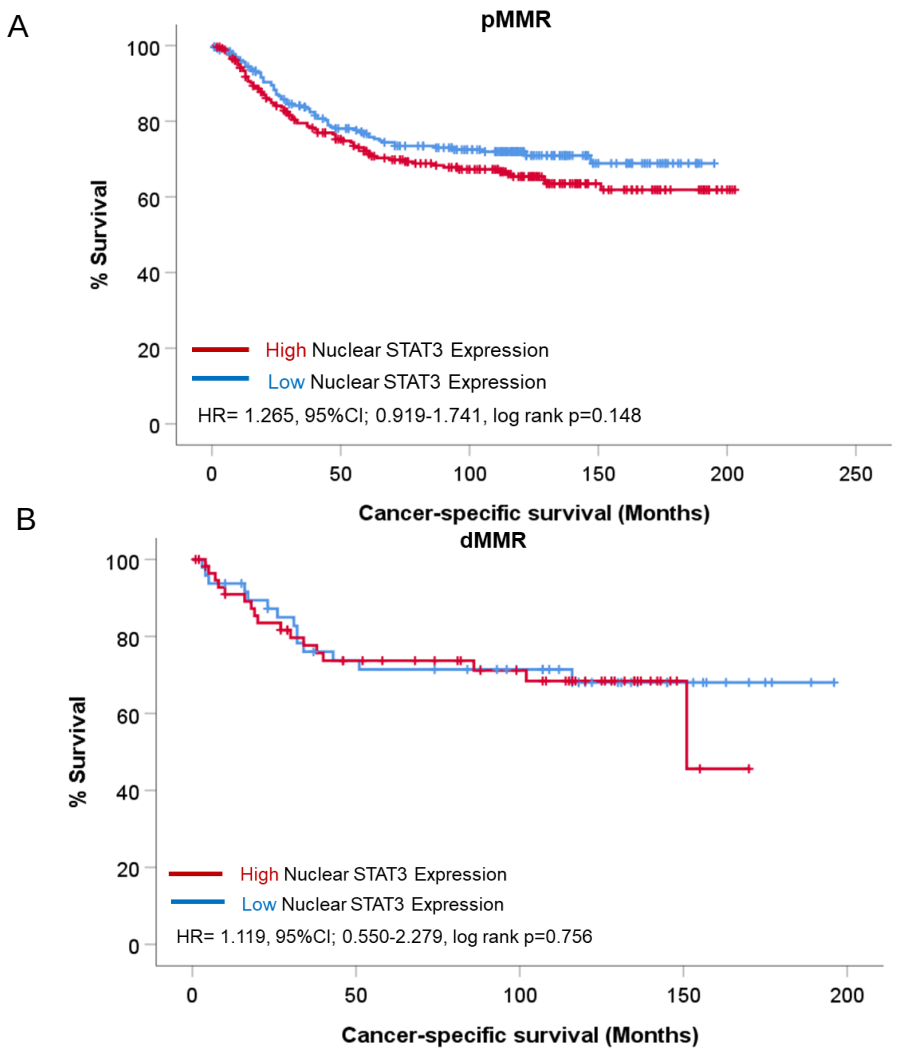


Figure 6.18 Nuclear expression of STAT3 expression and CSS relative to MMR status.

Kaplan Meier survival analysis of nuclear STAT3 expression in patients from the Glasgow combined cohort with MMR proficient tumours (A) and MMR deficient tumours (B). In MMR proficient cases the hazard ratio for nuclear STAT3 was 1.265, (95%CI; 0.919-1.741), log rank p=0.148 and in MMR deficient cases the hazard ratio was 1.119, (95%CI; 0.550-2.279), log rank p=0.756.

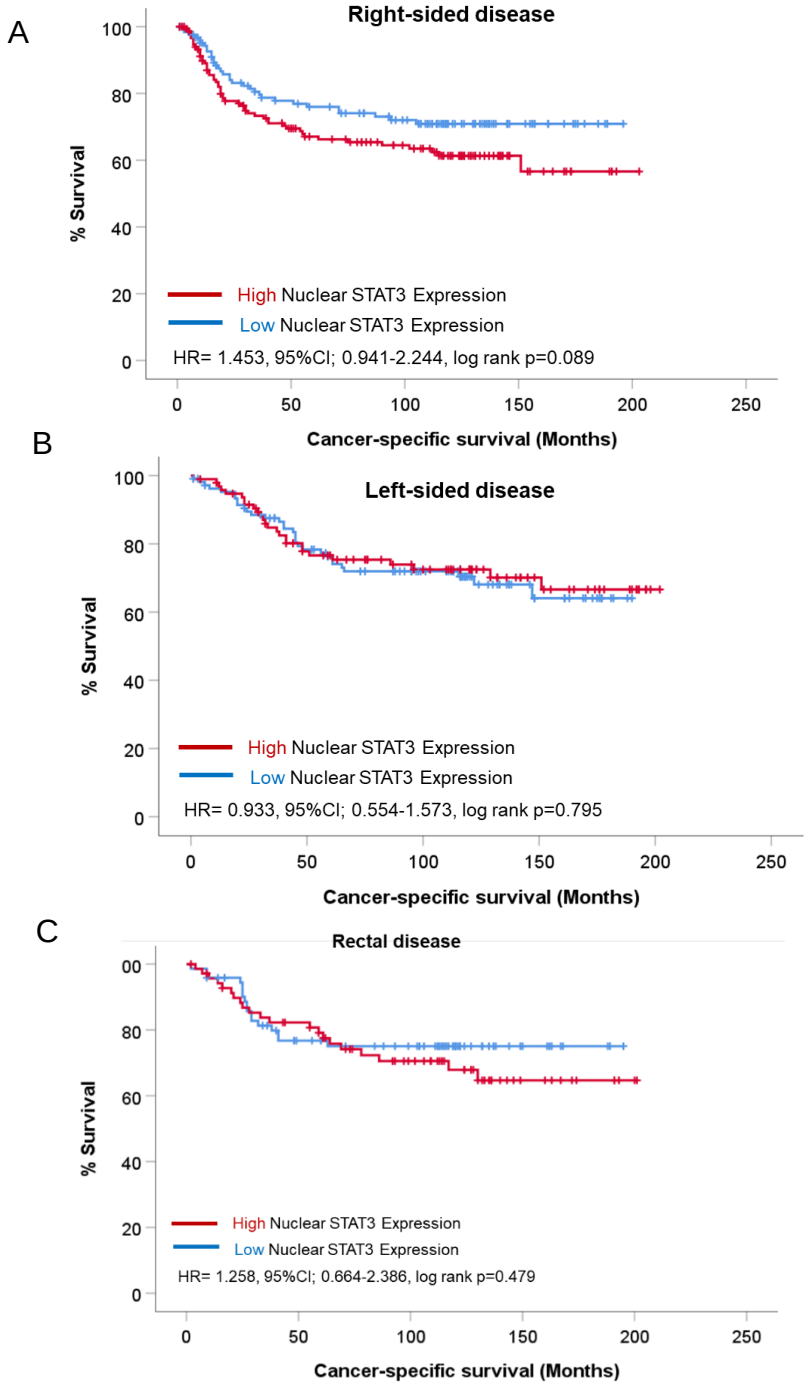


Figure 6.19 Nuclear expression of STAT3 expression and CSS relative to tumour subsite. *Kaplan Meier survival analysis of nuclear STAT3 expression in patients from the Glasgow combined cohort with right-sided (A), left-sided (B), and rectal disease (C). In patients with right-sided colonic tumours the hazard ratio for nuclear STAT3 was 1.453, (95%CI; 0.941-2.244), log rank p=0.089 and in left-sided colonic tumours the hazard ratio was 0.933, (95%CI;0.554-1.573), log rank p=0.795. In rectal cases the hazard ratio for nuclear STAT3 was 1.258, (95%CI; 0.664-2.386), log rank p=0.479.*

Chi-squared tests were conducted to determine any association between nuclear expression of STAT3 and host clinical characteristics. High nuclear STAT3 expression was associated with stromal tumour phenotypes GMS2 ($p=0.003$) and the stromal phenotypic subtype 4 ($p=0.009$) (Table 6.2).

Clinical Factor	Nuclear STAT3 Expression		p
	Low (n=310)	High (n=328)	
Age			
<65	105 (33.9)	102 (31.1)	0.254
>65	205 (66.1)	226 (68.9)	
Sex			
Female	146 (47.1)	169 (51.5)	0.149
Male	164 (52.9)	159 (48.5)	
T Stage			
I	14 (4.5)	12 (3.7)	0.781
II	39 (12.6)	42 (12.8)	
III	168 (54.2)	1898 (57.6)	
IV	89 (28.7)	85 (25.9)	
N Stage			
0	190 (61.5)	207 (63.5)	0.305
I	91 (29.4)	81 (24.8)	
II	28 (9.1)	38 (11.7)	
Tumour subsite			
Right-sided colon	128 (41.6)	156 (48.0)	0.246
Left-sided colon	108 (35.1)	98 (30.2)	
Rectum	72 (23.4)	71 (21.8)	
GMS			
0	103 (34.7)	107 (33.3)	0.932
1	147 (49.5)	161 (50.2)	
2	47 (15.8)	53 (16.5)	
Phenotypic Subtype			
1	104 (35.0)	108 (33.9)	**0.009**
2	74 (24.9)	110 (34.5)	
3	72 (24.2)	48 (15.0)	
4	47 (15.8)	53 (26.6)	
mGPS			
0	157 (58.1)	121 (48.8)	**0.003**
1	77 (28.5)	65 (26.5)	
2	36 (13.3)	62 (25.0)	
MMR status			
pMMR	261 (84.5)	267 (82.2)	0.251
dMMR	48 (15.5)	58 (17.8)	
Tumour differentiation			
0	266 (85.8)	295 (89.9)	0.069
1	44 (14.2)	33 (10.1)	
Marginal involvement			
0	293 (94.5)	312 (95.1)	0.433
1	17 (5.5)	16 (4.9)	
Vascular invasion			
0	207 (66.8)	204 (66.2)	0.130
1	103 (33.2)	124 (37.8)	

Table 6.2 Nuclear STAT3 Expression and Clinical Features. *Table of Chi-squared associations between nuclear tumoural STAT3 expression and clinicopathological prognostic features including age, sex, T stage, N stage, tumour subsite, GMS, phenotypic subtype, MMR status, mGPS, tumour differentiation, marginal involvement, and venous invasion.*

6.2.3 Expression of pSTAT3^{705} and clinical outcome

The main phosphorylation site of STAT3 is at tyrosine 705. STAT3^{tyr705} expression was assessed in the Glasgow combined colorectal cohort via immunohistochemistry. Representative images of weak, moderate, and strong pSTAT3^{tyr705} staining are shown in figure 6.20 (A-C). Positive and negative control colorectal tissue were stained alongside the cohort and representative images of staining can be observed in Figure 6.20 (D-E). True positive tissue, lung, and true negative, healthy colorectal epithelium were also stained alongside the cohort (Figure 6.20) (F-G). Antibody specificity for pSTAT3^{tyr705} was initially performed via staining COLO205 cell pellets treated with either 0.01% DMSO or 1mM JAK1/2 inhibitor Ruxolitinib (H) (Figure 6.20). Western blotting was performed for additional validation of specificity utilising Sw620 cell line lysates treated with either 0.01% DMSO or 1mM Ruxolitinib. Single bands at 88KDa were obtained in triplicate for the vehicle control samples with no bands detected at 88KDa in the treated samples (Figure 6.20). Tubulin remained consistent across all 6 samples indicating that similar levels of protein were present in all samples (Figure 6.20).

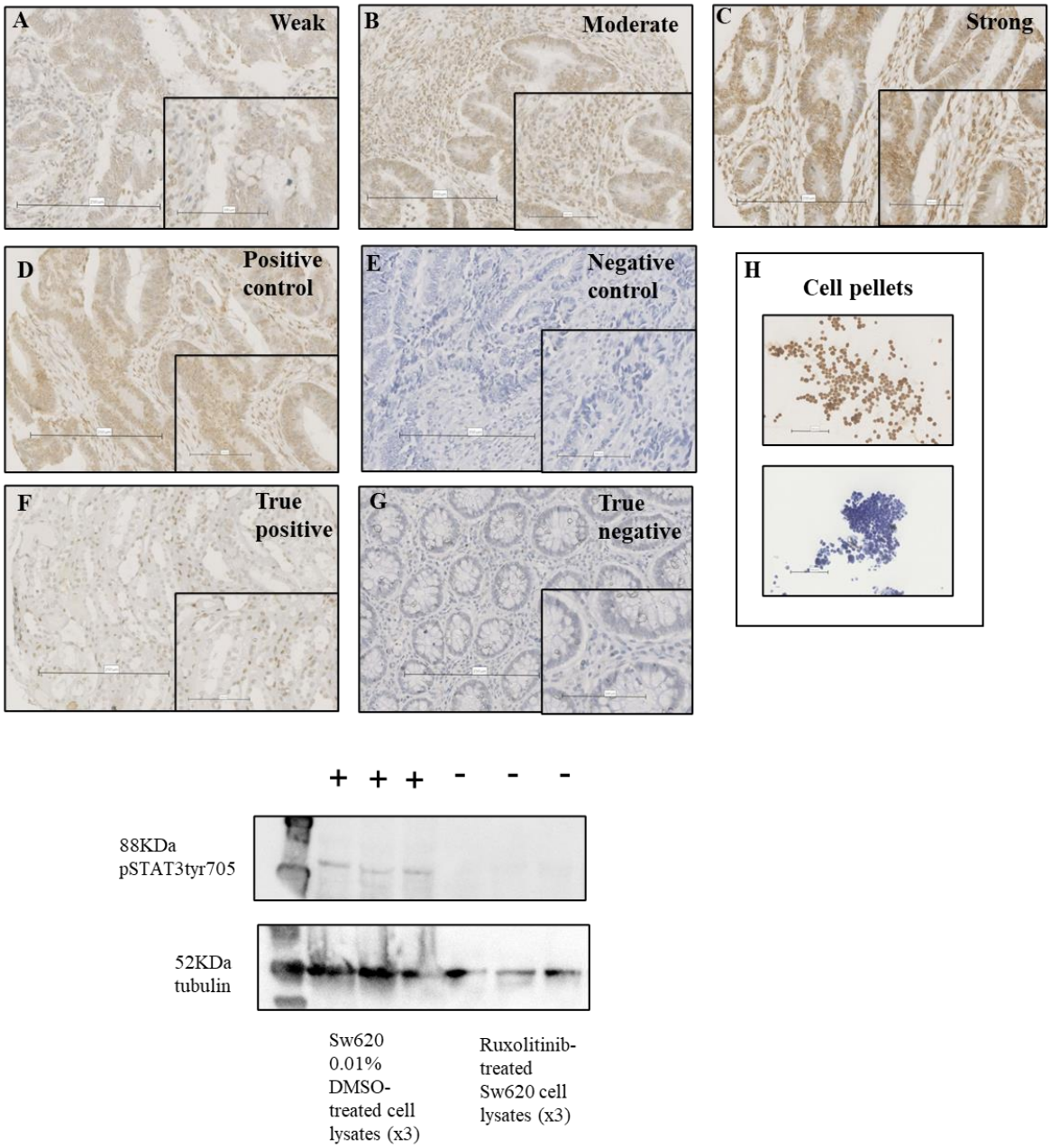


Figure 6.20 Representative images and antibody specificity for pSTAT3^{tyr705}. *Images showing representative images of weak (A), moderate (B) and strong (C) cytoplasmic staining of pSTAT3^{tyr705} within tumour cells. Representative images of positive (D) and negative (E) colorectal tissue, and true positive tissue(F). COLO205 cell pellets treated with vehicle control or JAK1/2 inhibitor and then stained for pSTAT3^{tyr705}(G). Image of western blot showing single band in triplicate for SW620 cell lysates (20μL loaded per lane) treated with vehicle control probed for pSTAT3^{tyr705}and no visible bands in triplicate for JAK1/2 inhibited SW620 cells (20μL loaded per lane) at 88KDa (H).*

Expression of pSTAT3^{705} in tumour nuclei was scored manually by a single observer (James H Park) in 921 patients from the Glasgow combined cohort. Of these, 80 patients were excluded due to mortality within 30 days of surgery, or administration of neoadjuvant chemotherapy leaving 841 available for statistical analysis. Weighted histoscores ranged from 0 to 205 with a mean score of 47.45. A histogram was plotted to visualise the distribution pattern of scores, which showed positive skewing (Figure 6.20). Scoring was validated using QuPath digital platform by KP on 10% of TMA cores. A correlation coefficient of 0.83 was obtained indicating a strong positive correlation between manual and digital scores. A scatter plot was constructed to visualise this correlation (Figure 6.21). A Bland Altman plot was used to visualise the bias and difference between manual and digital scoring (Figure 6.22). An intra-class correlation coefficient fell shy of acceptable at 0.612, however after assessing the cores with the biggest differences it was determined that manual scores were correct. Log rank statistics were used to determine optimal cut offs for high and low expression groups using survminer and maxstat packages in R Studio. Density and scatter plots were constructed to visualise the optimum cut point of 38.75 (Figure 6.23). Patients with a weighted histoscore for nuclear pSTAT3^{tyr705} of ≥38.75 (n=399) were classified as high for pSTAT3^{tyr705} expression and ≤38.75 were classified as low expression (n=442).

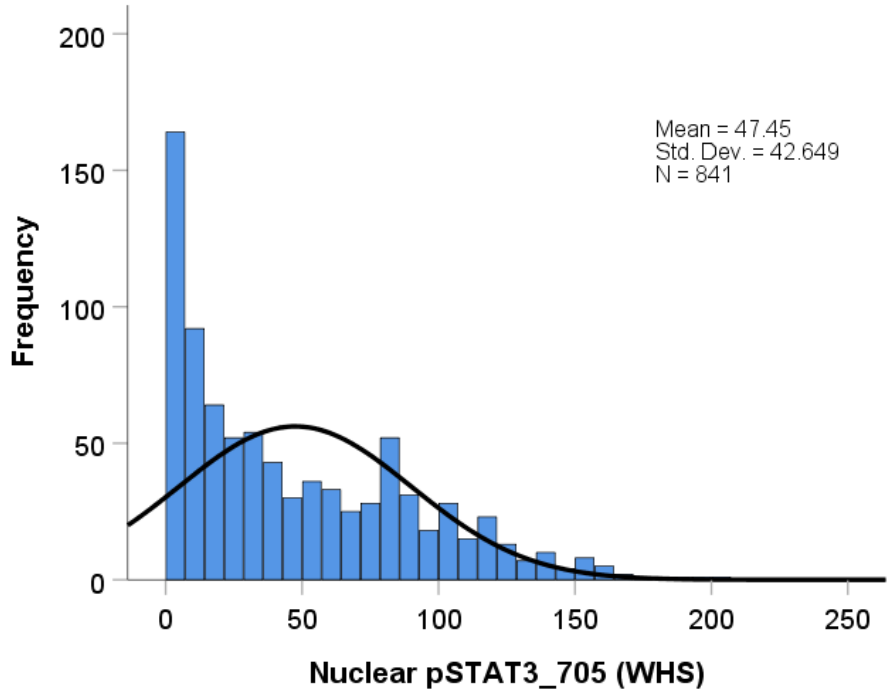


Figure 6.21 Distribution of weighted histoscores for nuclear pSTAT3^{tyr705}. *Histogram showing the range of scores obtained for nuclear tumour pSTAT3^{tyr705} expression and positively skewed distribution pattern of the data. The mean score for nuclear STAT3^{tyr705} was 47.45 and scores ranged from 0-205.*

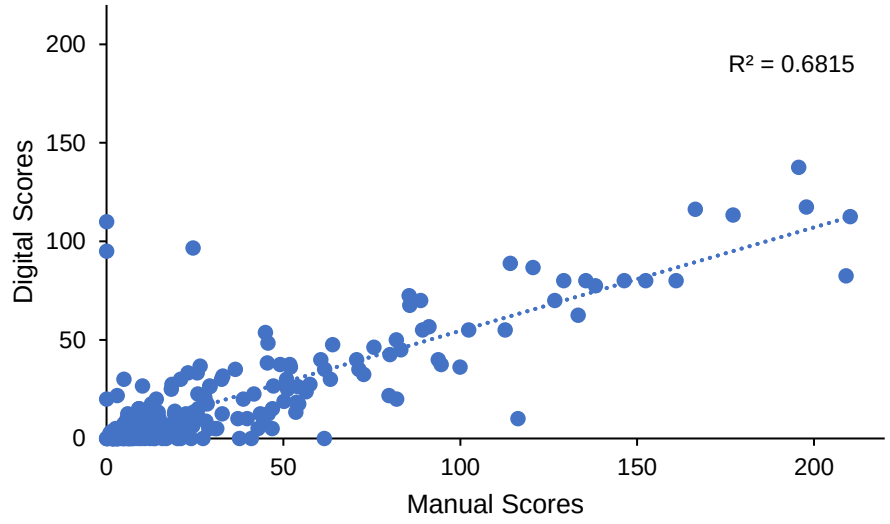


Figure 6.22 Correlation between manual and digital weighted histoscoring of pSTAT3^{tyr705}. *Scatter plot showing correlation between tumour cytoplasm weighted histoscores for pSTAT3^{tyr705} in 10% of patients from the Glasgow combined cohort co-scored for validation. A correlation coefficient of 0.6815 was obtained.*

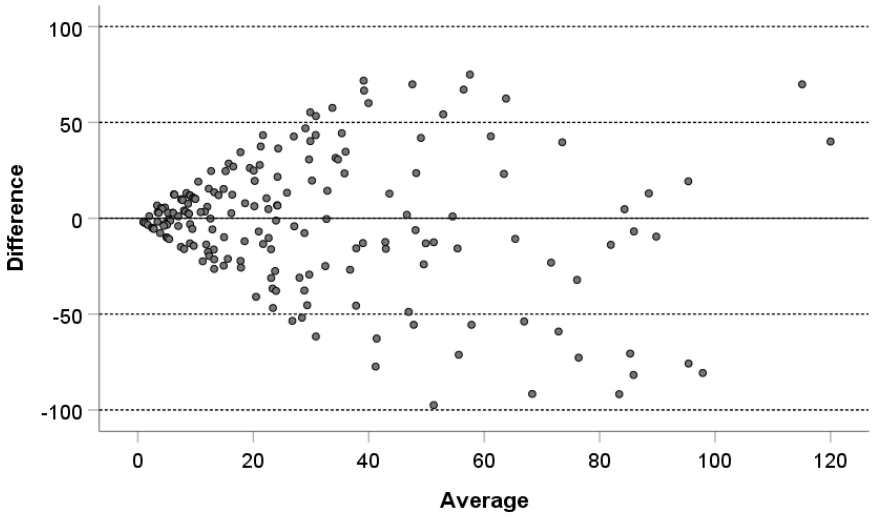


Figure 6.23 Validation of pSTAT3^{tyr705} manual scoring. *Bland Altman plot showing difference between manual and digital QuPath scores for pSTAT3^{tyr705} in tumour cytoplasm in 10% of patients from the Glasgow combined cohort co-scored for validation purposes.*

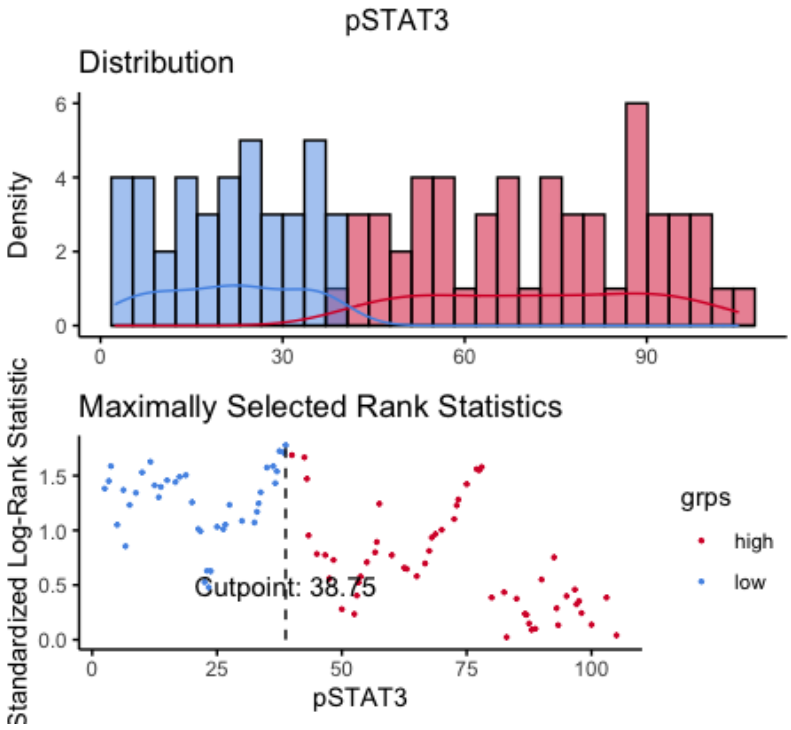


Figure 6.24 Defining cut off point for nuclear pSTAT3^{tyr705} expression high and low groups. *Density plot and scatter plot for visualisation of optimal cut off point for high and low expression of tumour nuclear pSTAT3^{tyr705}. The optimal cut point determined was 38.75 and therefore patients with a weighted histoscore of ≥38.75 were classed as high for nuclear pSTAT3^{tyr705} and patients with a weighted histoscore ≤38.75 were classed as low.*

Kaplan Meier survival analysis was performed to determine any association between pSTAT3^{tyr705} expression and CSS. In terms of the full cohort, there was a trend towards high nuclear expression of pSTAT3^{tyr705} associating with reduced CSS, however significance was not reached (HR= 1.236, 95%CI; 0.790-1.710, log rank p=0.101) (Figure 6.24). When patients were stratified by GMS, there was a profound association of pSTAT3^{tyr705} expression and CSS in the GMS2 patients with stromally-dense tumours (HR= 2.086, 95%CI; 1.224-3.554, log rank p = 0.006) (Figure 6.25). The mean survival time of GMS2 patients with high pSTAT3^{tyr705} expression was 90.666 months versus those with low expression at 134.157 months. This association with outcome was not observed in patients with GMS0 immune tumours or GMS1 tumours (Figure 6.26). When patients were segregated by MMR status, there was no significant association between pSTAT3^{tyr705} expression and CSS (Figure 6.27). Kaplan Meier survival analysis was then performed on data from patients grouped by tumour subsite. High expression of pSTAT3^{tyr705} in rectal cases showed a trend towards an association with reduced CSS (HR= 1.778, 95%CI; 0.956-3.306, log rank p= 0.065) (Figure 6.28). There was no significant association between pSTAT3^{tyr705} status and CSS in patients with tumours of right or left-sided colon origin (Figure 6.27).

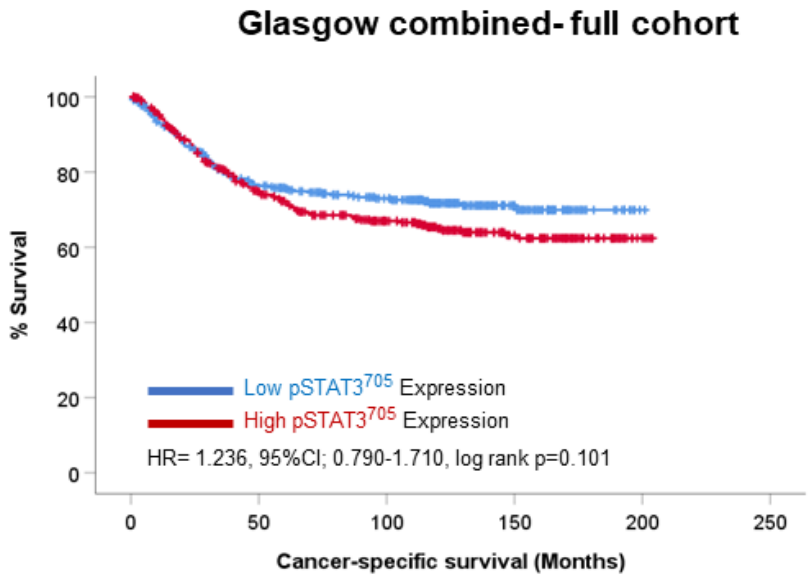


Figure 6.25 Nuclear expression of pSTAT3^{tyr705} expression and CSS. *Kaplan Meier survival analysis of nuclear pSTAT3^{tyr705} expression in patients from the Glasgow combined cohort. The hazard ratio for pSTAT3^{tyr705} in the full cohort was 1.236, (95%CI; 0.790-1.710), log rank= 0.101.*

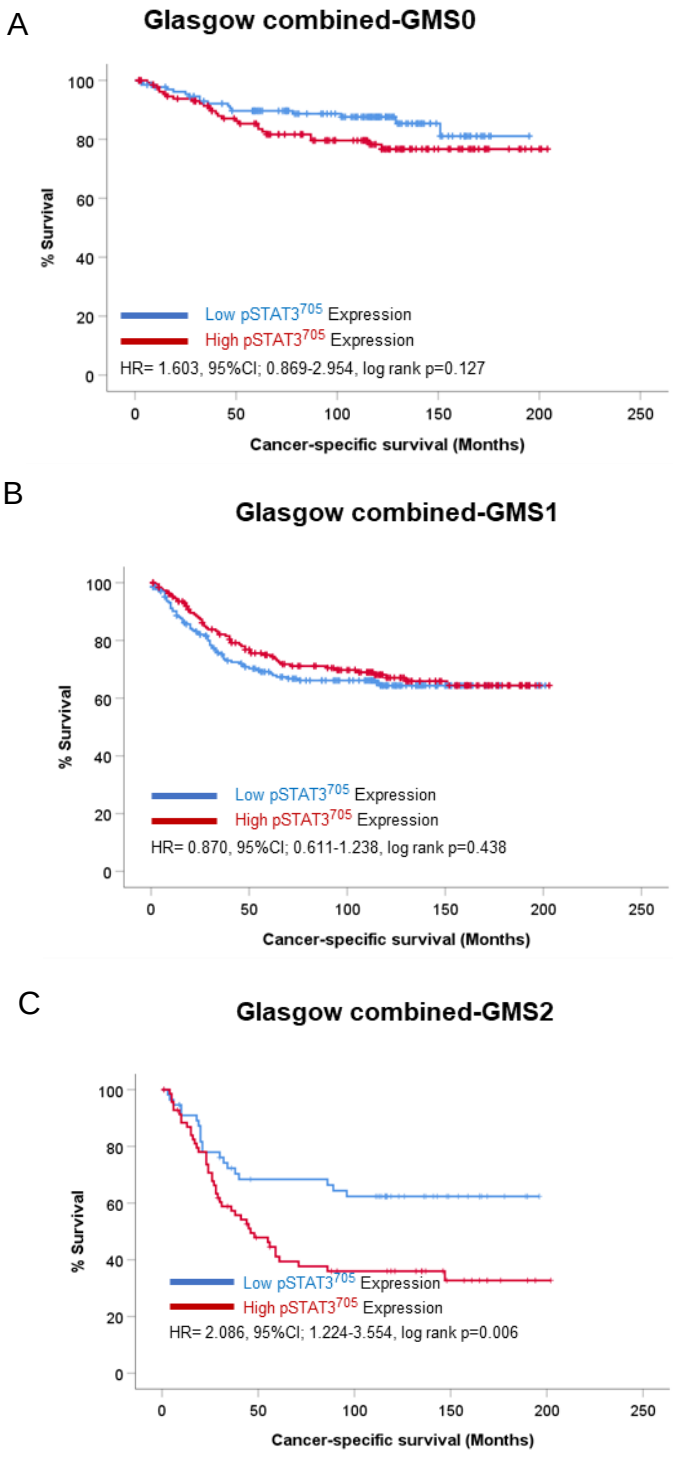


Figure 6.26 Nuclear expression of pSTAT3^{tyr705} expression and CSS relative to GMS classification. *Kaplan Meier survival analysis of nuclear pSTAT3^{tyr705} expression in patients from the Glasgow combined cohort classified as GM0 (A), GMS1 (B), and GMS2 (C). In GMS0 immune patients the hazard ratio for nuclear pSTAT3^{tyr705} was 1.603, (95%CI; 0.869-2.954), log rank p=0.127 and in GMS1 cases the hazard ratio was 0.870, (95%CI; 0.611-1.238), log rank p=0.438. In GMS2 cases the hazard ratio for nuclear STAT3^{tyr705} was 2.085, (95%CI; 1.224-3.554), log rank p=0.006.*

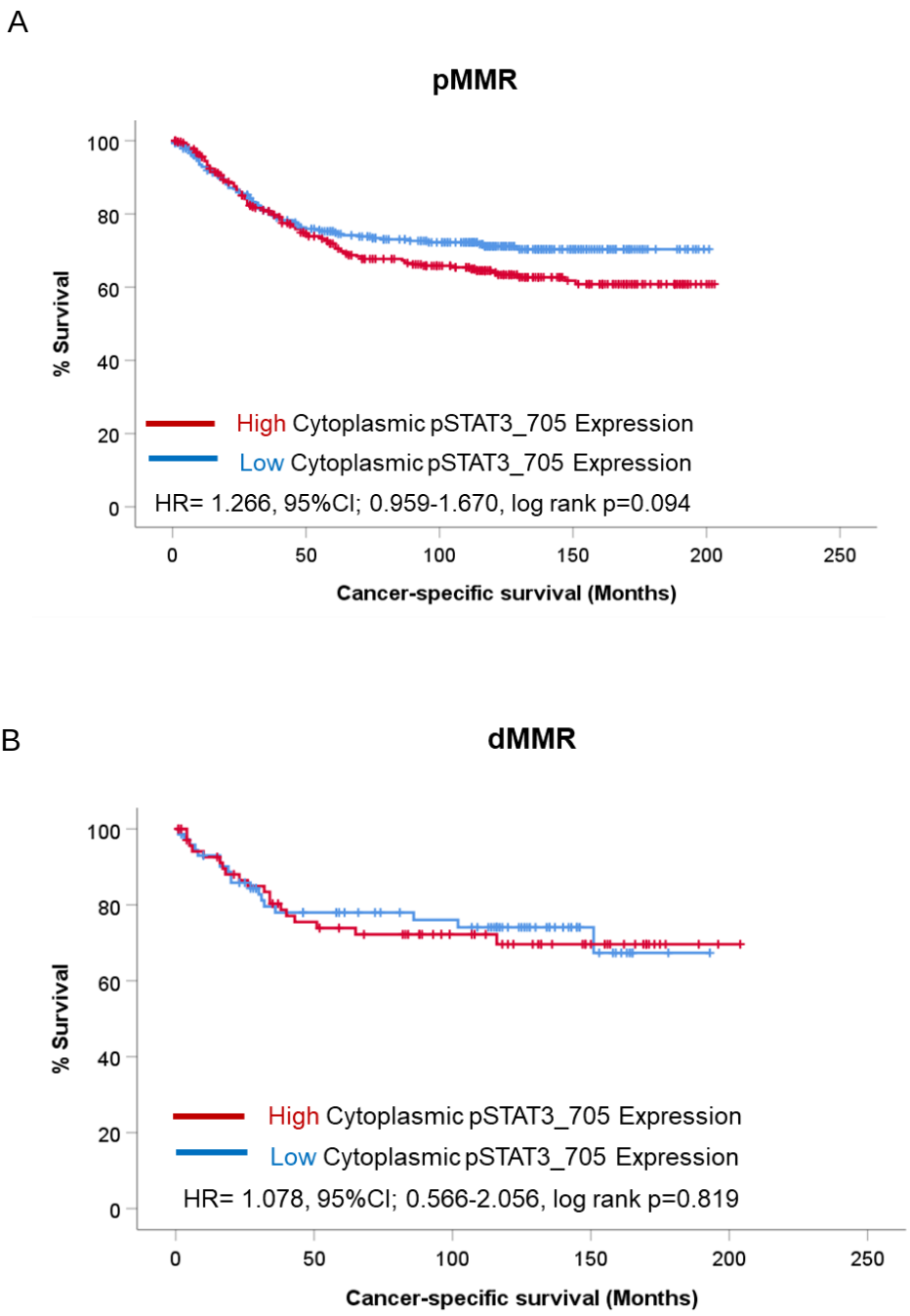


Figure 6.27 Nuclear expression of pSTAT3^{tyr705} expression and CSS relative to MMR status. *Kaplan Meier survival analysis of nuclear pSTAT3^{tyr705} expression in patients from the Glasgow combined cohort with MMR proficient tumours (A) and MMR deficient tumours (B). In MMR proficient cases the hazard ratio for pSTAT3^{tyr705} expression was 1.266, (95%CI;0.959-1.670), log rank p=0.094 and in MMR deficient cases the hazard ratio was 1.078, (95%CI; 0.566-2.056) log rank p=0.819.*

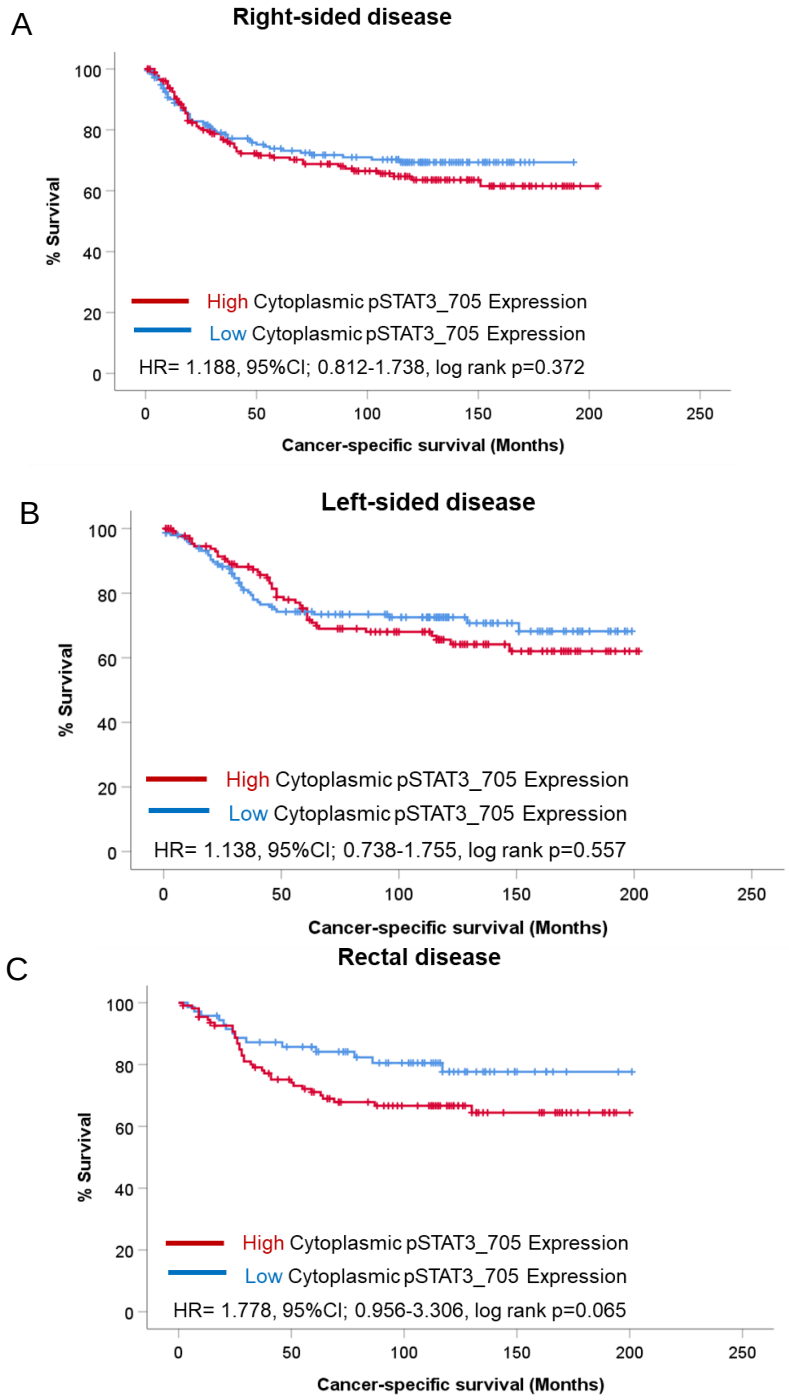


Figure 6.28 Nuclear expression of pSTAT3^{tyr705} expression and CSS relative to tumour subsite. *Kaplan Meier survival analysis of nuclear pSTAT3^{tyr705} expression in patients from the Glasgow combined cohort with right-sided (A), left-sided (B), and rectal disease (C). In patients with right-sided colonic tumours the hazard ratio for pSTAT3^{tyr705} expression was 1.188, (95%CI; 0.812-1.738), log rank p=0.372 and in left-sided colon cases the hazard ratio was 1.138, (95%CI; 0.738-1.755), log rank p=0.557. In rectal cases the hazard ratio for pSTAT3^{tyr705} was 1.778, (95%CI; 0.956-3.306, log rank p=0.065.*

Chi-squared tests were performed to determine any associations between $pSTAT3^{tyr705}$ expression and clinicopathological characteristics. High expression of $pSTAT3^{tyr705}$ was significantly associated with higher age (p=0.015), rectal disease subsite (p=0.011), latent and stromal phenotypic subtypes (p<0.001) and lower systemic inflammation as measured by modified Glasgow prognostic score (p=0.014) (Table 6.3).

Clinical Factor	Nuclear pSTAT3^{tyr705} Expression Low (n=442)	High (n=399)	p
Age			
<65	115 (27.8)	150 (35.0)	**0.015**
>65	298 (72.2)	278 (65.0)	
Sex	210 (50.8)	204 (47.7)	0.196
Female	203 (49.2)	224 (52.3)	
Male			
T Stage	23 (5.6)	15 (3.5)	0.391
I	56 (13.6)	50 (11.7)	
II	218 (52.8)	239 (55.8)	
III	116 (28.1)	124 (29.0)	
IV			
N Stage	267 (65.3)	254 (59.3)	0.133
0	95 (23.2)	125 (29.2)	
I	47 (11.5)	49 (11.4)	
II			
Tumour subsite			
Right-sided colon	68 (44.7)	87 (41.6)	**0.011**
Left-sided colon	61 (40.1)	74 (35.4)	
Rectum	23 (15.1)	48 (23.0)	
GMS			
0	138 (33.9)	137 (34.2)	0.260
1	212 (52.1)	192 (47.9)	
2	57 (14.0)	72 (18.0)	
Phenotypic Subtype			
1	138 (33.9)	139 (34.8)	**<0.001**
2	146 (35.9)	88 (22.0)	
3	66 (16.2)	102 (25.5)	
4	57 (14.0)	71 (17.8)	
mGPS			
0	180 (54.1)	184 (53.3)	**0.014**
1	83 (24.9)	113 (32.8)	
2	70 (21.0)	48 (13.9)	
MMR status	339 (82.3)	355 (83.1)	0.406
pMMR	73 (17.7)	72 (16.9)	
dMMR			
Tumour differentiation	369 (89.3)	384 (89.7)	0.474
0	44 (10.7)	44 (10.3)	
1			
Marginal involvement	392 (94.9)	405 (94.6)	0.487
0	21 (5.1)	23 (5.4)	
1			
Vascular invasion	271 (65.6)	292 (68.2)	0.233
0	142 (34.4)	136 (31.8)	
1			

Table 6.3 Nuclear pSTAT3^{tyr705} Expression and Clinical Features. *Table of Chi-squared associations between nuclear pSTAT3^{tyr705} expression and clinicopathological prognostic features including age, sex, T stage, N stage, tumour subsite, GMS, phenotypic subtype, mGPS, MMR status, tumour differentiation, marginal involvement, and vascular invasion.*

6.2.4 Expression of pSTAT3^{727} and clinical outcome

Expression of pSTAT3^{ser727} was assessed in the Glasgow combined colorectal array by immunohistochemistry. The staining pattern was mainly cytoplasmic within the tumour with very few patients expressing pSTAT3^{ser727} in the nucleus. Cytoplasmic staining represents that the protein has been active, whereas nuclear staining as in the case of pSTAT3^{tyr705} indicates that the protein is active. Given the frequency of patients expressing cytoplasmic STAT3^{ser727} versus nuclear expression, it was deemed more appropriate to score the cytoplasmic staining. Representative images showing weak, moderate, and strongly stained TMA cores are shown in figure 6.29 (A-C). Examples of true positive (spleen) and true negative (tonsil) tissue shown in 6.29 (D-E). Positive control colorectal tissue identified during antibody optimisation is and negative control colorectal tissue with no primary antibody added during overnight incubation are also shown in representative images (Figure 6.29) (F-G). The pSTAT3^{ser727} antibody specificity was validated by single band at the correct molecular weight on western blot in Sw620 cell lysates treated with 0.01% DMSO and no bands detected in JAK1/2 inhibited cell lysates (Figure 6.29). Further specificity was proven via staining COLO205 (treated/untreated) cell pellets for pSTAT3^{ser727} (Figure 6.29) (H).

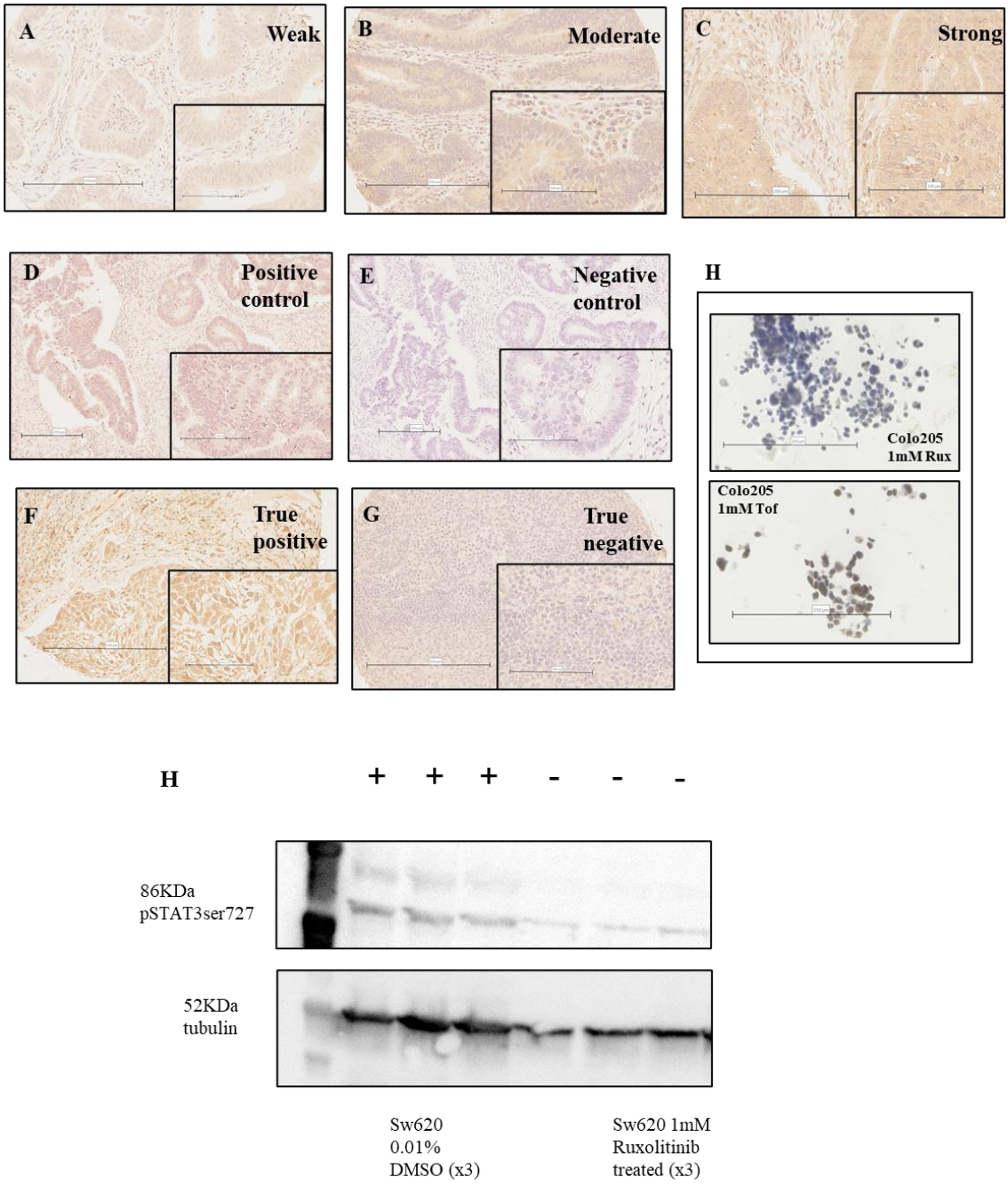


Figure 6.29 Representative images and antibody specificity for pSTAT3^{ser727}. *Images showing representative images of weak (A), moderate (B) and strong (C) cytoplasmic staining of pSTAT3^{ser727} within tumour cells. Representative images of positive (D) and negative (E) colorectal tissue (F), and true positive tissue(G). COLO205 cell pellets treated with vehicle control or JAK1/2 inhibitor and then stained for pSTAT3^{ser727}(H). Image of western blot showing single band in triplicate at 86KDa for SW620 cell lysates (20μL loaded per well) treated with vehicle control probed for pSTAT3^{ser727} and no visible bands in triplicate for pSTAT3^{tyr705} inhibited SW620 cells (20μL loaded per well) at 86KDa, and tubulin detected in all samples at 52KDa (H).*

Cytoplasmic expression of pSTAT3^{ser727} was manually scored by a single observer (KP) in 622 patients from the Glasgow combined cohort. Of these patients, 102 were excluded due to administration of neoadjuvant therapy or mortality within 30 days of surgery leaving 520 patients included in downstream analysis. Manual weighted histoscores ranged from 0 to 232.72 with a mean score of 120.66. A histogram was plotted to assess the distribution pattern of the data, which showed a normal distribution (Figure 6.30). Digital scoring was performed using QuPath to validate manual scoring in 10% of TMA cores. A scatter plot was constructed to visualise correlation between scores and a correlation coefficient of 0.725 was obtained (Figure 6.31). A bland Altman was plotted to show minimal variation between manual and digital scores (Figure 6.32). An intra-class correlation coefficient of 0.778 indicated the scores were not significantly different. Cut offs for high and low expression were determined by log rank statistics using survminer and maxstat in R studio which devised an optimum threshold of 146.02. Density and scatter plots were constructed to visualise the optimal cut point (Figure 6.33). This resulted in 159 patients classified as high for pSTAT3^{ser727} expression and 362 as low expression of pSTAT3^{ser727}.

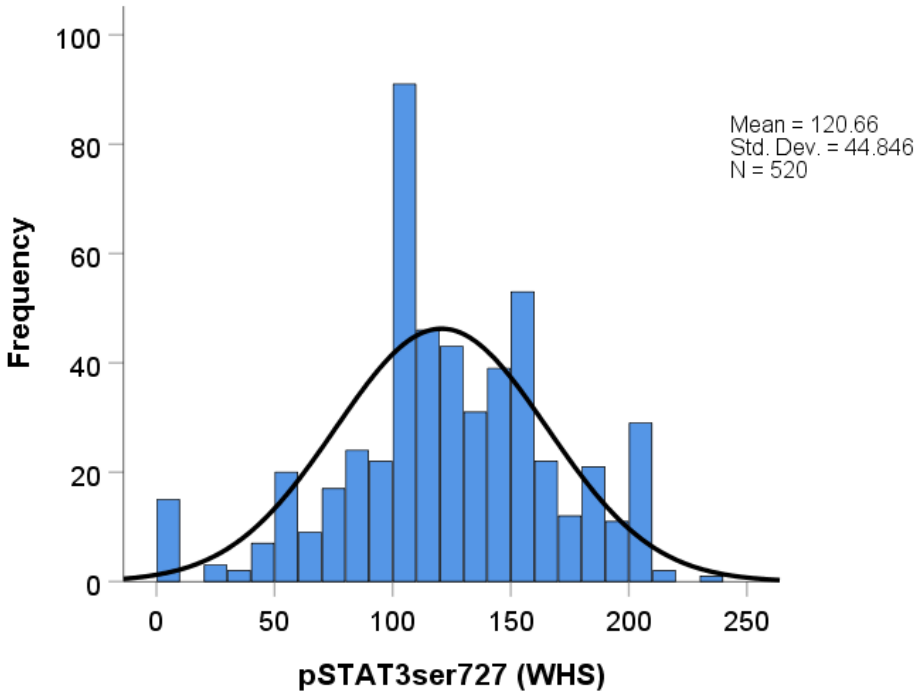


Figure 6.30. Distribution of weighted histoscores for nuclear pSTAT3^{ser727}. *Histogram showing the range of scores obtained for nuclear tumour pSTAT3^{ser727} expression and normal distribution pattern of the data (n=520). The mean score for pSTAT3^{ser727} was 120.66 and scores ranged from 0-232.72.*

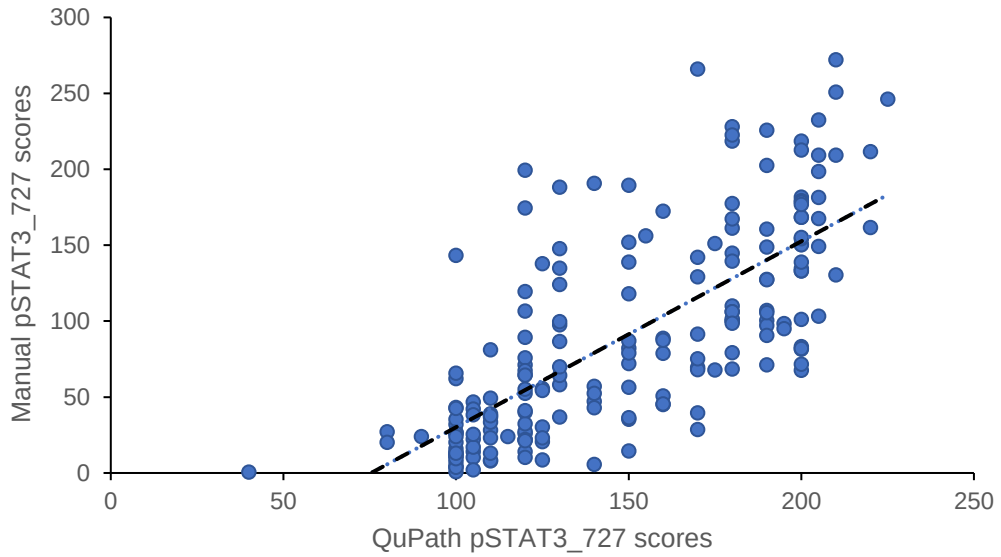


Figure 6.31 Correlation between manual and digital weighted histoscoring of pSTAT3^{ser727}. *Scatter plot showing correlation between tumour cytoplasm weighted histoscores for pSTAT3^{ser727} in 10% of cases from the Glasgow combined cohort co-scored for validation. A correlation coefficient of 0.725 was obtained.*

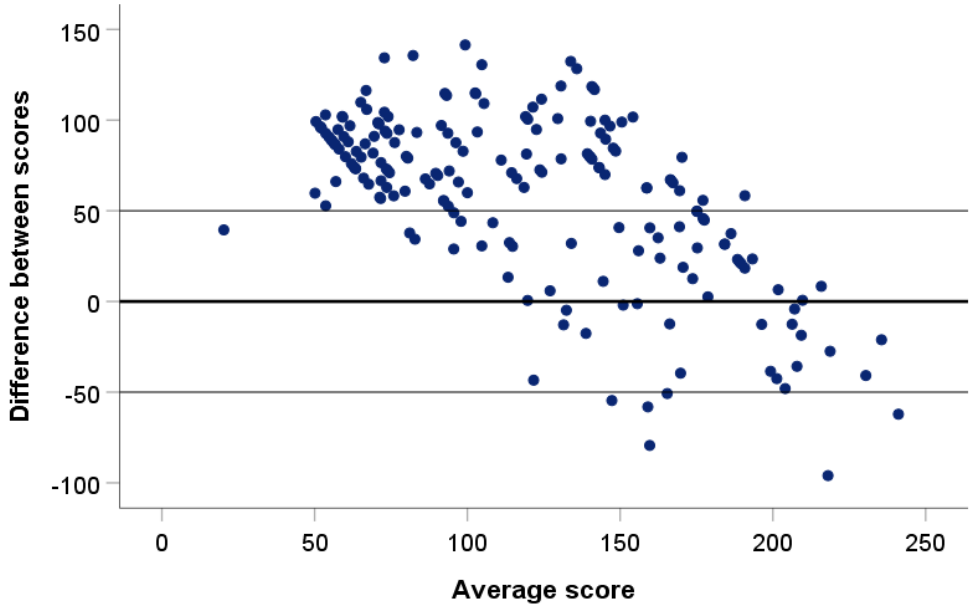


Figure 6.32 Validation of pSTAT3^{ser727} manual scoring. *Bland Altman plot showing difference between manual and digital QuPath scores for pSTAT3^{ser727} in tumour cytoplasm in the 10% of cases from the Glasgow combined array co-scored for validation.*

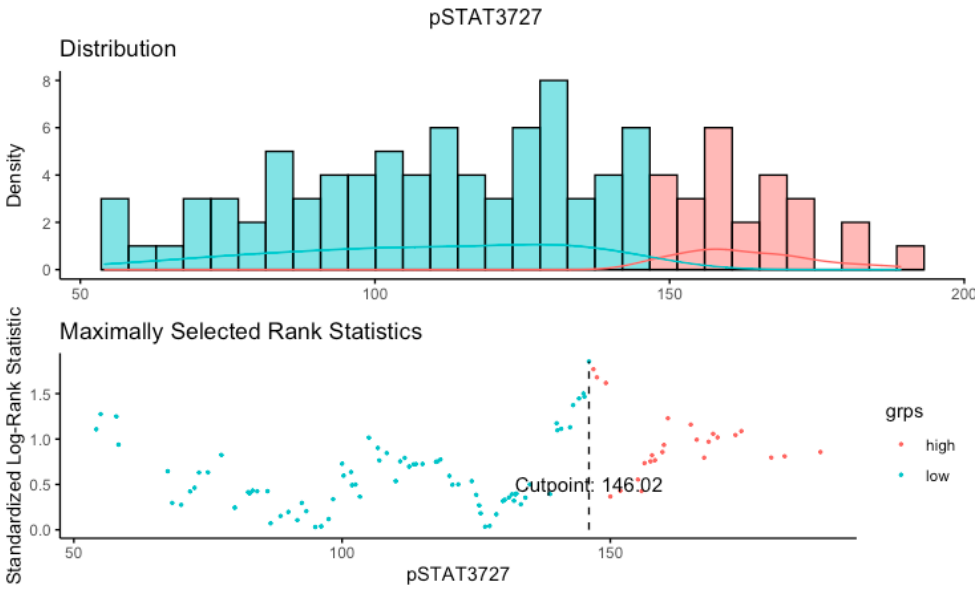


Figure 6.33 Defining cut off point for nuclear pSTAT3^{ser727} expression high and low groups. *Density plot and scatter plot for visualisation of optimal cut off point for high and low expression of tumour nuclear pSTAT3^{ser727}. The optimal cut-off point determined was 146.02 and therefore patients with a weighted histoscore of ≤146.02 were considered low for pSTAT3^{ser727} and patients with a weighted histoscore of ≥146.02 were classified as high for pSTAT3^{ser727} expression.*

Kaplan Meier curves were plotted to determine any association between pSTAT3^{ser727} expression and cancer-specific survival in the full cohort and in patient subsets. When analysed with respect the full cohort, pSTAT3^{ser727} was not associated with cancer-specific survival (HR=0.895, 95%CI;0.629-1.274, log rank p=0.535) (Figure 6.34). When stratified by histological measure GMS, cytoplasmic pSTAT3^{ser727} expression and CSS failed to reach statistical significance in any group (Figure 6.35). Similarly, when stratified by MMR status no association of cytoplasmic pSTAT3^{ser727} expression was observed for MMR proficient or MMR deficient patients (Figure 6.36). There was no significant association between pSTAT3^{ser727} expression and CSS with respect to disease subsite (Figure 6.37).

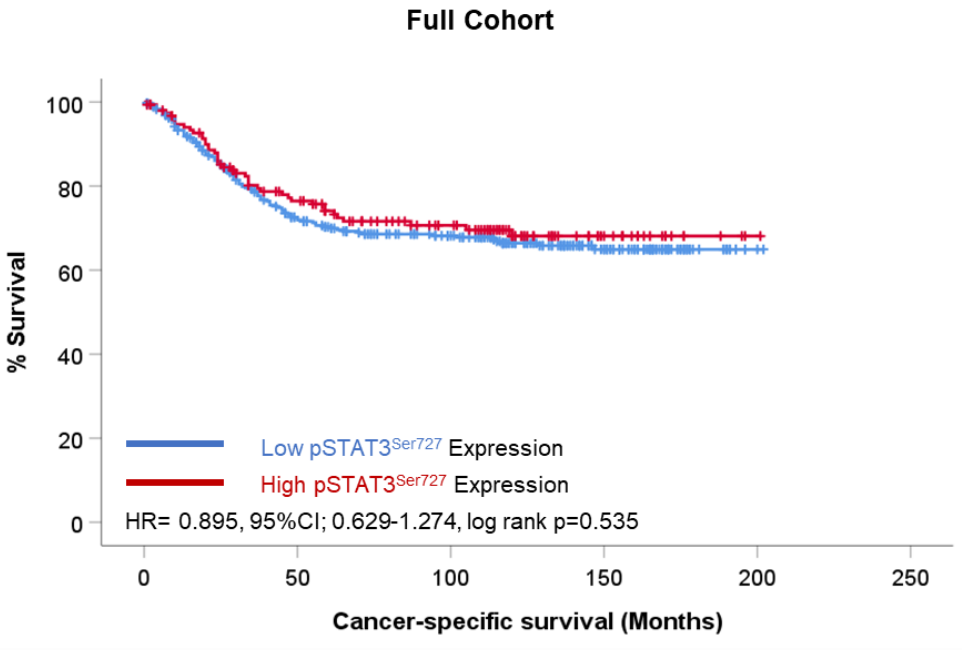


Figure 6.34 Nuclear expression of pSTAT3^{ser727} expression and CSS. *Kaplan Meier survival analysis of nuclear pSTAT^{tser725} expression in patients from the Glasgow combined cohort. In the full cohort the hazard ratio associated with pSTAT3ser727 expression was 0.895, (95%CI; 0.629-1.274), log rank p= 0.535.*

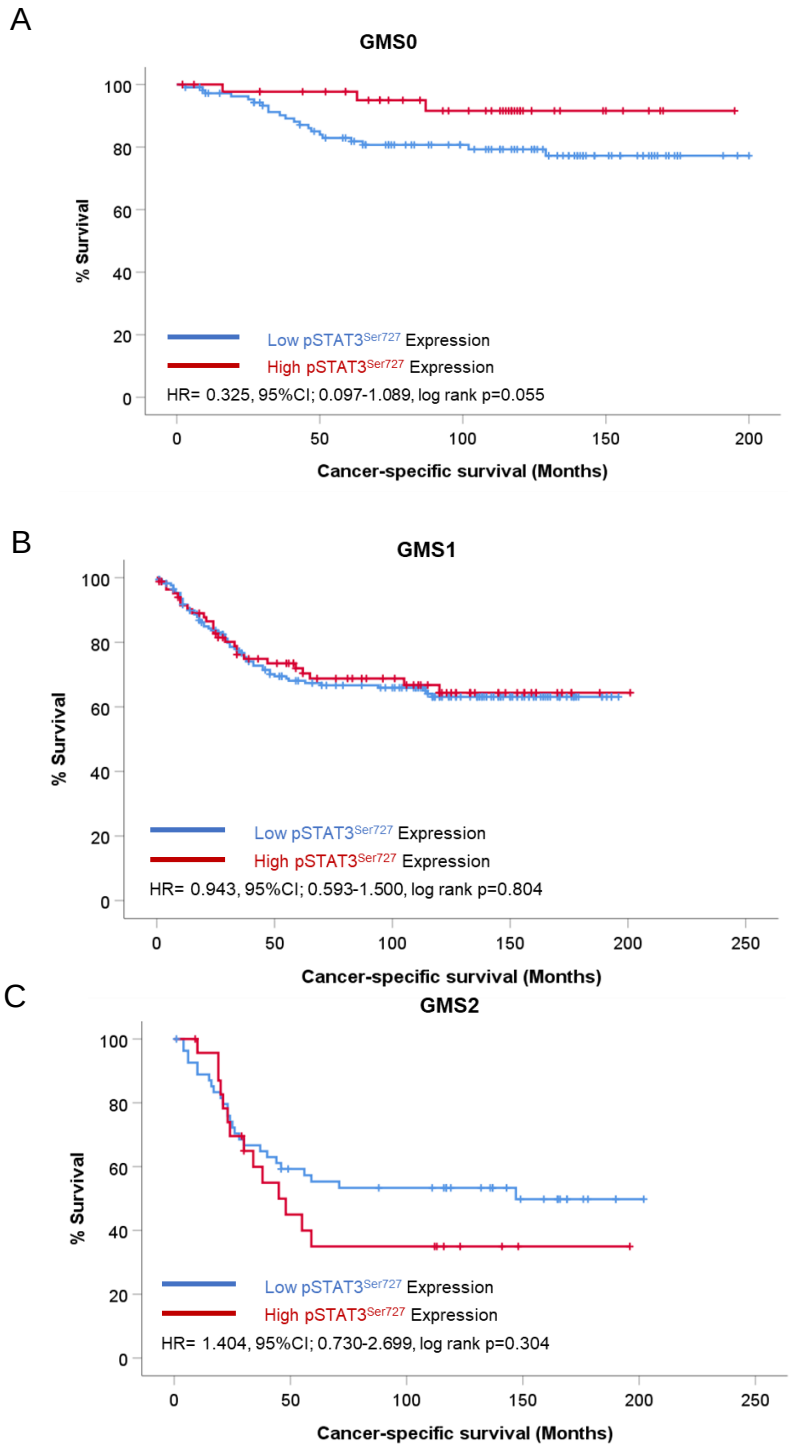


Figure 6.35 Nuclear expression of pSTAT3^{ser727} expression and CSS, *Kaplan Meier survival analysis of nuclear pSTAT3^{ser727} expression in patients from the Glasgow combined cohort classified as GM0 (A), GMS1 (B), and GMS2 (C). Patients with GMS0 immune graded tumours observed a hazard ratio of 0.325, (95%CI; 0.097-1.089), log rank p=0.055 for pSTAT3^{ser727} expression. In patients with GMS1 tumours the hazard ratio for pSTAT3^{ser727} expression was 0.943, (95%CI; 0.593-1.500), log rank p=0.804 and for GMS2 stromal-rich tumours the hazard ratio was 1.404, (95%CI; 0.730-2.699), log rank p=0.304.*

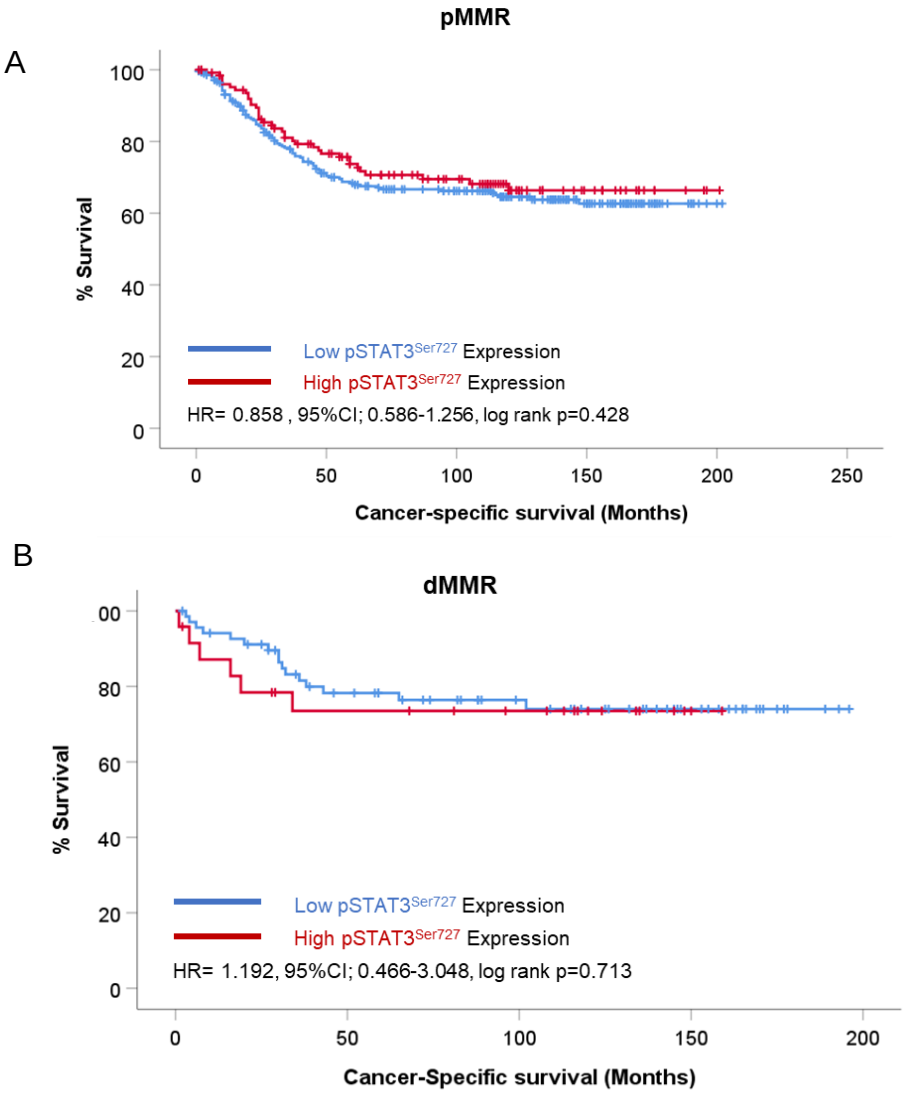


Figure 6.36 Nuclear expression of pSTAT3^{ser727} expression and CSS relative to MMR status. *Kaplan Meier survival analysis of nuclear pSTAT3^{ser727} expression in patients from the Glasgow combined cohort with MMR proficient tumours (A) and MMR deficient tumours (B). In MMR proficient cases the hazard ratio for pSTAT3^{ser727} expression was 0.858, (95%CI; 0.586-1.256), log rank p= 0.428 and in MMR deficient cases the hazard ratio was 1.192, (95%CI; 0.466-3.048), log rank p=0.713.*

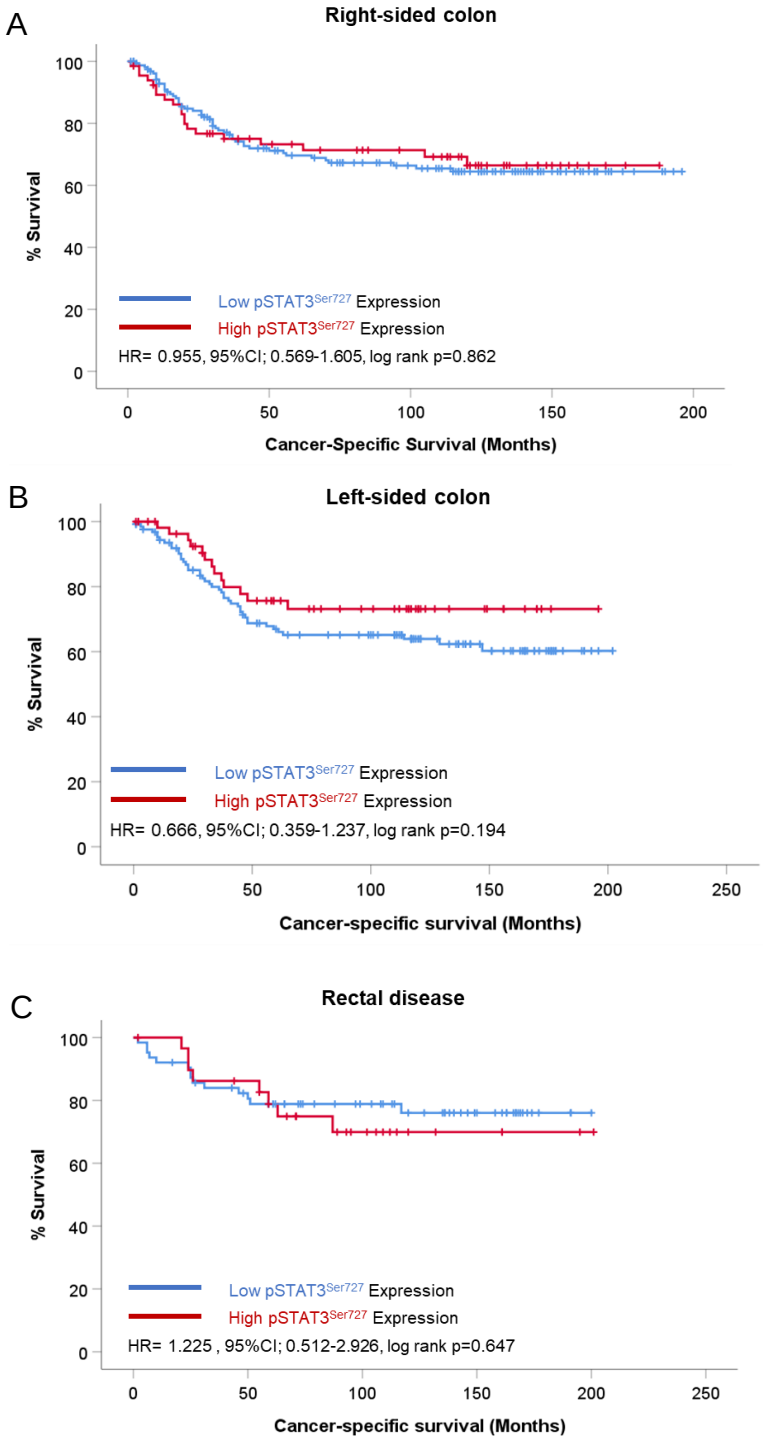


Figure 6.37 Nuclear expression of pSTAT3^{ser727} expression and CSS relative to tumour subsite. *Kaplan Meier survival analysis of nuclear pSTAT3^{ser727} expression in patients from the Glasgow combined cohort with right-sided (A), left-sided (B), and rectal disease (C). In patients with right-sided colonic tumours the hazard ratio for pSTAT3^{ser727} was 0.955, (95%CI; 0.569-1.605), log rank p=0.862 and in left-sided cases this was 0.666, (95%CI; 0.359-1.237, log rank p=0.194. In patients with rectal tumours the hazard ratio for pSTAT3^{ser727} was 1.225, (95%CI; 0.512-2.926), log rank p=0.647.*

Chi-squared tests were performed to determine any association between pSTAT3^{ser727} expression and clinicopathological features. No significant associations were observed in this cohort (Table 6.4). These data validate other observations in the literature which suggest phosphorylation at tyrosine 705 is a more important activation site than pSTAT3^{ser727}.

Clinical Factor	Cytoplasmic pStat3^{ser727} Expression Low (n=362)	High (n=159)	p
Age			
<65	98 (27.1)	42 (26.4)	0.483
>65	264 (72.9)	117 (73.6)	
Sex			
Female	167 (46.1)	86 (54.1)	0.057
Male	195 (53.9)	73 (45.9)	
T Stage			
I	217 (4.7)	7 (4.4)	0.892
II	47 (13.0)	24 (15.1)	
III	193 (53.3)	80 (50.3)	
IV	105 (29.0)	48 (30.2)	
N Stage			
0	223 (62.1)	109 (68.6)	0.350
I	92 (25.6)	35 (22.0)	
II	44 (12.3)	15 (9.4)	
Tumour subsite			
Right-sided colon	168 (46.8)	68 (43.3)	0.731
Left-sided colon	128 (35.7)	58 (36.9)	
Rectum	63 (17.5)	31 (19.7)	
GMS			
0	115 (33.0)	49 (31.0)	0.860
1	176 (50.6)	84 (53.2)	
2	57 (16.4)	25 (15.8)	
Phenotypic Subtype			
1	116 (33.3)	49 (31.2)	0.832
2	118 (33.9)	52 (33.1)	
3	57 (16.4)	31 (19.7)	
4	57 (16.4)	25 (15.9)	
mGPS			
0	132 (46.8)	67 (56.3)	0.197
1	95 (33.7)	31 (26.1)	
2	55 (19.5)	21 (17.6)	
MMR status			
pMMR	291 (80.4)	133 (84.7)	0.147
dMMR	71 (19.6)	24 (15.3)	
Tumour differentiation	324 (89.5)	146 (91.8)	0.258
0	38 (10.5)	13 (8.2)	
1			
Marginal involvement	341 (94.2)	150 (94.3)	0.565
0	21 (5.8)	9 (5.7)	
1			
Vascular invasion	242 (66.9)	105 (66.0)	0.466
0	120 (33.1)	54 (34.0)	
1			

Table 6.4 Nuclear pSTAT3^{ser727} Expression and Clinical Features. *Table of Chi-squared associations between nuclear pSTAT3^{ser727} expression and clinicopathological prognostic features including age, sex, T stage, N stage, GMS, phenotypic subtype, MMR status, mGPS, tumour differentiation, marginal involvement and venous invasion.*

6.2.5 Expression of pSTAT3^{705} and pSTAT3^{727} and clinical outcome

There is evidence from mouse data that for maximum pathway activation, phosphorylation at both tyrosine 705 and serine 727 is necessary(71). Therefore, scores for both sites were combined to form 2 new groups: Low for both pSTAT3^{tyr705} and pSTAT3^{ser727} or high for 1 marker, and high for both markers. Kaplan Meier survival analysis was performed to determine any association between expression of both pSTAT3^{tyr705} and pSTAT3^{ser727} and CSS. No significant difference in survival time was observed (p=0.733) (Figure 6.38). However, when analysed with respect to GMS, patients with stromal-rich GMS2 tumours had significant worse outcomes if they were high for both pSTAT3^{tyr705} and pSTAT3^{ser727} expression (HR=2.758, 95%CI; 1.183-6.429, log rank p=0.014) (Figure 6.39). The mean survival time of patients with high expression for both markers was profoundly low at 38.969 months compared to 126.007 months for patients with both low/1 high.

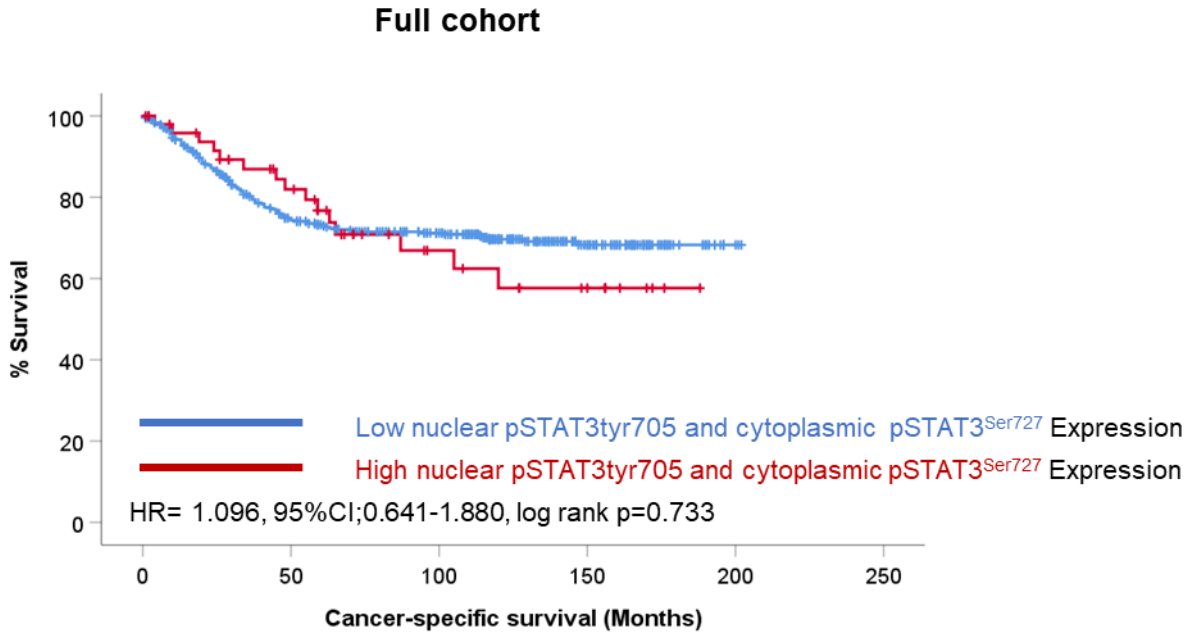


Figure 6.38 Nuclear expression of pSTAT3^{tyr705} and cytoplasmic pSTAT3^{ser727} expression and CSS relative to tumour subsite. *Kaplan Meier survival analysis of nuclear pSTAT3^{tyr705} and cytoplasmic pSTAT3^{ser727} expression in patients from the Glasgow combined cohort. In the full cohort a combined score of pSTAT3^{ser727} and pSTAT3^{tyr705} yielded a hazard ratio of 1.096, (95%CI 0.641-1.880), log rank p=0.733.*

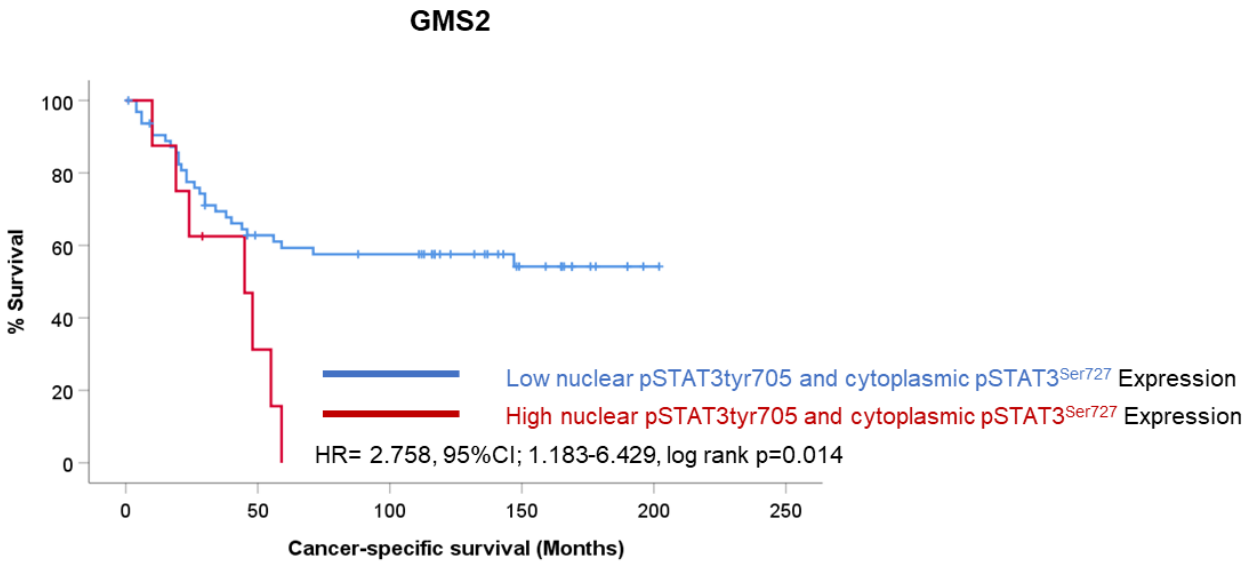


Figure 6.39 Nuclear expression of nuclear pSTAT3^{tyr705} and cytoplasmic pSTAT3^{ser727} expression and CSS relative to tumour subsite. *Kaplan Meier survival analysis of nuclear pSTAT3^{tyr705} and cytoplasmic pSTAT3^{ser727}expression in patients from the Glasgow combined cohort with GMS2 classified tumours. In GMS2 stromal-rich cases the hazard ratio for a combined pSTAT3^{tyr705} and pSTAT3^{ser727} score was 2.758, (95%CI; 1.183-6.429), log rank p=0.014.*

6.4 Discussion

Colorectal cancer is characterised by dysregulation of numerous cellular signalling pathways. Constitutive activation or overproduction of upstream signals of certain transcription factors results in pro-tumour gene transcription. STAT3 is an example of a master regulator that can become hyperactivated in colorectal cancer (145). In its inactive state, STAT3 is located within the cytoplasm and when activated it dimerizes and translocates to the nucleus where it can promote transcription of a variety of cancer-promoting genes (142). Dysregulation of STAT3 transcriptional activity in CRC is associated with many of the hallmarks of cancer including angiogenesis, proliferation, invasion, migration, and differentiation (146). More recently, it has also been shown that STAT3 can support tumourogensis irrespective of its transcriptional activity (9). For example, STAT3 activation can modulate the TME to an immunosuppressive phenotype by repressing production of inflammatory cytokines including type I interferons(147). This has been corroborated by recent tissue-based histology studies, with one study showing expression of STAT3 within colorectal cancer specimens associates with unfavourable clinical characteristics and reduced inflammatory infiltrates, specifically T lymphocytes (148). STAT3 also has roles in altering mitochondrial function, blocking autophagy and promotion of tumour cell motility leading to metastases (9, 149, 150).

In this chapter it was observed that high expression of cytoplasmic STAT3 in tumour cells was associated with reduced CSS (HR=1.384, 95%CI; 1.040-1.841, log rank p=0.025). In contrast to hypotheses, nuclear STAT3 was not associated with CSS in this cohort overall, although in patients with GMS1 tumours specifically, high expression was associated with worse outcome (HR=1.729, 95%CI; 1.128-2.652, log rank p=0.011). Although STAT3 nuclear localisation is an indication of pathway activation, perhaps a more conclusive measure is expression of predominant phosphorylation site, pSTAT3^{tyr705}. In published work investigating pSTAT3^{try705} expression in retrospective CRC cohorts, high expression has previously been shown to associate with cancer-specific and overall survival (143). Data from this chapter investigating pSTAT3^{tyr705} expression in the Glasgow combined CRC array identified a profound association between high nuclear expression and reduced CSS in patients with GMS2 stromal dense tumours (HR= 2.086, 95%CI; 1.224-3.554, log rank p = 0.006). High expression of pSTAT3^{tyr705} was associated with unfavourable phenotypic subtypes latent and stromal by chi-squared analysis ($p<0.001$).

In addition to canonical activation at tyrosine 705, STAT3 can undergo post-translational modifications to enhance its transcriptional action via phosphorylation at serine 727 (71). Murine models have indicated phosphorylation at both sites is necessary for maximal activation, however there is very limited evidence for the role pSTAT3^{ser727} plays in CRC (71). In this chapter, we found no association between pSTAT3 ^{ser727} and outcome or clinical features. When pSTAT3 ^{ser727} and pSTAT3^{tyr705} expression were combined to form 1 or both low and both high groups, no association with CSS was observed in the full cohort but in GMS2 patient's survival outcomes were significantly worse in patients with high expression of both markers (HR=2.758, 95%CI; 1.183-6.429, log rank p=0.014). This indicates that maximal pathway activation in patients with stromal-rich tumours has a profound influence on survival time, which is highlighted by the drop in survival at ~50 months.

In terms of the inhibiting STAT3 activation therapeutically, one of the main mechanisms proposed is through inhibition of upstream JAKs. This is due to the similarity in structural shape of STAT3 with other STAT family members STAT1 and STAT5, making it difficult to target STAT3 specifically. There are a number of JAK inhibitors in ongoing clinical trials in numerous solid tumour and haematological disorders. There are ongoing phase 3 clinical trials investigating the use of JAK2 specific inhibitor Paracritinib for myeloproliferative disorders. Baracitinib (JAK1/2 inhibitor) was approved in 2016 for RA after success in the RA-BUILD trial(151). Fedratinib (JAK2 inhibitor) has recently been approved for use in myeloproliferative neoplasm associated myelofibrosis(152). Abrocitinib (JAK1 inhibitor) is currently in phase 2 and 3 clinical trials for patients with atopic dermatitis (153, 154). The number of clinical trials highlights the broad spectrum of JAK inhibitors available, and future work could include screening all the repurposable Jakinibs in a more high-throughput model to identify the optimal JAK inhibitor for CRC patients and determining the optimal JAK inhibitor for specifically inhibiting STAT3 signalling.

In conclusion, the main findings of this chapter were that GMS2 patients had significantly reduced survival time if classified as high for pSTAT3^{tyr705} expression alone, or high for both pSTAT3 ^{ser727} and pSTAT3^{tyr705}. Taken together these data, coupled with results from previous chapters, suggest patients with stromal-rich GMS2 tumours would be the subgroup most likely to benefit from therapeutic inhibition of STAT3 activation. Repeating the staining in a validation cohort would be necessary to confirm the results from this discovery cohort. Data from this, and the two prior chapters have implicated IL6/JAK/STAT3 in CRC, however the underlying genomics driving over-expression of pathway members at the protein level is not yet known and may be of importance.

Chapter 7: Assessment of the mutational landscape associated with Glasgow Microenvironment Score and STAT3 activation in clinical colorectal cancer specimens

7.1 Introduction

Data from previous chapters have highlighted a prognostics role for Glasgow microenvironment score and the IL6/JAK/STAT3 pathway in CRC, however the underlying genomics driving these phenotypes unknown. In this chapter the mutational background of patients was assessed relative to GMS and phosphorylated STAT3.

Colorectal cancers are heterogenous with numerous genetic alterations associated with oncogenesis. Most CRCs arise sporadically, however ~4-10% of cases are of hereditary origin (155). Patients with familial adenomatous polyposis (FAP) exhibit germline mutations in the most frequently mutated CRC-associated gene, Adenomatous polyposis coli (APC) (155). FAP patients have a lifetime risk of 100% for developing CRC and account for ~1% of CRC patients (155). Hereditary non-polyposis colorectal cancer (HNPCC) or Lynch syndrome accounts for 3-10% of CRCs and results from germline mutations in one of 5 mismatch repair (MSI genes, MLH1, MSH2, MSH6, PMS1 and/or PMS2 (156), unlike sporadic cases Lynch syndrome tumours with mutations in the MSI genes do not have BRAF V600E mutation. Patients with HNPCC have ~80% lifetime risk of developing CRC (156). The high risk of malignancy associated with these inherited cases of CRC highlights the importance of mutational profiles. In addition to mutations observed in hereditary cancers, somatic mutations in many different genes are observed in non-hereditary CRC, with a high degree of heterogeneity between patients. Perhaps the most well-defined genomic alteration linked to CRC is the chromosomal instability (CIN) pathway (157). Development and progression follow a pattern of mutation acquisition which corresponds to distinct histological changes. Loss of tumour-suppressor gene APC results in formation of an early adenoma from normal tissue (157, 158). APC mutations are found in 50-83% of sporadic CRCs, and initiate cancer progression via enabling β-catenin accumulation in enterocytes which leads to sustained activation of WNT signalling (158, 159). APC mutation status in sporadic cases of CRC is not itself associated with prognosis (160). APC loss leads to activation mutations in KRAS and subsequently late adenoma establishment. Alterations in KRAS are generally associated with poorer outcomes in CRC (160) .The final stage of the CIN pathway involves loss of p53 resulting in development of an adenocarcinoma via disruption to DNA repair processes and apoptotic cell death (157). In addition to CIN, MSI instability via methylation or mutation of the MLH1 gene is also strongly linked to CRC oncogenesis (161). Somatic mutations in many other genes have been linked to CRC progression. Research has aimed to identify mutational signatures that predict patient prognosis and treatment response. However, next generation sequencing (NGS) technologies are not used routinely in clinical laboratories due to extensive costs and time consuming protocols (8). A more translatable method for using mutational profiling to enhance therapeutics may be to link mutational signatures to distinct, easily conducted phenotypic measures like GMS, phenotypic subtype, or prognostic IHC-based protein staining such as for $pSTAT3^{tyr705}$.

GMS is now an extensively validated prognostic measure of tumour phenotype in colorectal cancer(47, 103, 108, 162). Due to the number of patients with mutation profile data available, the 3 GMS groups were analysed for mutational backgrounds as opposed to 4 phenotypic subtype groups (n=252). The aim of this analysis was to determine any patterns of genetic alterations associated with driving immune (GMS0), intermediate (GMS1) and stromally-dense (GMS2) phenotypes.

Data from the Glasgow combined cohort in chapters 3, 4 and 5 suggest that IL6-STAT3 signalling is an important inflammatory pathway involved in CRC oncogenesis, particularly in the stromal GMS2 subtype. To investigate the underlying biology of patients with STAT3 activation, mutational landscaping was performed in a subset of patients from the Glasgow combined cohort with high/low $pSTAT3^{tyr705}$ phenotypic status available (n=207). The aim of this analysis was to determine a pattern of mutation associated with aberrant STAT3 signal transduction and to identify potential genes suitable for dual therapeutic targeting with JAK-STAT3 inhibitors for CRC patients. It was hypothesised that a distinct mutational pattern would be associated with each GMS group, and with respect to high versus low $pSTAT3^{tyr705}$ protein groups.

Finally, the Cancer Genome Atlas (TCGA) PanCancer CRC cohort (n=594) was analysed for alterations in IL6-STAT3 pathway members and association with outcomes/clinicopathological features. It was hypothesised that alterations in IL6/JAK/STAT3 pathway genes would be rare but associated with unfavourable clinical characteristics and reduced survival times.

7.2 Mutational landscape of colorectal cancer patients

7.2.1 Mutational landscape in a subset of the Glasgow combined cohort

Mutational landscaping was performed by Glasgow Precision Oncology Laboratory in a subset of patient specimens from the Glasgow combined cohort (n=252). Bulk whole exome sequencing using a custom panel of genes designed in-house enabled analysis of 151 cancer-associated genes of interest. The median number of variants was 7 alterations per patient. Most (42.51%) mutations were missense single nucleotide polymorphisms with cytosine to thymine transitions the most frequent (58.49%) SNV class. To visualise the most frequently occurring mutations an oncoplot was constructed using the 'oncoplot' function in maftools package for R Studio (Figure 7.2). Genomic mutations in one or more of the genes in the custom panel were identified in 99.21% of the patient's tumours investigated. Lynch syndrome patients were identified by combined presence of alteration/s in ≥ 1 MSI gene and absence of BRAF mutation (n= 17, (6.74%))(163). A summary plot was created in maftools using the 'plotmafSummary' function for R studio (Figure 7.1). APC was mutated in 189 (75%) of patients,

TP53 in 60% of patients and KRAS in 48% of patients. Pathway enrichment analysis showed 83% of patients with alterations in the RTK-RAS pathway, 88% in the WNT pathway and 72% in the TP53 signalling pathway (Figure 7.3). These data are reflective of broad mutational patterns of CRC observed in recent scientific literature confirming the validity of the NGS performed in this cohort (164).

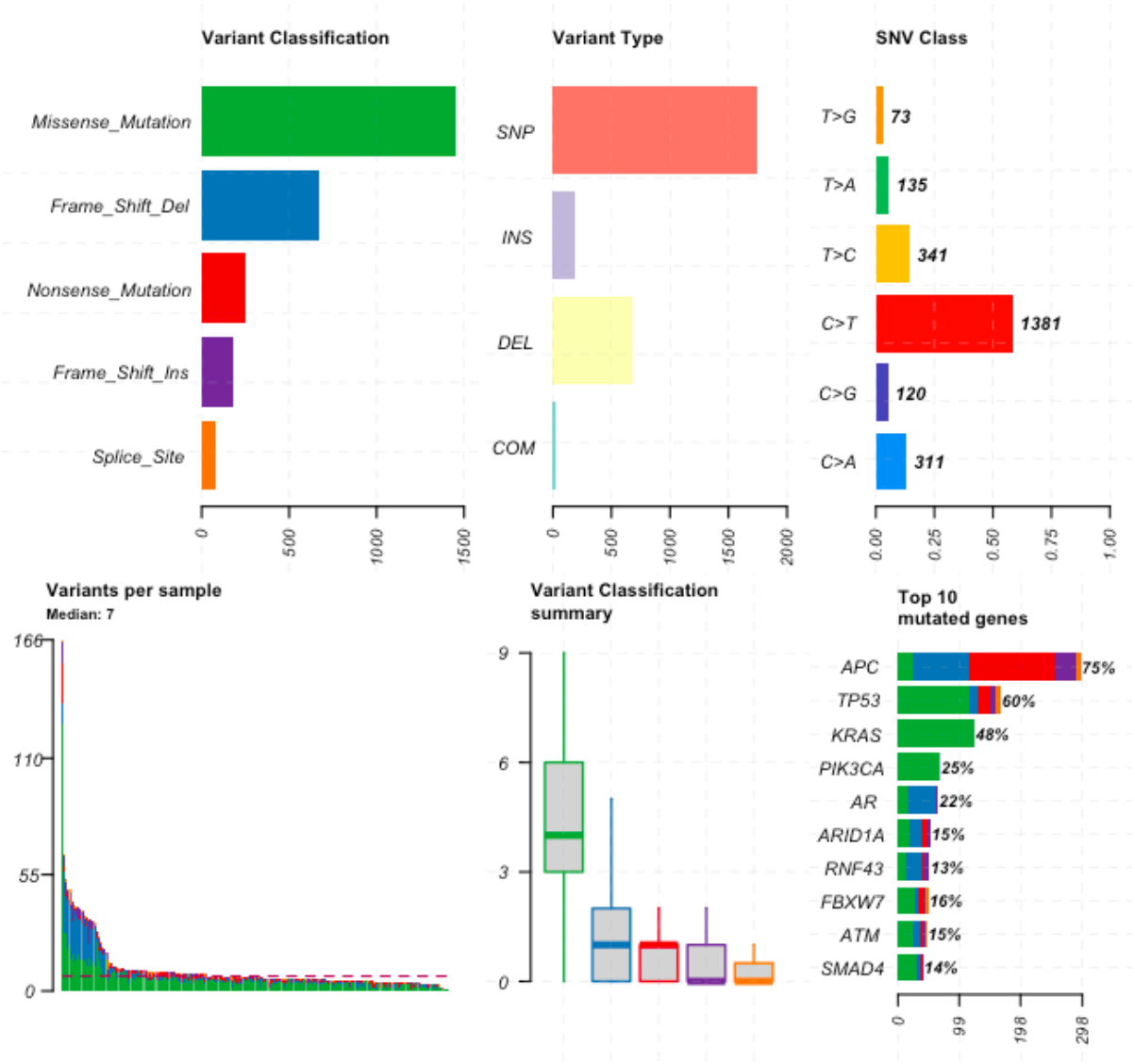


Figure 7.1 Overview of cohort mutational background. *Summary plot showing the distribution of variant classifications, variant types, SNV classes, mutation burden and top 10 mutations within the subset of 252 patients from the Glasgow combined cohort created using maftools in R studio version 1.3.*

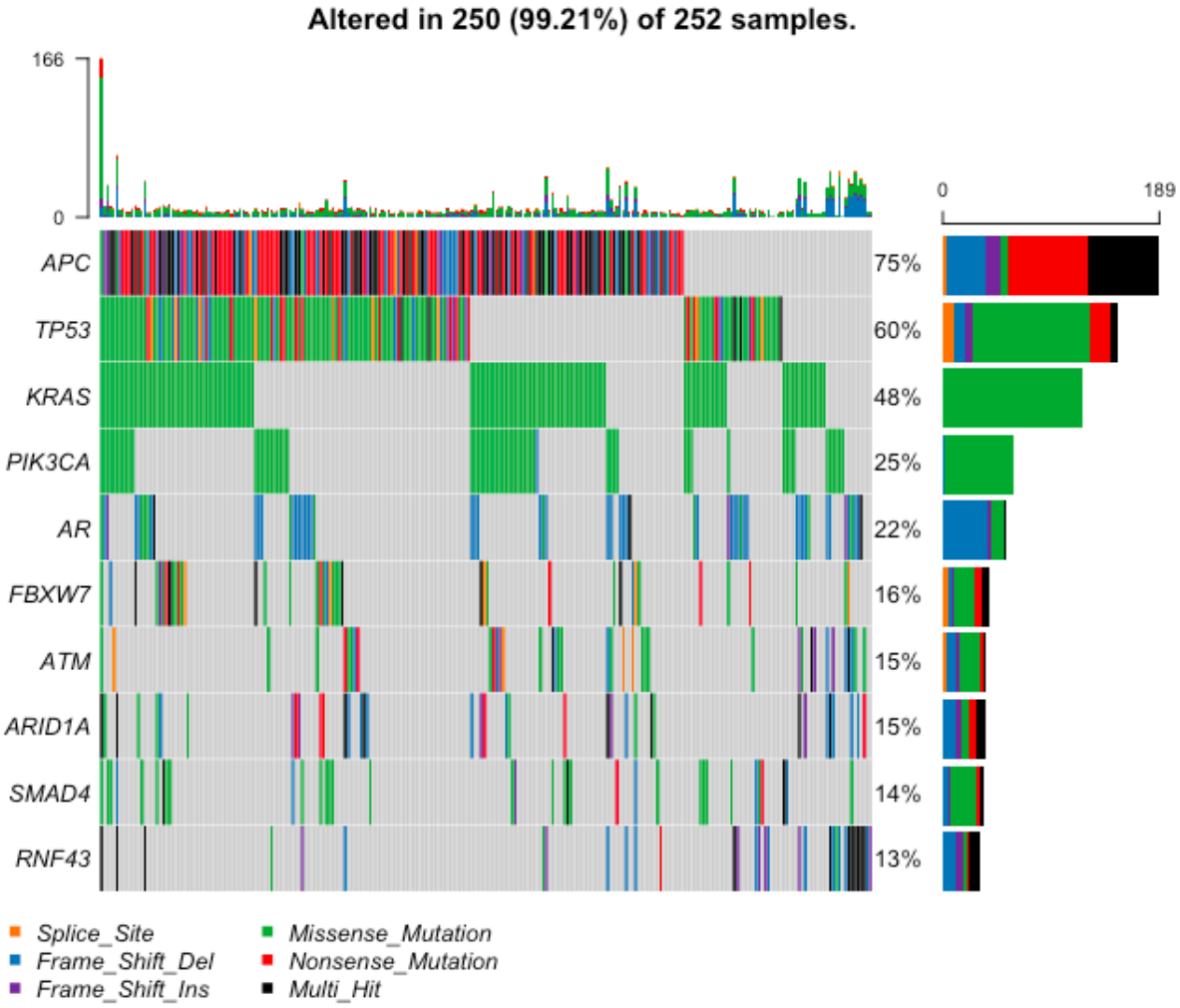


Figure 7.2 Mutational landscape in a subset of patients from the Glasgow combined cohort. *Oncoplot showing the top 10 mutated genes in 252 patient samples from the stage I-IV CRC Glasgow combined cohort created using maftools in R studio version 1.3 from custom 151-gene panel sequencing run by Glasgow Precision Oncology Laboratory. In the full cohort APC was mutated in 75% of cases, TP53 in 60% and KRAS in 48% of patients. Mutation types were a mix of missense mutations, nonsense mutations, frameshift insertion, frameshift deletion, splice site and multi-hit mutations.*

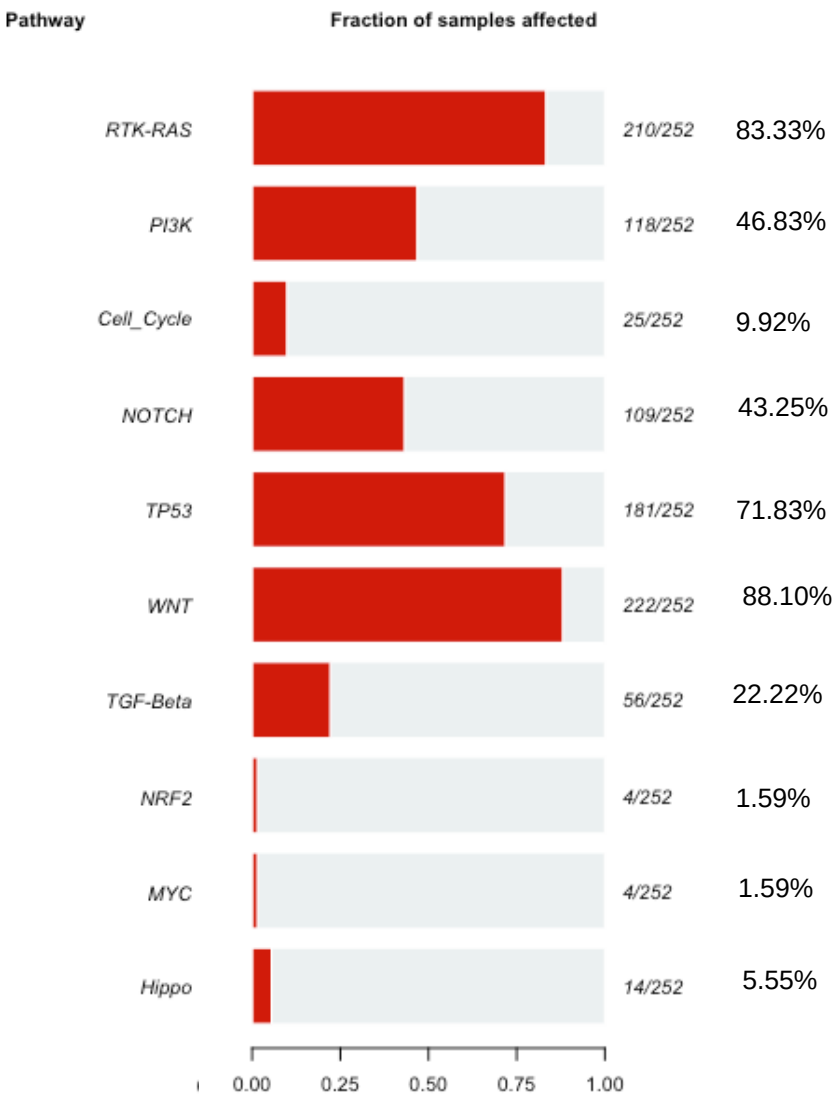


Figure 7.3 Pathway enrichment in a subset of patients from the Glasgow combined cohort. *Bar chart showing common cancer-associated pathways and frequency of alteration in 252 patients from the Glasgow combined cohort created in R Studio version 1.3 using data generated from custom 151-gene panel sequencing outsourced and performed by Glasgow Precision Oncology Laboratory.*

7.3 Mutational landscape underlying colorectal cancer phenotypes

7.3.1 Mutational landscape and Glasgow Microenvironment Score

The mutational landscape of each GMS group was investigated in patients from the Glasgow combined cohort with sequencing data (n=252). In terms of Immune GMS0 (n=72) patients, the

median number of variants was 7 and the most common mutation types were missense single nucleotide polymorphisms (SNP) (39.36%) and cytosine to thymine transitions were the most frequently occurring SNV class (56.18%) (Figure 7.4). To visualise the top 10 most mutated genes amongst GMS0 patients an oncoplot was constructed (Figure 7.5). APC was mutated in 74% of patients, TP53 in 56% and KRAS in 46% of patients (Figure 7.5). Pathway enrichment analysis showed a high proportion of GMS0 patients with alterations in the WNT (87.5%), TP53 (72.22%) and RTK-RAS (77.77%) pathways (Figure 7.6). Fishers exact tests were used to determine any statistically significant differences in mutation patterns between GMS0 and all other patients (GMS1 and GMS2 combined). A total of 7 genes were found to be enriched in GMS0 (Figure 7.7). MSI genes MLH1 and MSH2 were altered more frequently in GMS0 patients (p=0.000253, p=00769 respectively) (Figure 7.7). Akt signalling pathway member PTEN was also more commonly mutated in GMS0 patients compared to low immune patients (p=0.014). ARID1A (chromatin remodelling, p=0.0166), RPL22 (ribosomal protein, p=0.0228), ERBB4/HER4 (EGFR family, p=0.025) and B2M (component of MHCI, p=0.033) genes were all more frequently mutated within the GMS0 group (Figure 7.7).

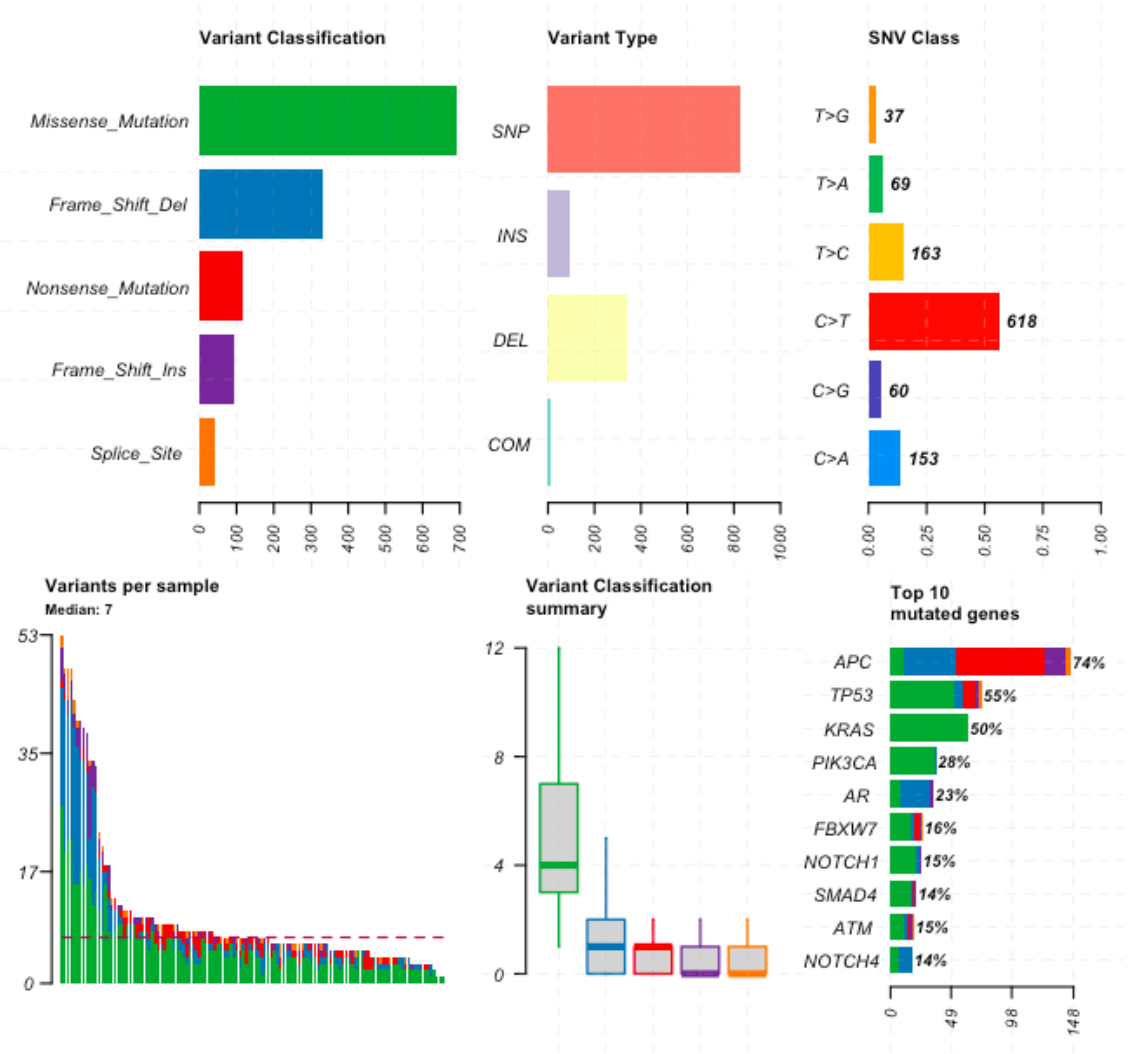


Figure 7.4 Overview of mutational landscape in GMS0 patients. *Summary plot showing the most frequent variant classifications, types, SNV classes, mutation burden and top 10 mutated genes in GMS0 patients from a subset of the Glasgow combined cohort (n=72) run in RStudio version 1.3 using maftools package.*

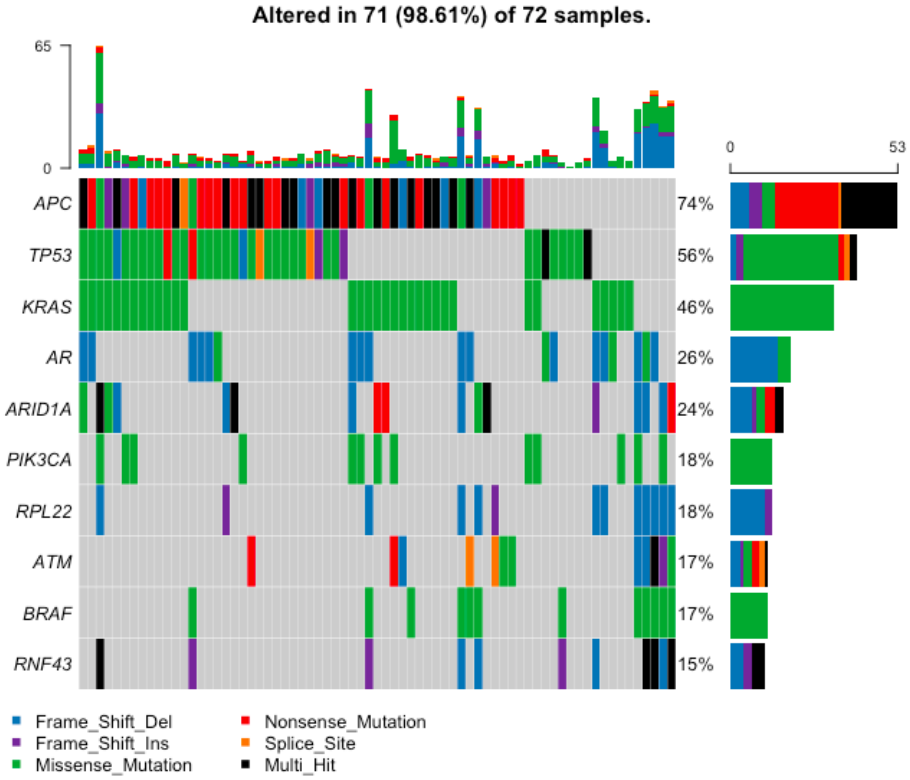


Figure 7.5 Mutational landscape of GMS0 patients. *Oncoplot showing the top 10 mutated genes in patients classified as GMS0 created in R Studio version 1.3 using the maftools package. In GMS0 patients APC was mutated in 74% of cases, TP53 in 56% and KRAS in 46% of patients. Sequencing was performed by Glasgow Precision Oncology Laboratory using a 151-gene custom panel.*

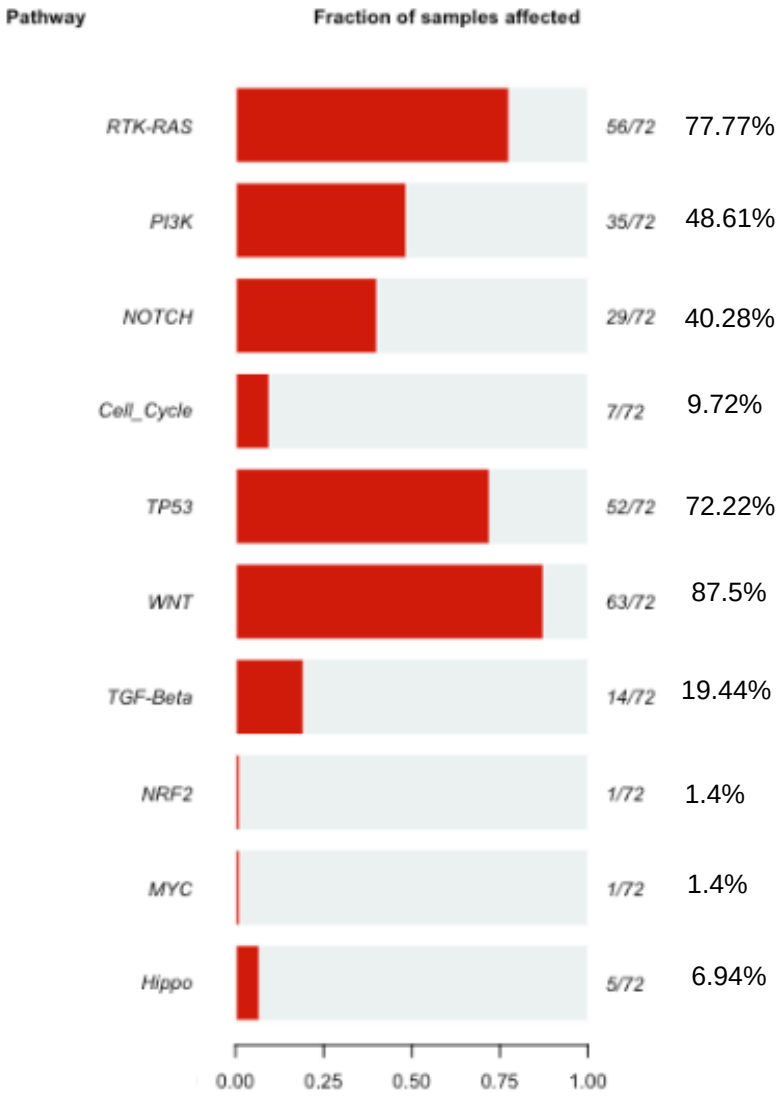


Figure 7.6 Pathway enrichment analysis in GMS0 patients. *Bar plot showing percentage of patients with a mutation in at least one gene from common cancer-associated signalling pathways created in R Studio version 1.3 using maftools package. Sequencing was performed by Glasgow Precision Oncology Laboratory using a 151-gene custom panel.*

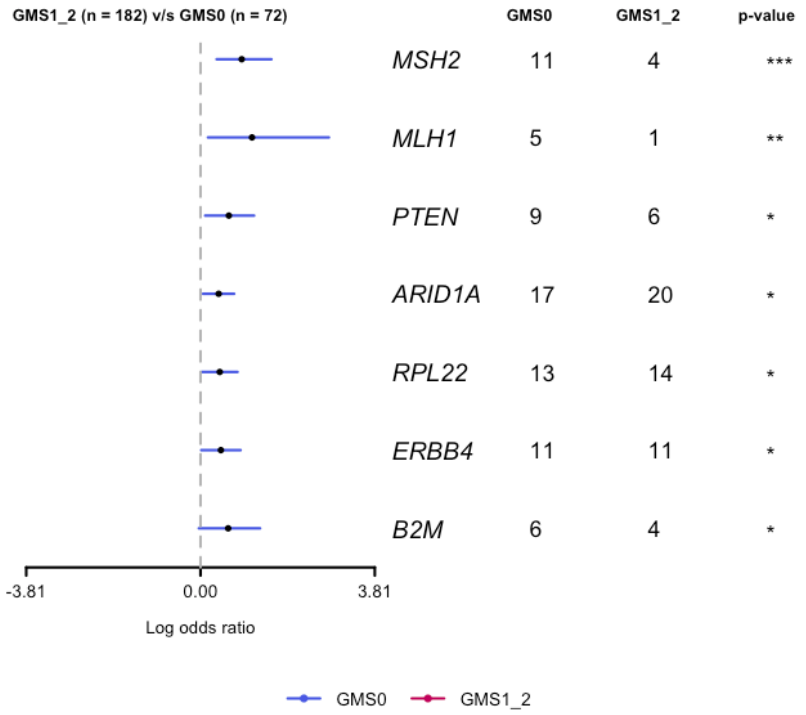


Figure 7.7 Frequently mutated genes in GMS0 versus GMS1/2. *Forest plot showing the significantly enriched mutation in GMS0 patients compared to GMS1 and GMS2 patients grouped together based on fishers exact tests and run in R Studio version 1.3 using the maftools package.*

When GMS1 patients were investigated, the median number of variants was also 7 and the most common variant classifications were also missense mutations (41.79%). Alterations tended to be single nucleotide polymorphisms (66.62%) and SNV classes were most frequently cytosine to thymine transitions (58.31%) (Figure 7.8). The top 10 mutated genes were the same for GMS 1 as GMS0 with APC mutated in 74% of patients, TP53 in 55% and KRAS in 50% of patients as shown in an oncoplot (Figure 7.9). PIK3CA was mutated in 28% of GMS1 patients compared to only 18% of GMS0 patients. Pathway enrichment analysis showed strong involvement of RTK-RAS (86.18%), WNT (86.18%) and TP53 signalling (67.48%), similar to that observed in GMS0 phenotypes (Figure 7.10). Fishers exact tests revealed significant associations between mutations in MSI gene MSH2 and GMS1 versus all other patients, with GMS1 patients exhibiting less mutations in MSH2 (p=0.00611) (Figure 7.11). In terms of genes more frequently mutated in the GMS1 group, death associated protein 6 DAXX (MAPK signalling, p=0.0319), ERBB2/HER2 (EGFR family, p=0.037) and MTOR (p=0.046) were observed at higher rates than in GMS0/2 patients (Figure 7.11).

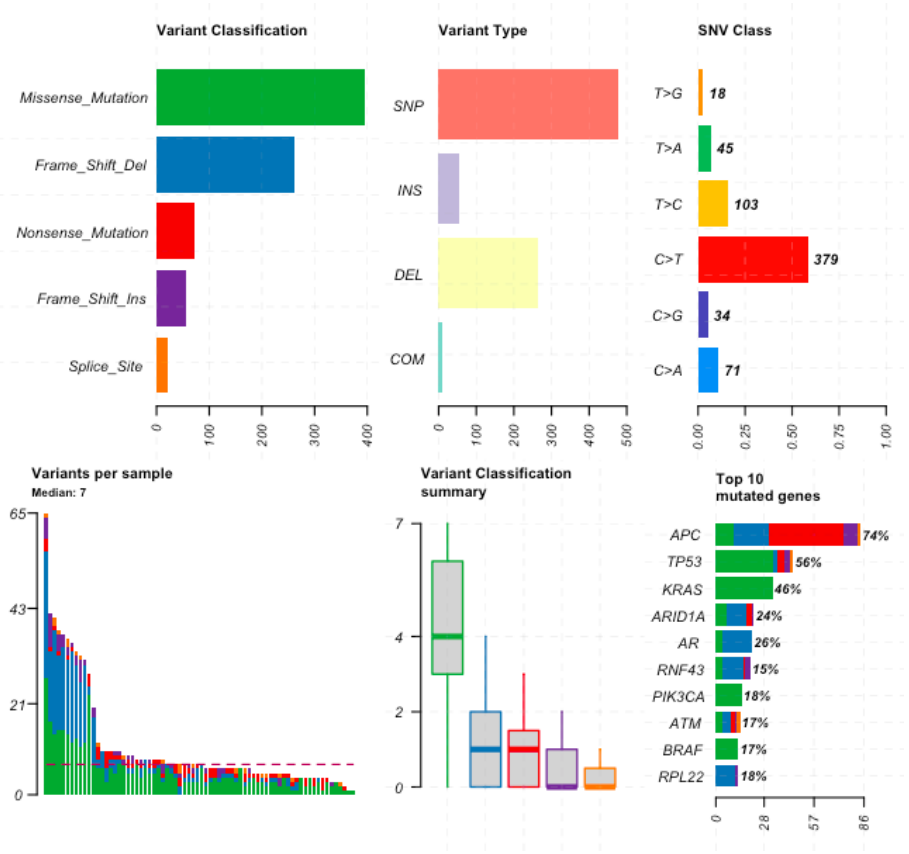


Figure 7.8 Overview of mutational landscape in GMS1 patients. *Summary plot showing distribution of variant classifications, types, SNV classes, mutation burden and top 10 mutated genes in a subset of the Glasgow combined cohort (n=123). Analysis was performed in R Studio version 1.3 using the maftools package.*

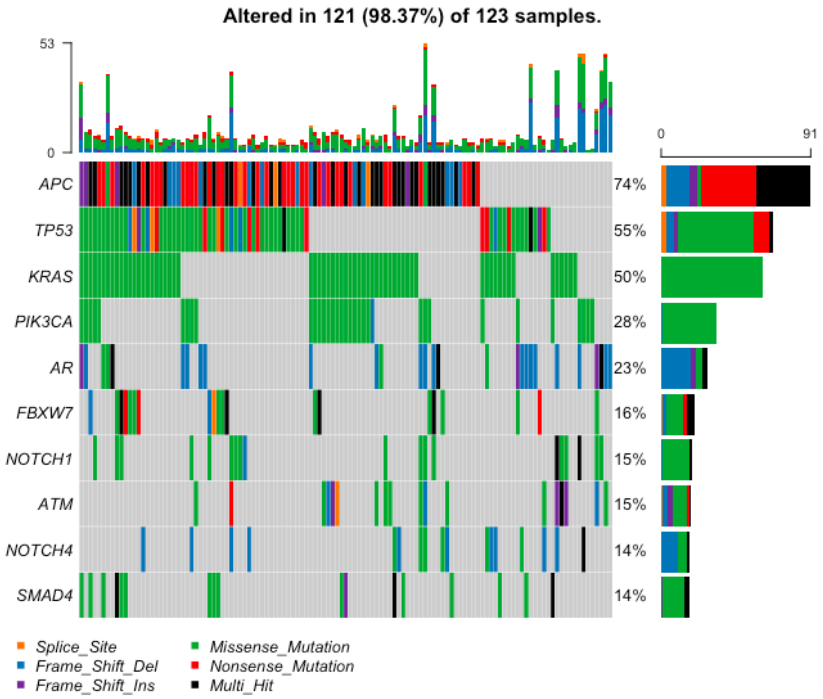


Figure 7.9 Mutational landscape in GMS1 patients. *Oncoplot showing the top 10 mutated genes amongst GMS1 patients from a subset of the Glasgow combined cohort (n=123). In GMS1 patients APC was mutated in 74% of cases, TP53 in 55% and KRAS in 50% of patients. Sequencing was performed by Glasgow Precision Oncology Laboratory and downstream analysis was performed in R Studio version 1.3 using the maftools package.*

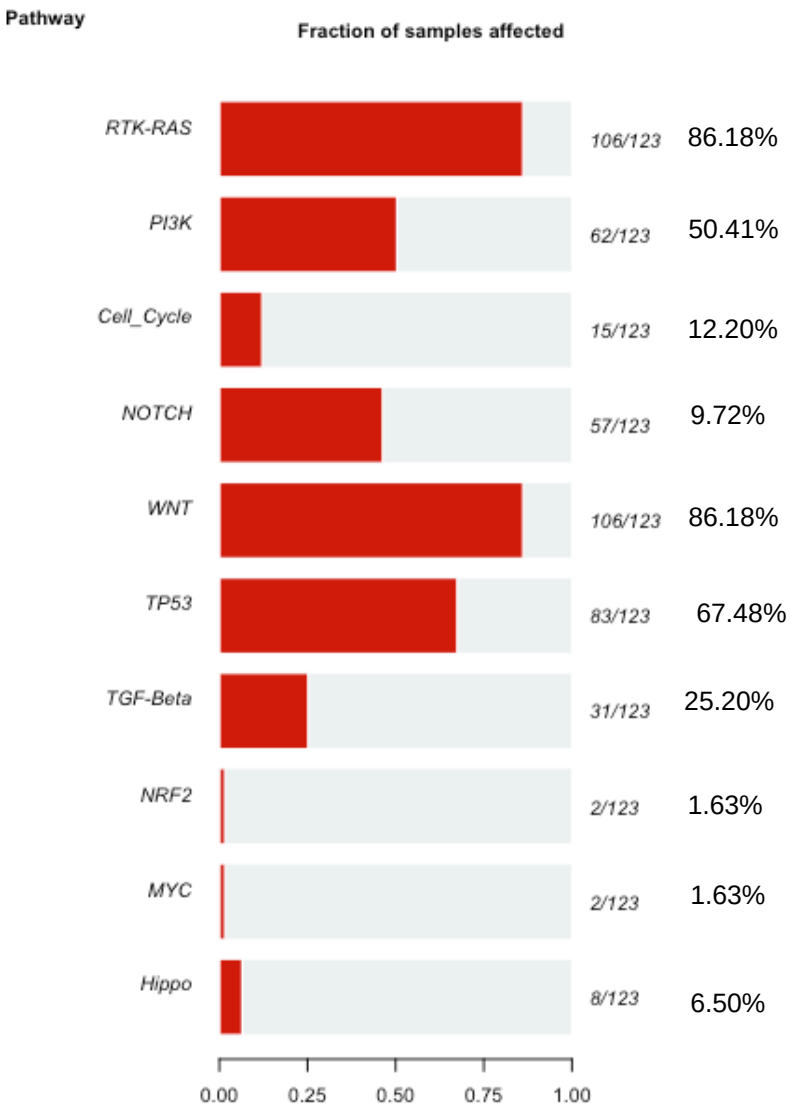


Figure 7.10 Pathway enrichment in GMS1 patients. *Bar plot showing the percentage of GMS1 patients in a subset of the Glasgow combined cohort showing an alteration/s in at least one gene from common cancer-associated pathways. Analysis was performed in R Studio version 1.3 using the maftools package.*

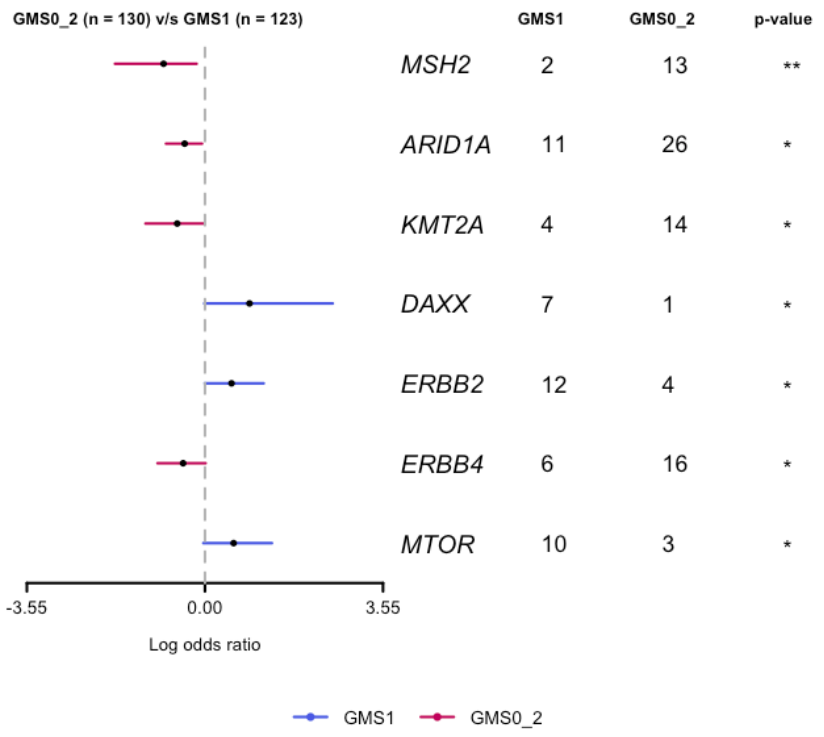


Figure 7.11 GMS1 versus GMS0/2 mutation patterns. *Forest plot showing differences in mutation landscape between GMS1 and GMS0/2 patients analysed via fishers exact tests in R Studio version 1.3 using the maftools package.*

In concordance with data from GMS0 and GMS1 groups, missense mutations were the most common variant classification observed (49.26%), SNPs the most common variant type (82.01%) and cytosine to thymine transitions the most common SNV class observed amongst GMS2 patients (62.85%) (Figure 7.12). The most frequently mutated gene observed in GMS2 patients was APC (78%). Interestingly, TP53 was mutated at a much higher frequency in GMS2 patients at 76% compared to 56% in GMS0 and 55% in GMS1 patients (Figure 7.13). KRAS alterations were observed in 47% of GMS2 patients (Figure 7.13). Pathway enrichment analysis showed high rates of mutations in RTK-RAS pathway (82.76% patients affected), WNT (91.4%) and TP53 (79.31%) similar to GMS0 and GMS1 (Figure 7.14). Fishers exact tests revealed significantly higher rates of TP53 (p=0.0059) and NTRK1 (nerve growth factor receptor, p=0.00175) genes in GMS2 patients (Figure 7.15). There were profoundly less TGFβ2R mutations seen in GMS2 patients than the other groups (p=0.0018).

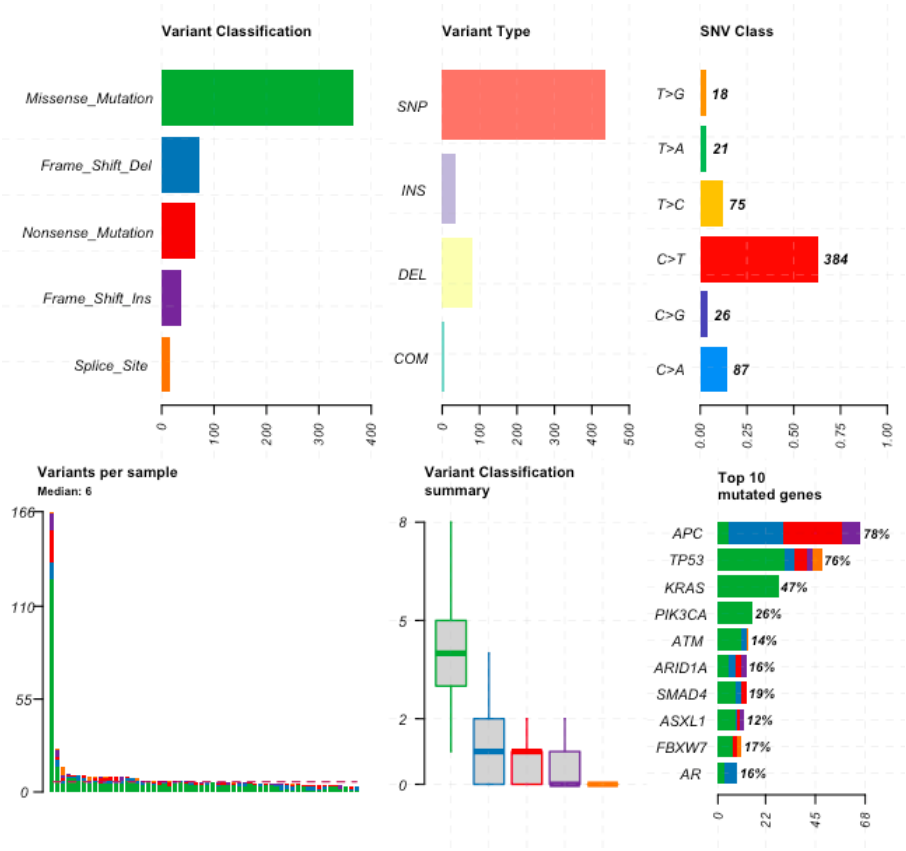


Figure 7.12 Summary plot for GMS2 patients. *Summary plot showing distribution of variant classifications, types, SNV classes, mutation burden and top 10 mutated genes in a subset of the Glasgow combined cohort (n=58). Panel sequencing of a custom 151 cancer-associated genes was performed by Glasgow Precision Oncology Laboratory and downstream analysis was performed in R Studio version 1.3 using the maftools package.*

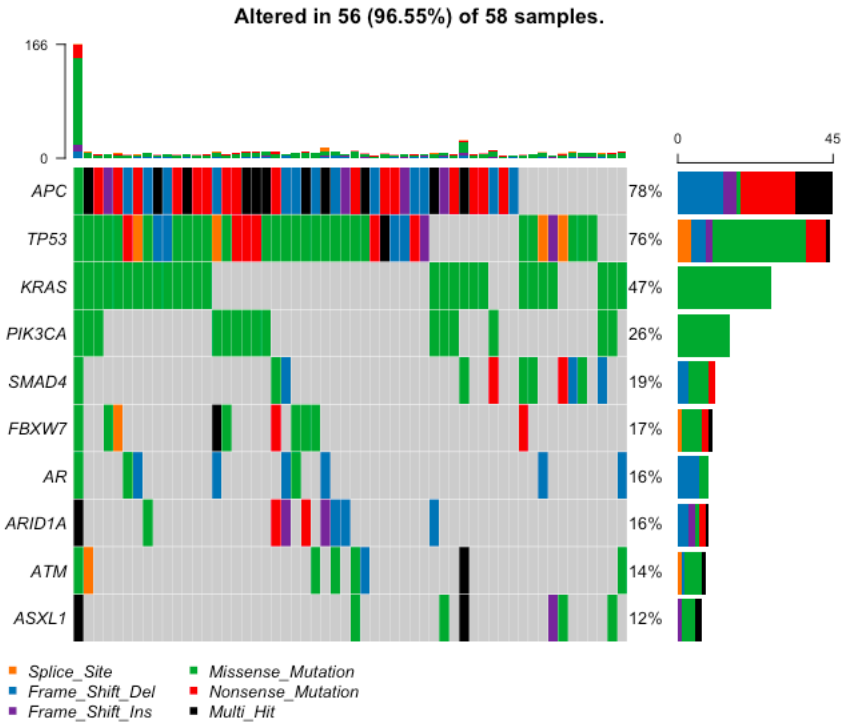


Figure 7.13 Mutational landscape in GMS2 patients. *Oncoplot showing the top 10 mutated genes amongst GMS1 patients from a subset of the Glasgow combined cohort (n=58). In GMS2 patients APC was mutated in 78% of cases, TP53 in 76% of cases and KRAS in 47% of cases. Analysis was performed in R Studio version 1.3 using the maftools package.*

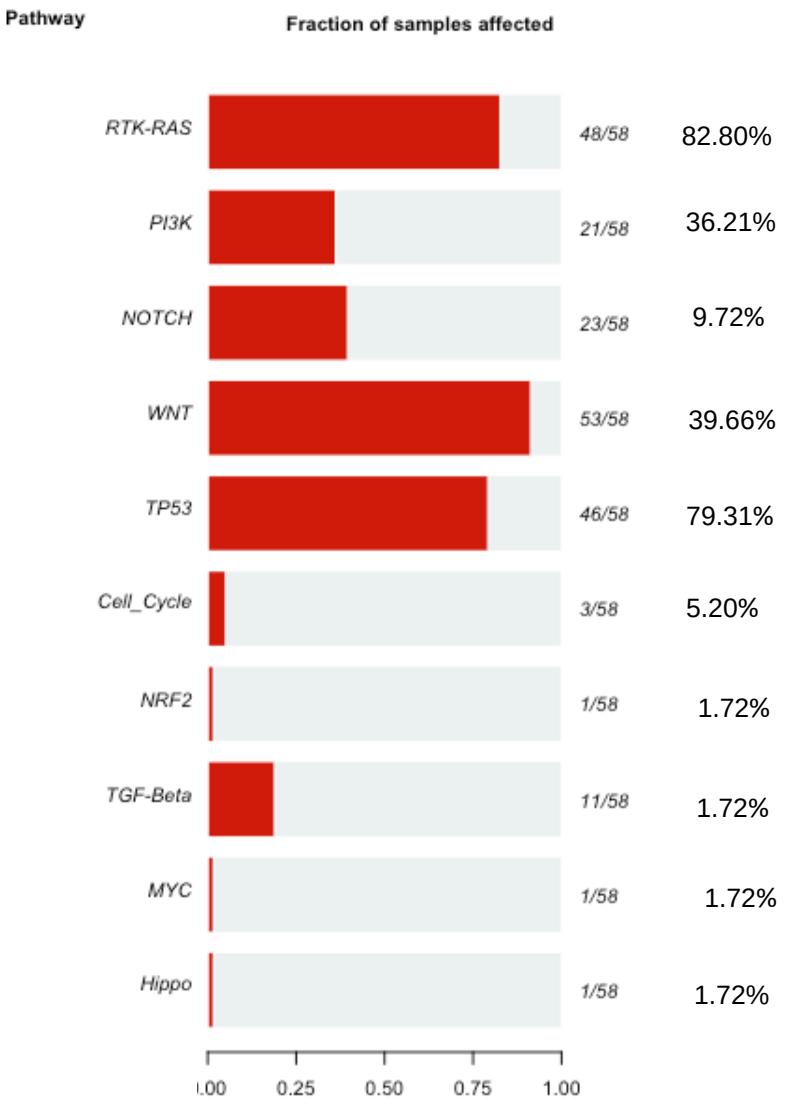


Figure 7.14 GMS2 versus GMS0/1 mutation patterns. *Forest plot showing differences in mutation landscape between GMS2 and GMS0/1 patients analysed via fishers exact tests performed in R Studio version 1.3 using the maftools package.*

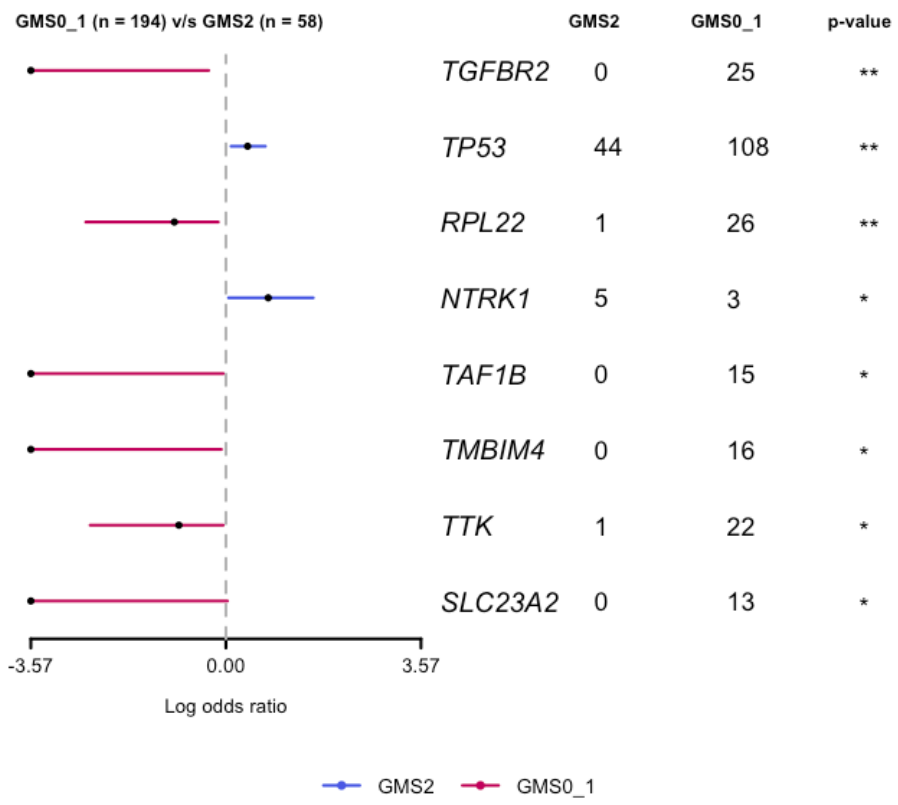


Figure 7.15 GMS2 versus GMS0/GMS1 patients. *Forest plot showing results from fisher exact test to identify difference sin frequency of alterations between GMS2 and all other patients. Analysis was performed in R Studio version 1.3 using the maftools package and statistical significance was set to $p<0.05$.*

Glasgow microenvironment scores were then investigated for association with differences in copy number alterations. Raw copy number data was transformed by calculating the difference between copy number state for the sample and copy number state for the gene. Negative values were considered deletions and positive values were considered gains. Patients were divided intro respective GMS groups and the mean score for each gene was calculated within each group. The difference in mean copy number for each gene was calculated between groups and the top 10 amplifications in both directions were plotted on swarm and box plots using Python package and seaborn function. When GMS0 patients were compared to GMS1 and 2 patients the top amplified genes in the GMS0 group were STK1, PIK3R3, CIITA, SMAD4, LINC00290, FGF3, FBXW7, PARP10, IDO1 and IDO2 (Figure 7.16). The top 10 downregulated genes in the GMS0 groups compared to GMS1 and 2 were ERBB2, SOX9, AR, CDK8, AMER1, NOTCH1, FANCC, RFXAP, BRCA2 and EGFR (Figure 7.16). Perhaps of most interest were IDO1 and IDO2, which are involved in metabolic regulation of

immune checkpoints and increased expression is associated with T cell anergy and tumour evasion (165). IDO1 and IDO2 genes showed strong amplification in the immune subgroup.

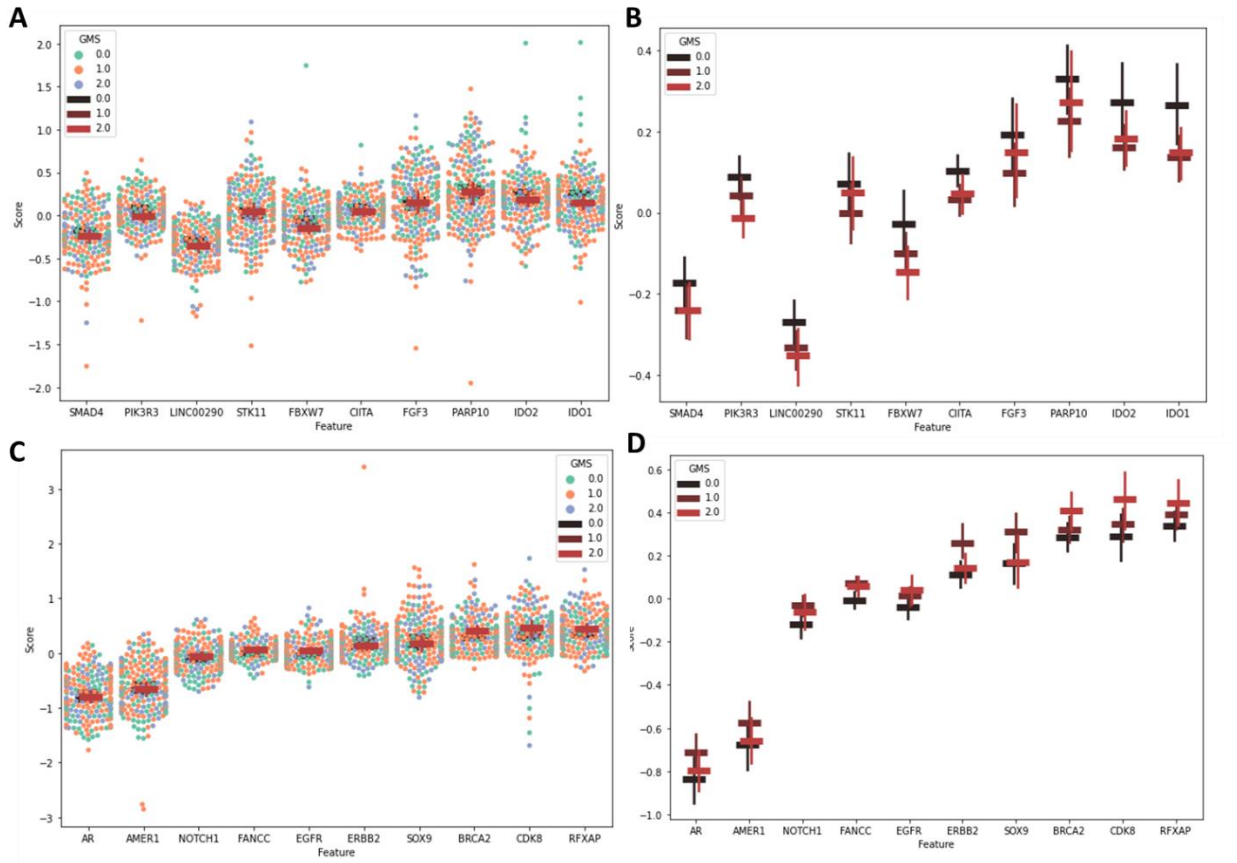


Figure 7.16 Copy number alterations in GMS0. *Swarm plot showing top 10 amplified copy number alterations for GMS0 versus GMS1/2 (A). Point plot showing top 10 amplified copy number alterations for GMS0 versus GMS1/2 (B). Swarm plot showing top 10 downregulated copy number alterations for GMS0 versus GMS1/2 (C). Point plot showing top 10 downregulated copy number alterations for GMS0 versus GMS1/2 (D).*

In terms of GMS1 versus the GMS0/2 the top 10 amplified genes were SOX9, ERBB2, AR, KDM6A, AMER1, MET, CDK4, GATA6, NOTCH1 and KDM5A (Figure 6.17). The top 10 downregulated genes in GMS1 were AURKA, IGF2, CCND3, HNF4A, PARP10, IDO1, FGF3, IDO2, PDCD1LG2 and ALOX15B (Figure 7.17). Interestingly, the most amplified genes, SOX9, is involved in activations of WNT signalling and AR signalling in castrate resistant prostate cancer. AR and AMER1 (WNT pathways member) are both also upregulated in GMS1 patients suggesting crosstalk between these pathways may be important in patients with low immune infiltrates and low stromal invasion (166).

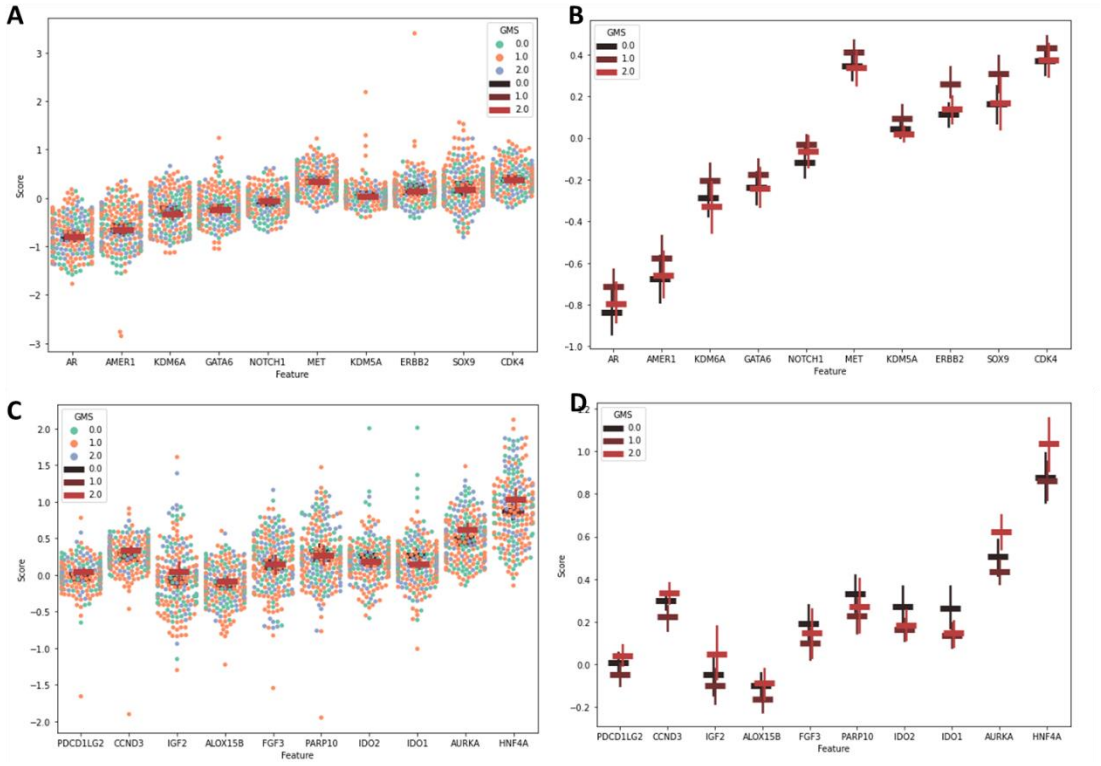


Figure 7.17 Copy number alterations in GMS1. *Swarm plot showing top 10 amplified copy number alterations for GMS1 versus GMS0/2 (A). Point plot showing top 10 amplified copy number alterations for GMS1 versus GMS0/2 (B). Swarm plot showing top 10 downregulated copy number alterations for GMS1 versus GMS0/2 (C). Point plot showing top 10 downregulated copy number alterations for GMS1 versus GMS0/2 (D).*

The top 10 amplified genes amongst GMS2 patients compared to GMS0/1 patients were HNF4A, AURKA, CDK8, IGF2, ZNF217, IRS2, BCL2L1, BRCA2, RB1, PIK3CA (Figure 7.18). The top 10 downregulated genes in GMS2 tumours compared to GMS0/1 were KDM6A, PPP2R2A, FAT1, SOX9, PIK3R3, FBXW7, KRAS, KIT, ERBB2, KDR (Figure 7.18). The most amplified gene in stromally dense tumours compared to high immune/low both tumours was HNF4A, a transcriptional regulator associated with tumour suppressor properties in gastrointestinal cancers (167).

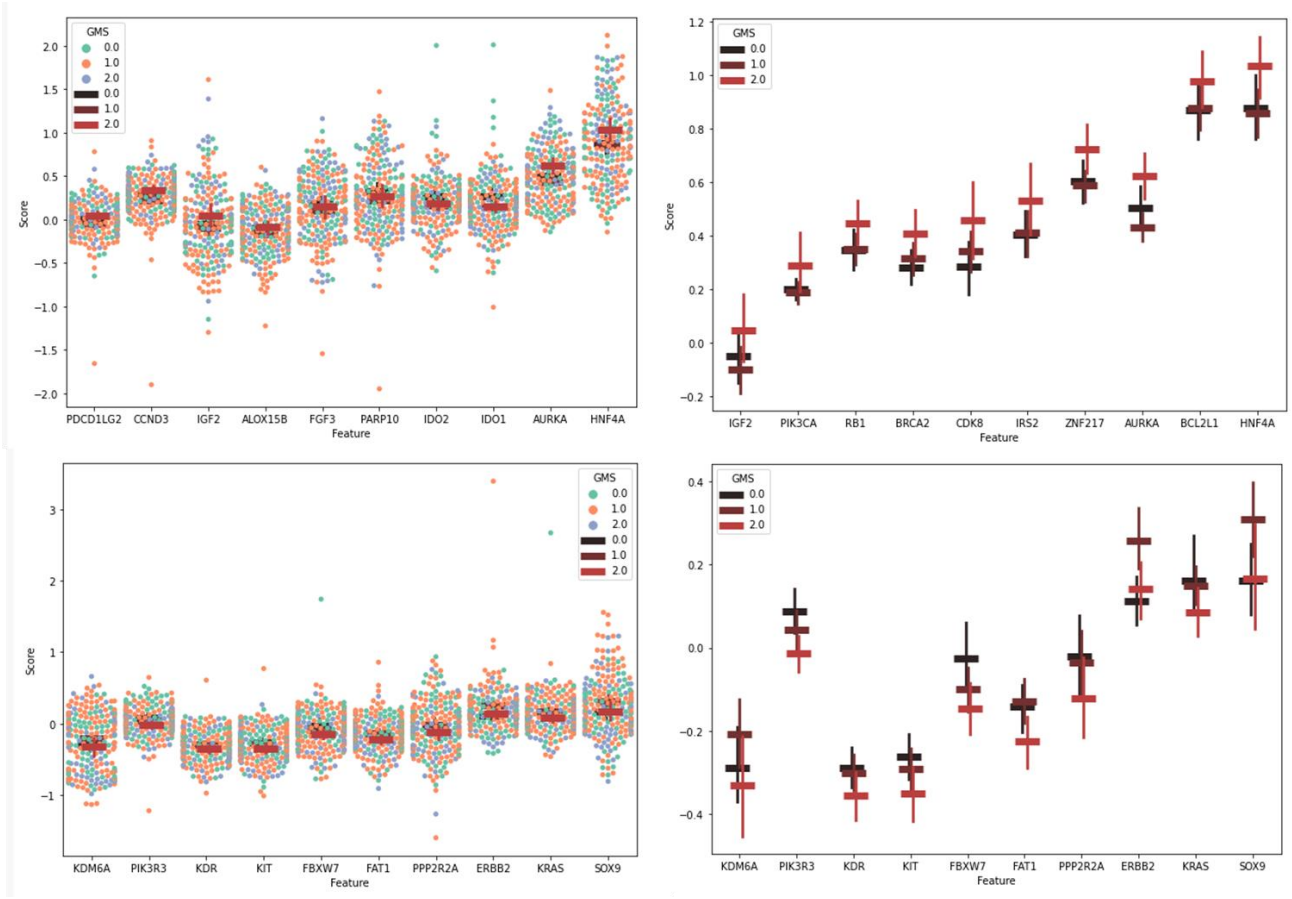


Figure 7.18 Copy number alterations in GMS1. *Swarm plot showing top 10 amplified copy number alterations for GMS2 versus GMS1/2 (A). Point plot showing top 10 amplified copy number alterations for GMS2 versus GMS1/2 (B). Swarm plot showing top 10 downregulated copy number alterations for GMS2 versus GMS1/2 (C). Point plot showing top 10 downregulated copy number alterations for GMS2 versus GMS1/2 (D).*

6.4 Mutational landscape of patients with activated STAT3

In addition to mutational landscaping for each GMS group, the underlying genomic biology was also investigated in relation to inflammatory pathway STAT3 signalling. Patients classified as high pSTAT3^{tyr705} protein expression via prior IHC analysis were selected from the 252 Glasgow combined cohort patients with mutational profiling. This identified 62 patients that had high expression of pSTAT3^{tyr705} protein for mutational analysis. 91% of tumours in this subgroup had mutations present, this is lower than the 97% observed for the cohort as a whole. The median number of variants observed in each patient tumour was 7. To visualise the mutational landscape of these high pSTAT3^{tyr705} patients a summary plot was constructed (Figure 7.19). APC mutations were mostly

nonsense mutations (41.74%) or frameshift deletions (35.69%). The majority of TP53 and KRAS mutations were missense mutations (61.02%, 92.7%, respectively) (Figure 7.19). An oncoplot was constructed using the 'oncoplot' function in maftools package in R Studio (Figure 7.20). The top 10 mutated genes in this subset of patients included APC (74% of patients harboured ≥1 mutation), TP53 (66%), KRAS (52%), AR (35%), PIK3CA (23%), ARID1A (21%), TGFβR2 (18%), NOTCH3 (18%), ASXL1 (18%), ALK (18%) (Figure 7.20).

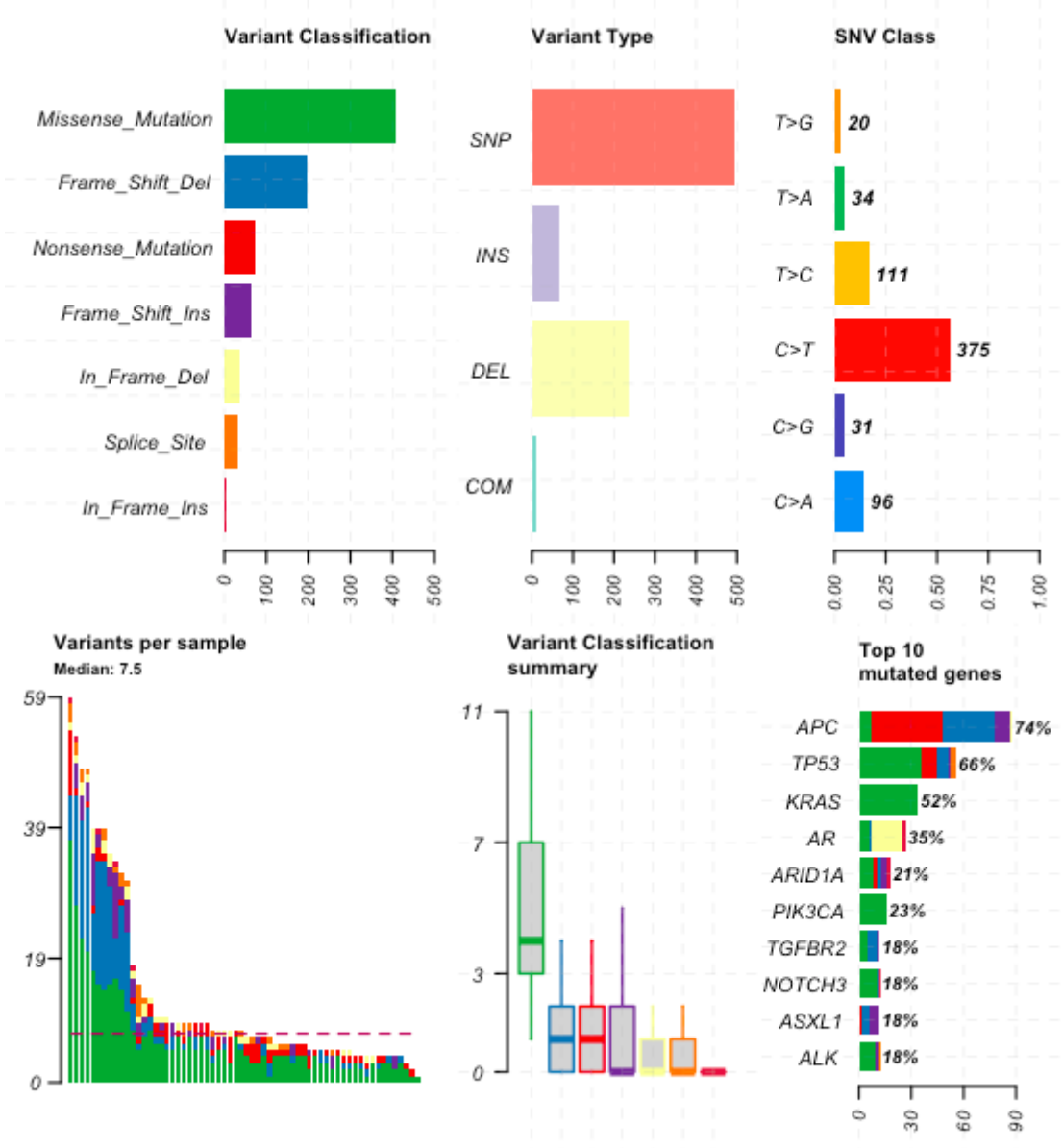


Figure 7.19 Mutation landscape in patients with activated STAT3. *Summary plot showing distribution of variant classifications, types, SNV classes, mutation burden and top 10 mutated genes in a subset of the Glasgow combined cohort with high STAT3^{tyr705} protein expression (n=62). Sequencing was performed by Glasgow Precision Oncology Laboratory focused on a panel of 151 cancer-associated genes. Analysis was performed in R Studio version 1.3 using the maftools package.*

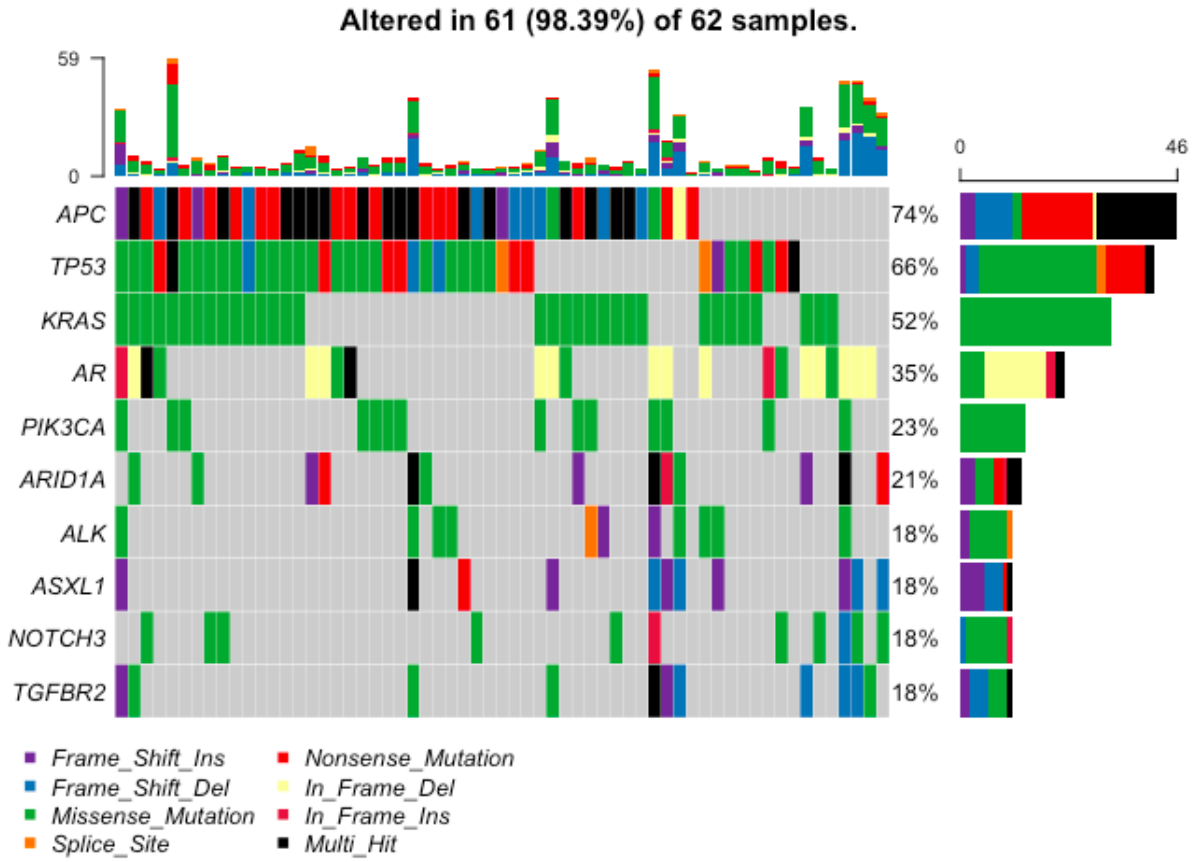


Figure 7.20 Mutational background of patients with activated STAT3 phenotypes. *Oncoplot showing the top 10 mutated genes in patients with high protein expression of pSTAT3^{tyr705}. In patients with high pSTAT3^{tyr705} protein expression APC was mutated in 74% of cases, TP53 in 66% and KRAS in 52% of cases. Analysis was performed in R Studio version 1.3 using the maftools package.*

Differences in the most frequently observed mutations between patients with high pSTAT3 ^{tyr705} and patients with low pSTAT3^{tyr705} protein expression were visualised in a co-oncoplot (constructed using the 'mafcompare' and 'coOncoplot' functions in maftools in RStudio) (Figure 7.21). The top 3 mutations for both groups were concordant, with APC mutated in 64% of high pSTAT3 ^{tyr705} patients versus 66% of low pSTAT3^{tyr705} patients, TP53 mutated in 54% high pSTAT3^{tyr705} group versus 50% of the low pSTAT3^{tyr705} group and KRAS mutated in 44% of high pSTAT3^{tyr705} group versus 39% of low pSTAT3^{tyr705} classified patients.

In contrast, the high pSTAT3^{tyr705} group of patients observed less ATM mutations (13% of patients) than the low classified pSTAT3^{tyr705} group (23% of patients) and less PIK3CA mutations (17% of patients) than the low classified pSTAT3^{tyr705} group (23% of patients).

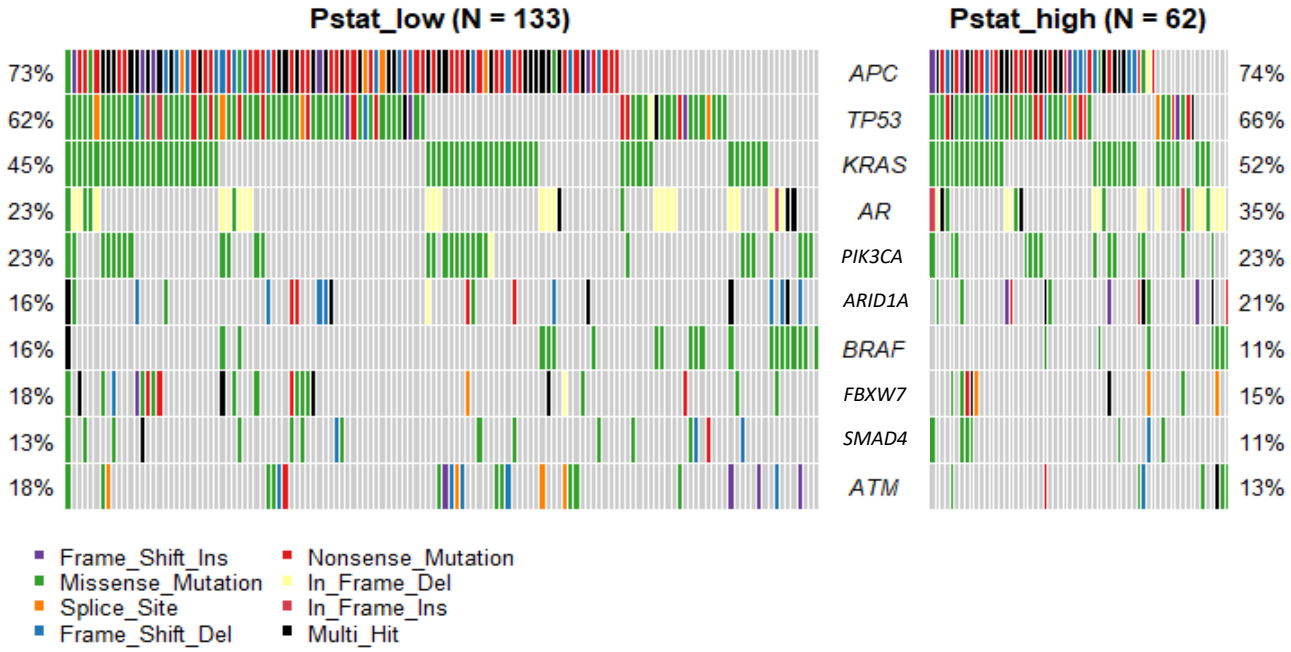


Figure 7.21 Mutational landscape in patients with activated STAT3 versus patients without STAT3 activation. *Oncoplot showing the top 6 mutated genes across high and low pSTAT3* ^{tyr705} *patient groups from a subset of the Glasgow combined cohort. Analysis was performed in R Studio version 1.3 using the maftools package. Mutation types included nonsense mutations, missense mutations, frameshift insertions, frameshift deletions, in-frame deletions, in-frame insertions, splice site and multi-hit mutations.*

A forest plot based on Fishers exact test was used to determine any significant differences in the presence of mutations in the panel (151 genes) between high and low pSTAT3^{tyr705} patient tumours (Figure 7.22). Genes mutated in less than 5 samples were excluded to avoid bias due to single mutated genes. Patients with high pSTAT3^{tyr705} expression exhibited a significantly higher frequency of AMER1 (p=0.010), RAF1 (p=0.030), MSH6 (p=0.033), PIK3CB (p=0.034), MAP2K4 (p=0.034) and SF3B1 (p=0.039) (Figure 7.22). AMER1 was mutated in 16.13% (10/62) of high pSTAT3^{tyr705} patients and 5.3% (6/133) of low pSTAT3^{tyr705} patients. RAF1 was mutated in 9.68% (6/62) of high pSTAT3^{tyr705} compared to 2.65% (3/113) low pSTAT3^{tyr705}. MSH6 was mutated in 16.13% (10/62) of high pSTAT3^{tyr705} compared to 7.1% (8/113) low pSTAT33^{tyr705}. PIK3CB and MAP2K4 genes were mutated in 8% (5/62) of high pSTAT3^{tyr705} compared to 1.8% (2/113) low pSTAT3^{tyr705} patients. SF3B1 was mutated in 11.3% (7/62) of high pSTAT3^{tyr705} compared to 3.54% (4/113) low pSTAT3^{tyr705}.

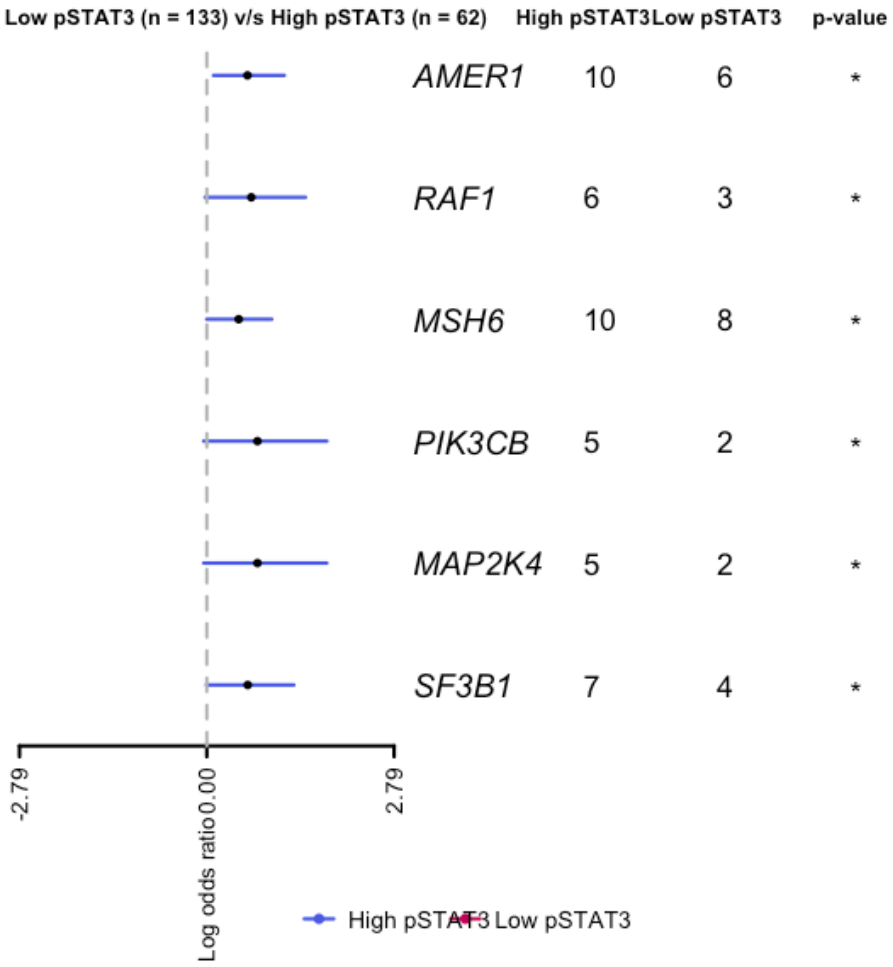


Figure 7.22 Comparison of mutational landscape of patients with high and low pSTAT3 ^{tyr705} protein expression. *Forest plot showing genes most differently mutated between high and low protein expression pSTAT3^{tyr705}groups. Analysis was performed in R Studio version 1.3 using the maftools package.*

To assess any patterns of co-occurrence or mutual exclusivity of mutated genes in patient with activated STAT3, pairwise Fishers' exact tests were performed. A somatic interactions plot was constructed using the 'somaticInteractions' function in maftools and RStudio visualise the results (Figure 7.23). Presence of TP53 and BRCA2 or BRAF mutations were significantly likely to not co-occur ($p<0.01$). TP53 mutations rarely occurred in the presence of RNF43 ($p<0.01$), TGFβR2 ($p<0.05$) and RPL22 mutations ($p<0.05$) in patients with high pSTAT3 ^{tyr705}. AR mutations tended to co-occur with TGFβR2 mutations ($p<0.01$).

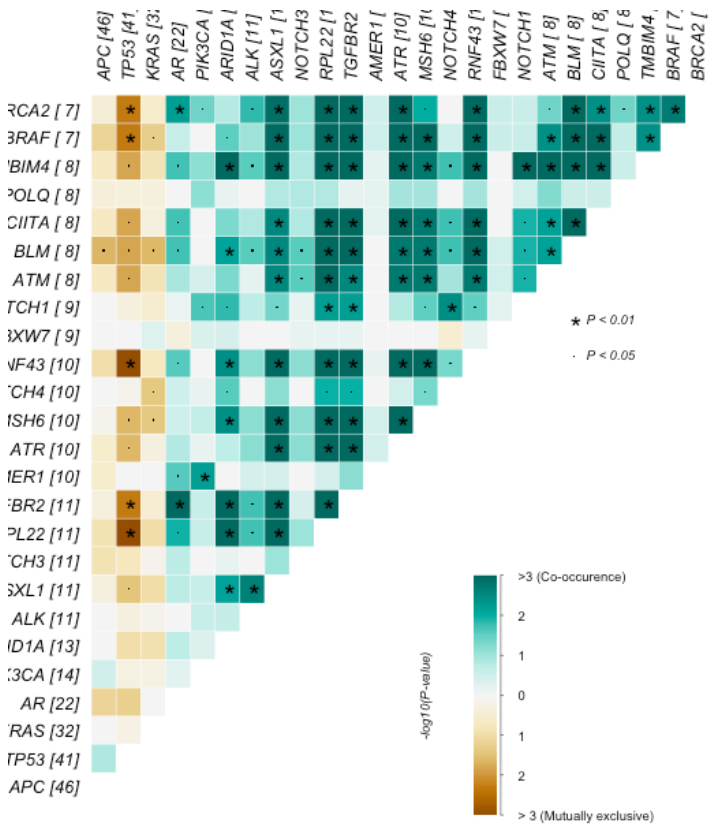


Figure 7.23 Co-occurrence of genetic mutations in high pSTAT3 ^{tyr705} patients. *Somatic interactions plot showing the co-occurrence or exclusivity of mutations in the activated STAT3 in a subset of the patients from the Glasgow combined cohort. Green colour indicates a high level of co-occurrence and orange colour represents a high level of exclusivity. Significance was set to p<0.01 represented by '*', and p<0.05 represented by '.'.*

Next, copy number alterations were analysed with respect to high and low pSTAT3^{tyr705} groups. Raw copy number data was transformed by calculating the difference between copy number state for the sample and copy number state for the gene. Negative values were considered deletions and positive values were considered gains. To determine any differences between groups, swarm plots were constructed by Dr Akhill Yeduresi using Python. The top ten copy number alterations with the biggest difference in mean between pSTAT3^{tyr705} groups are shown in Figure 6.24. Patients with activated pSTAT3^{tyr705} showed amplification of WNT signalling pathway member RNF43 compared low pSTAT3 ^{tyr705} expressing patients (Figure 7.24). Conversely, low pSTAT3 ^{tyr705} patients showed amplification of CDK8 compared to high pSTAT3^{tyr705} patients (Figure 7.24). No statistical testing was performed; however, these data show clear differences in copy number alterations between high/low pSTAT3^{tyr705} phenotypes.

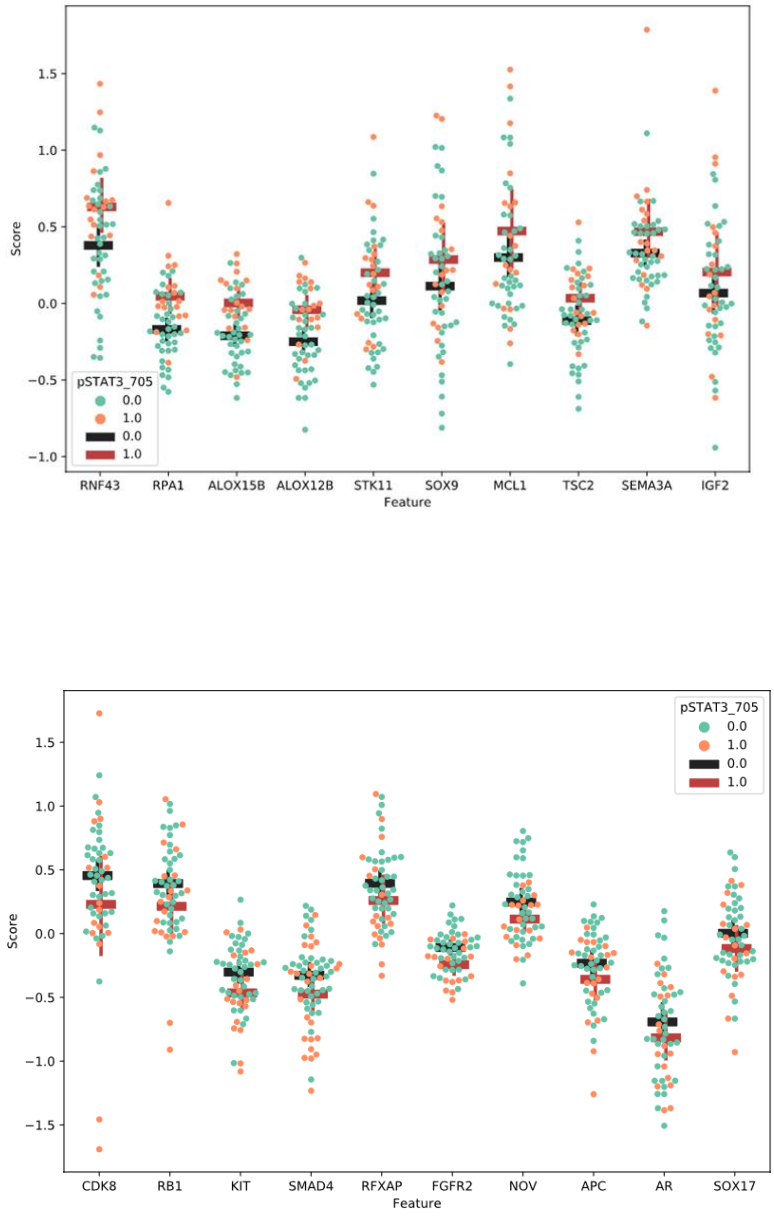


Figure 7.24 Copy number alterations in high versus low pSTAT3^{tyr705} groups. *Top 10 genes with highest difference in copy number alteration for high (A) and low (B) pSTAT3^{tyr705} groups. Black bar represents mean of low pSTAT3^{tyr705} group and red bar represents mean of high pSTAT3^{tyr705} group.*

6.5 Oncogenic signalling pathways

To investigate oncogenic signalling pathways in specimens with high pSTAT3^{tyr705} protein expression, pathway fraction alterations were assessed using 'OncogenicPathways' function in maftools with R Studio (Figure 7.25). TP53 signalling observed the highest fraction of pathway

affected by a mutation with one of 4 (out of 6) genes mutated in at least one patient. TP53 mutations were detected in a total of 96 patients. Mutations in at least one member of the WNT signalling pathway and RTK-RAS were observed in 111 patients (Figure 7.25).

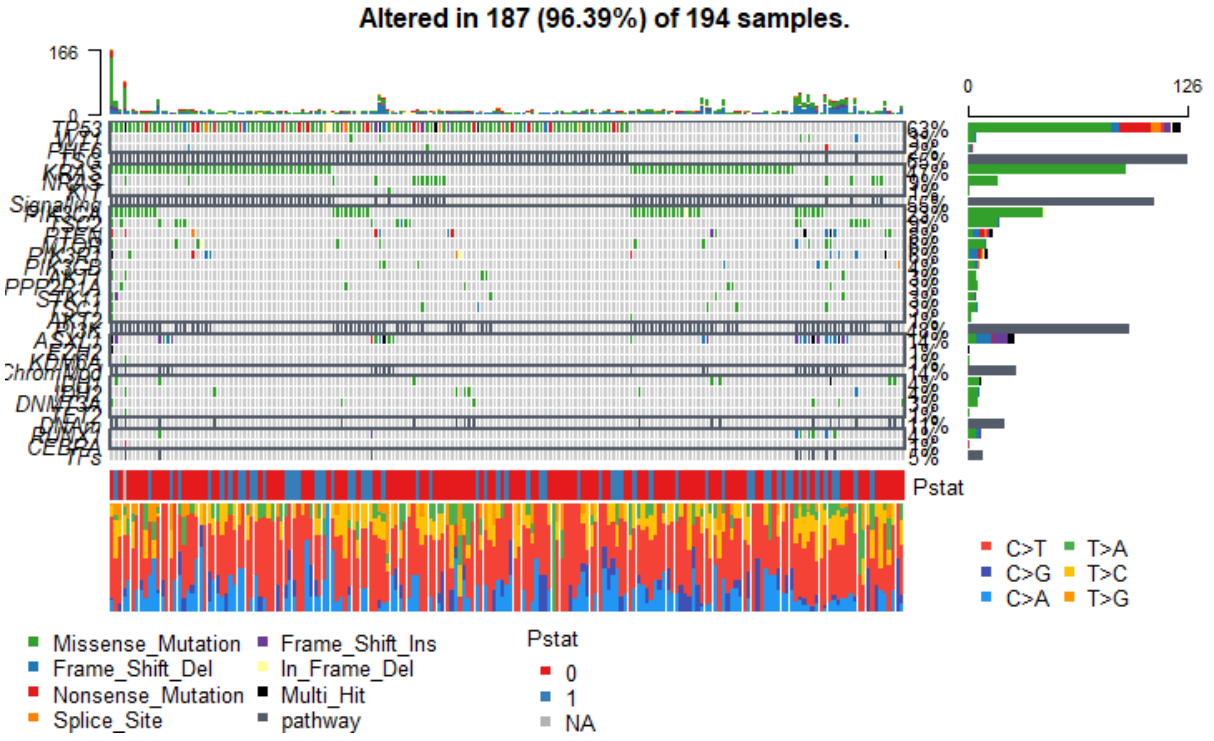


Figure 7.25 Alterations in signalling pathways with respect to STAT3 activation status. *Oncoplot showing presence of mutations in cancer-associated signalling pathways linked to STAT3^{tyr705} protein status. Analysis was performed in R Studio version 1.3 using the maftools package and created by Dr Akhill Yedhursi.*

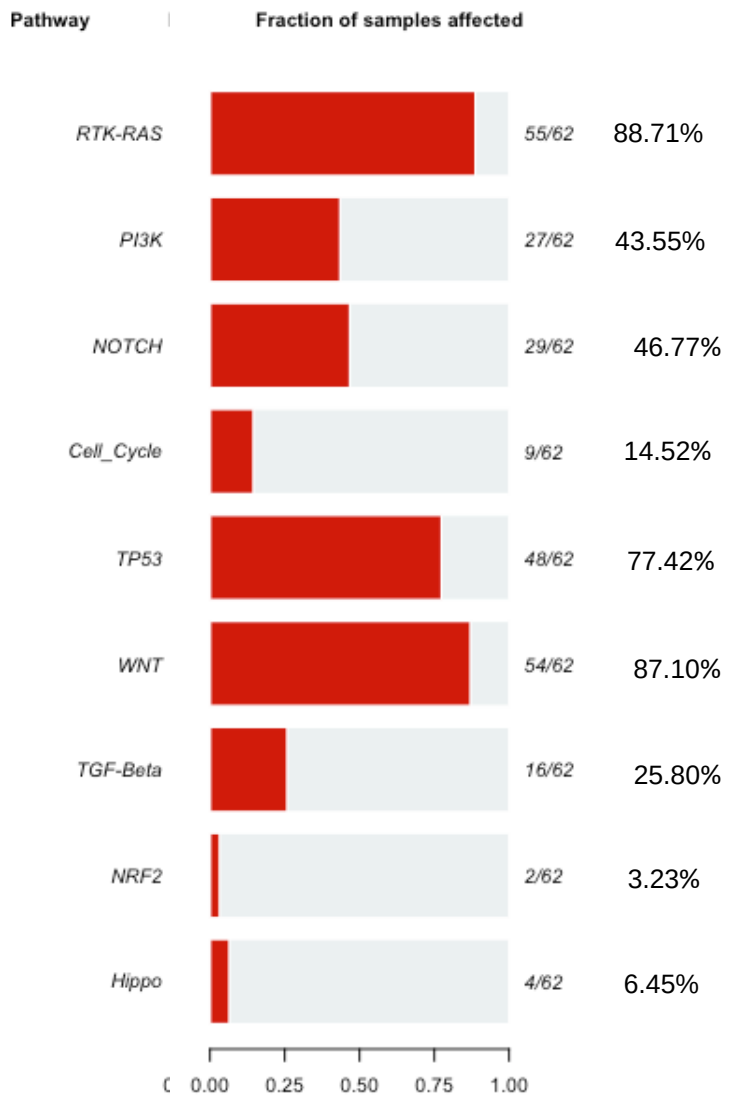


Figure 7.26 Pathway alteration fractions in high pSTAT3^{tyr705} patients. *Pathway enrichment analysed via OncogneicPathways in patients from the Glasgow combined cohort with high pSTAT3^{tyr705}protein expression and analysis was performed in R Studio version 1.3 using the maftools package.*

Using 'PlotOncogenicPathways' function, it was possible to visualise presence of mutations in specific genetic components of each pathway of interest (Figure 7.26). 87.1% of patients had an alteration in at least 1 gene in the WNT pathway, 88.71% were affected by the RTK-RAS pathway and 77.42% of patients displayed a mutation in TP53 signalling (Figure 7.26).

7.6 Amino acid changes

7.6.1 Mutations in members of the IL6-STAT3 pathway

In patients with high pSTAT3 ^{tyr705} protein expression, mutations in IL6-STAT3 signalling pathway members JAK1, JAK2 and JAK3 were observed. Oncogenes tend to have preferential sites where mutations occur, referred to as 'hot-spots'. To investigate the frequently mutated loci within JAK genes in this cohort, lollipop plots were constructed using the 'lolliPlot' function in maftools within R Studio. JAK1 was mutated in 1.99% of patients with activated STAT3, and mutations occurred at two hotspots (Figure 7.27). JAK2 was also mutated in 1.99% of patients, with mutations clustered in the SH2 domain (Figure 7.28). There were 4 hotspots identified within the JAK3 gene, which had an overall mutation rate of 3.31% in patients with high pSTAT3^{705} expression (Figure 7.29).

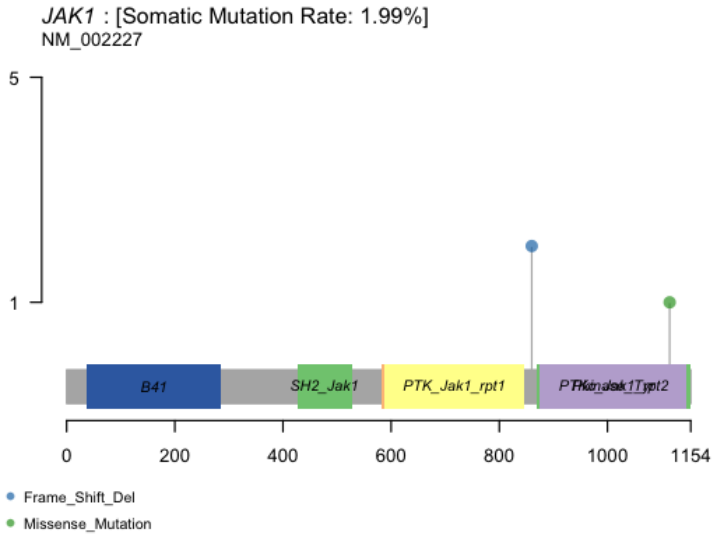


Figure 7.27 Preferential mutational loci within JAK1 gene in patients with activated STAT3. *Lollipop plot showing sites of mutation within the JAK1 gene in CRC patients categorised as high pSTAT3 ^{tyr705} and mutation rate within the JAK1 gene for the cohort (1.99%). Analysis was performed in R Studio version 1.3 using the maftools package.*

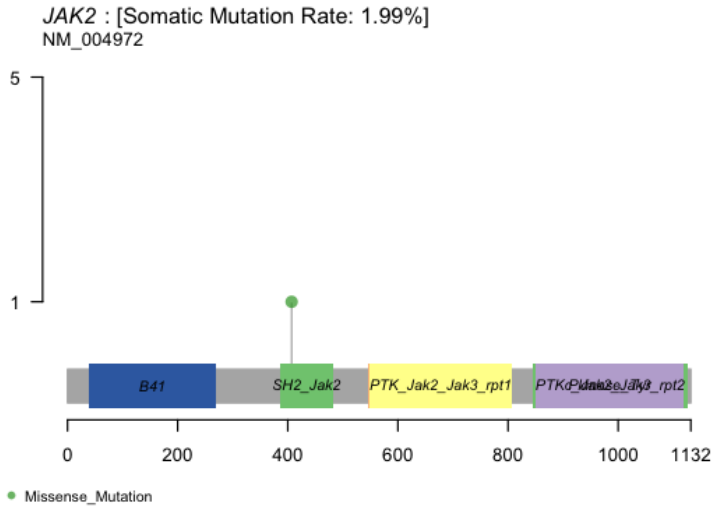


Figure 7.28 Preferential mutational loci within JAK2 gene in patients with activated STAT3. *Lollipop plot showing sites of mutation within the JAK2 gene in CRC patients categorised as high pSTAT3* ^{tyr705} *and the mutation rate for the JAK2 gene in the cohort (1.99%). Analysis was performed in R Studio version 1.3 using the maftools package.*

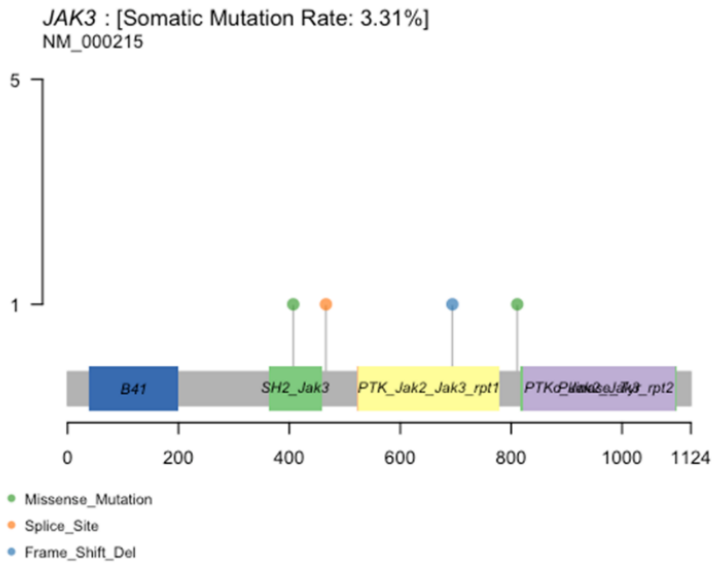


Figure 7.29 Preferential mutational loci within JAK3 gene in patients with activated STAT3. *Lollipop plot showing sites of mutation within the JAK3 gene in CRC patients categorised as high pSTAT3* ^{tyr705} *and the mutation rate for the JAK3 gene in the cohort (3.31%). Analysis was performed in R Studio version 1.3 using the maftools package.*

7.7 Mutation status, clinicopathological features and outcome

Presence of a mutation in JAK1, JAK2 and/or JAK3 was detected in 29 (11.6%) of all 255 Glasgow combined cohort patients with mutational data available. To determine if mutation(s) in any JAK gene was associated with CSS a Kaplan-Meier survival curve was plotted (Figure 7.30). No significant association was observed (HR= 1.374. 95%CI; 0.797-2.371, p=0.248). Chi-squared analysis was used to look for associations between mutation(s) in JAK genes and clinicopathological features (Table 7.1). Mutation/s in ≥ 1 JAK gene was significantly associated with T stage (p=0.014), MMR status (p<0.001), tumour subsite (p=0.001), peritoneal involvement (p=0.002) and marginal involvement (p=0.032) (Table 7.1). Patients with a JAK mutation were likely to be T stage IV (62.5%), right sided (79.2%) and MMR deficient (54.2%) (Table 7.1).

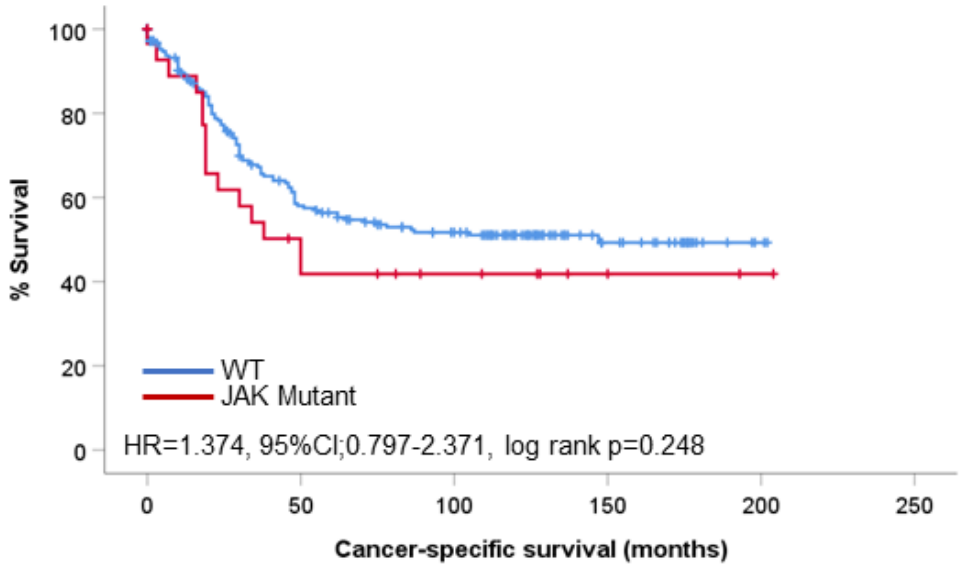


Figure 7.30 JAK mutations and association with clinical outcome. *Kaplan-Meier survival curve showing the association between presence of mutation in JAK1, JAK2 and/or JAK3 and cancer-specific survival in the Glasgow combined cohort. In the full cohort, presence of mutation in a JAK gene yielded a hazard ratio of 1.374, (95%CI; 0.797-2.371), log rank p=0.248.*

Clinicopathological feature	JAK status (JAK1/JAK2/JAK3)		p
	Wild type	Mutated	
T stage			
I	7 (3.7)	0 (0)	0.014
II	15 (7.9)	2 (8.3)	
III	111 (58.1)	7 (29.2)	
IV	58 (30.4)	15 (62.5)	
N Stage			
0	98 (51.6)	10 (41.7)	0.638
1	62 (32.6)	9 (37.5)	
2	30 (15.8)	5 (20.8)	
MMR Status			
pMMR	161 (84.7)	11 (45.8)	<0.001
dMMR	29 (15.3)	13 (54.2)	
Tumour subsite			
Right	77 (40.3)	19 (79.2)	0.001
Left	72 (37.7)	4 (16.7)	
Rectal	42 (22.0)	1 (4.2)	
Peritoneal involvement			
Absent	133 (69.6)	9 (37.5)	0.002
Present	58 (30.4)	15 (62.5)	
Marginal Involvement			
Absent	183 (95.8)	20 (83.3)	0.032
Present	8 (4.2)	4 (16.7)	
mGPS			
0	83 (57.6)	7 (38.9)	0.264
1	34 (23.6)	5 (27.8)	
2	27 (18.8)	6 (33.3)	
GMS			
0	58 (31.2)	7 (29.2)	0.979
1	91 (48.9)	12 (50.0)	
2	37 (19.9)	5 (20.8)	
Phenotypic subtype			
Immune	58 (31.4)	7 (29.2)	0.351
Canonical	52 (28.1)	10 (41.7)	
Latent	38 (20.5)	2 (8.3)	
Stromal	37 (20.0)0	5 (20.8)	

Table 7.1 JAK mutations and clinicopathological features. *Chi-squared table of associations for ≥ 1 mutation in JAK1, JAK2 and/or JAK3 genes and clinical features including T stage, N Stage, MMR status, tumour subsite, peritoneal involvement, marginal involvement, mGPS, GMS and phenotypic subtype.*

7.8 TCGA analysis

Given the association between JAK mutations and factors associated with poor prognosis and trend towards predicting poor outcomes in the Glasgow CRC combined cohort the Cancer Genome Atlas (TCGA) PanCancer Atlas cohort was analysed to validate these observations. Protein data for $pSTAT5^{tyr705}$ was not available for this cohort so JAK-STAT3 signalling was investigated only at the mutation and mRNA level. A total of 594 patients were included in analysis which was completed using cBioPortal for Cancer Genomics software. In concordance with mutational patterns observed in the Glasgow combined cohort, frequently mutated genes in the TCGA PanCancer Atlas cohort included APC (67%), TP53 (55%) and KRAS (37%) (Figure 7.31). Mutations in the AR gene were less common in the TCGA cohort than the Glasgow with only 9% of patients altered (Figure 7.31). In terms of Jak mutations, JAK1 was mutated in 4% of patients and JAK2 and JAK3 were mutated in 3% of patients.

6.8.1 JAK mutations, outcomes, and clinical factors

The main purpose of analysing mutation data in separate CRC cohort was to investigate JAK mutations and association with survival and clinical factors to validate that seen in the Glasgow cohort. In contrast to the Glasgow cohort, no association was observed between mutations in JAK1/2/3 and cancer-specific survival (p=0.526, figure 6.32). However, presence of ≥1 JAK mutation was associated with clinical measures including stage, with more stage IIA presentations in the mutated group assessed via Chi-Squared analyses $p=0.00110^{\wedge 10}$) (figure 7.33) Kruskal wallis tests revealed significant associations between JAK mutations and mutation count, with mutated patients showing a higher mutation burden (p<10^10, figure 6.34). Kruskal wallis tests also revealed a significant association between presence of JAK mutation and MSI Mantis score ($p=9.83e^{-9}$, figure 7.34). MSI mantis score is a clinical measure used to predict the likelihood of a patient being microsatellite instable, the higher the mantis score the more likely a patient is to be instable.

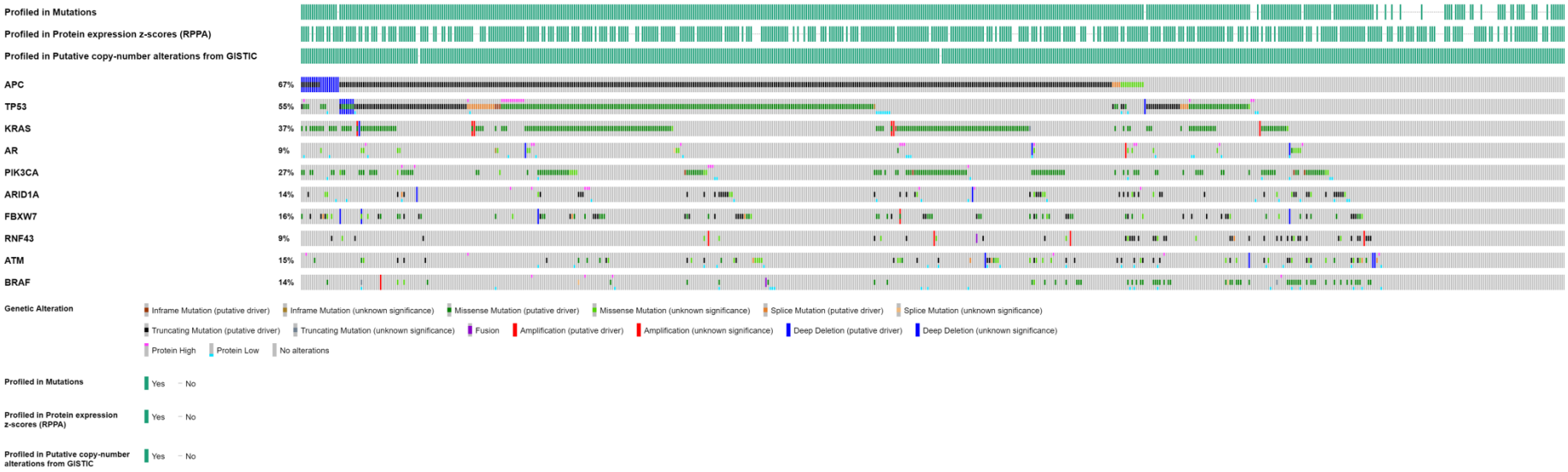


Figure 7.31 Glasgow cohort mutation landscape in TCGA dataset. *Oncoprint showing the mutation rates of the top 10 mutated genes observed in the Glasgow combined cohort, within the TCGA PanCancer Atlas CRC cohort. In this validation cohort the frequency of APC mutation was 67%, TP53 mutation was 55% and KRAs was mutated in 37% of cases.*

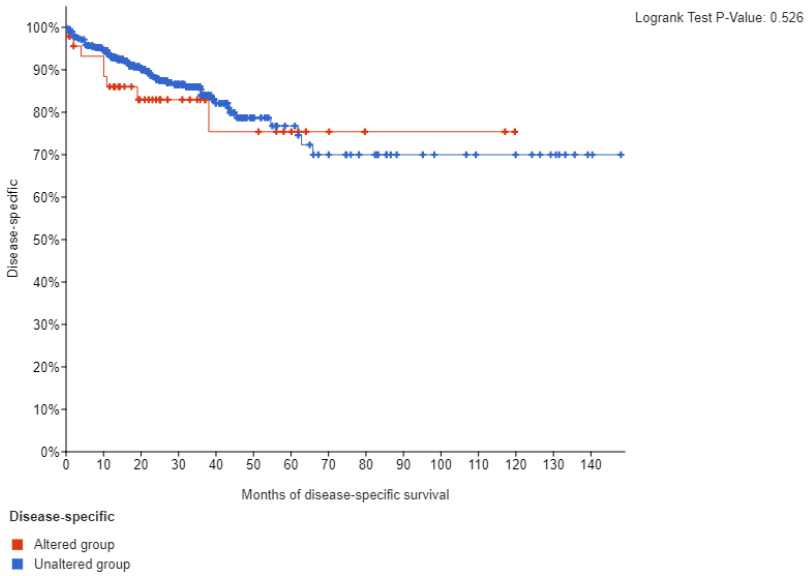


Figure 7.32 JAK mutations and Cancer-Specific Survival. *Kaplan Meier curve showing no association between presence of ≥1 JAK mutation and cancer-specific survival in the TCGA PanCancer Atlas CRC cohort. The red line represents patients with a mutation in at least 1 JAK gene and the blue line represents patients with no mutations detected in any JAK gene. Analysis was performed in cBioPortal.*

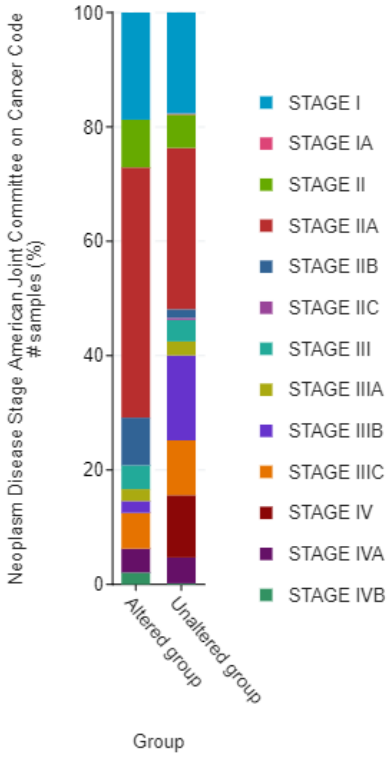


Figure 7.33 Jak mutations and CRC Stage. *Bar plot showing distribution of CRC disease stage (Stage I-IVB) at presentation in JAK-mutated (Altered) and JAK-wildtype (Unaltered) groups. Analysis was performed in cBioPortal.*

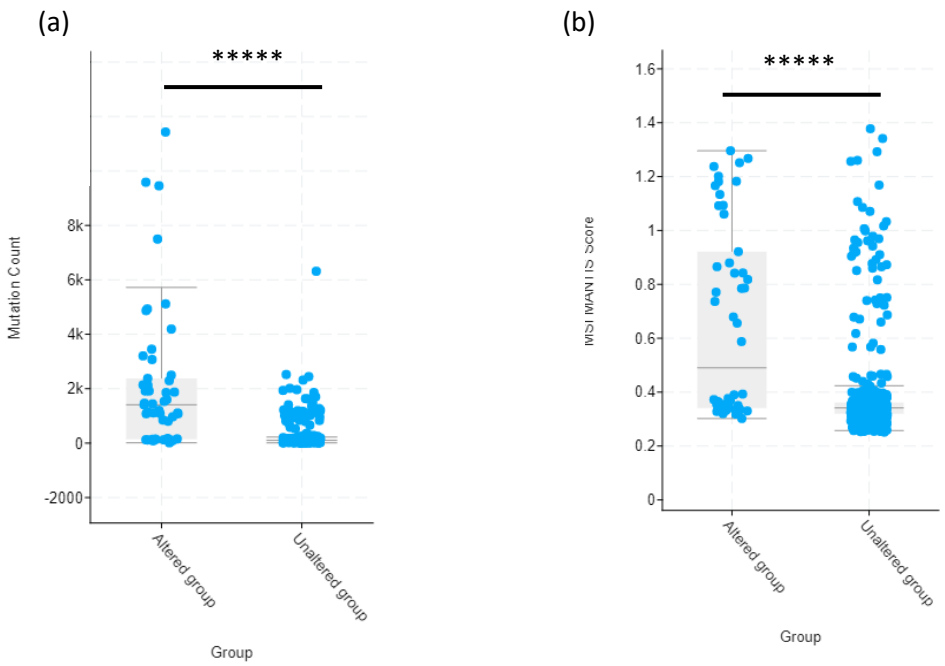


Figure 7.34 Jak mutation status and overall mutation count (a) and MSI mantis score (b). *Box plots showing association between presence of ≥1 JAK mutations and mutation burden (a) and likelihood of microsatellite instability (b). For both plots altered group represents the patients in the cohort that had a mutation detected in at least 1 of the JAK genes (JAK1, JAK2, JAK3 or TYK2) and patients in the unaltered group did not present with mutation in any JAK gene. Analysis was performed in cBioPortal.*

The TCGA PanCancer Atlas CRC cohort benefits from associated protein data for numerous pre-clinical biomarkers measured via reverse phase protein arrays (RPPAs). To investigate differences in protein expression between JAK mutated and JAK wildtype patients a volcano plot was constructed (Figure 7.35). The top two differentially expressed protein between groups were middle phase apoptotic marker Caspase7 (CASP7) and Ras signalling pathway member RAF1. CASP7 was significantly enriched in the JAK altered group (p=0.00003414, Figure 7.36). RAF1 expression was also significantly increased in the JAK mutant group (p=0.0167, Figure 7.37).

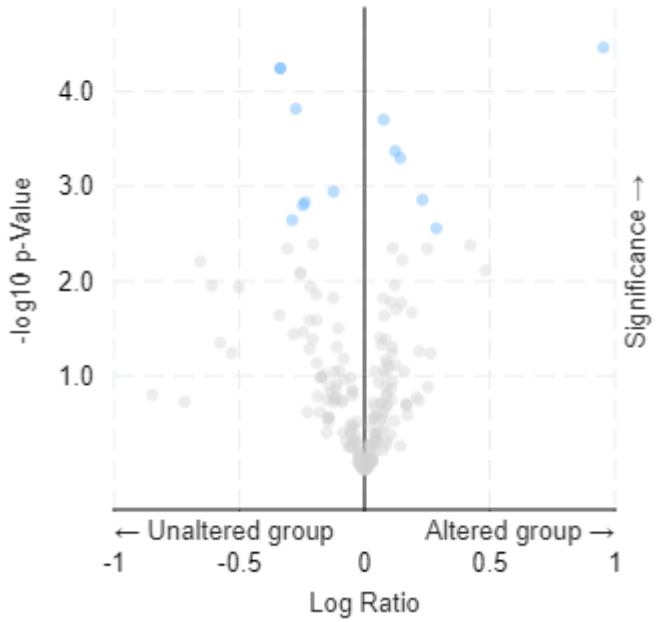


Figure 7.35 JAK alterations and protein associations. *Volcano plot showing associations between presence of ≥1 JAK mutation and available protein expression, significant associations highlighted in blue. There were data on 12 proteins which were significantly enriched in either JAK mutated or JAK wildtype patients. Analysis was performed in cBioPortal.*

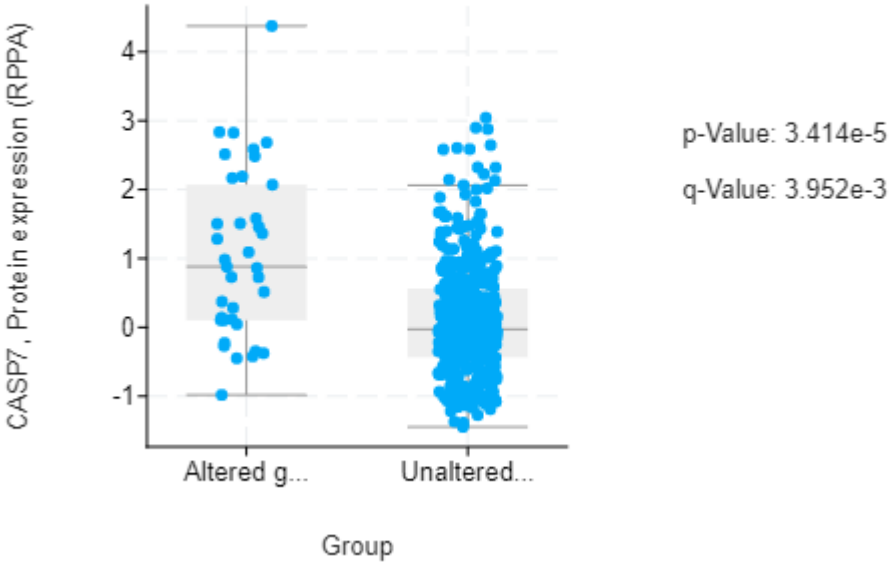


Figure 7.36 JAK alterations and Caspase 7. *Box plot showing Caspase 7 protein expression (reverse phase protein array) for JAK mutated and JAK wildtype patient groups (p<0.001). Analysis was performed in cBioPortal and data is publicly available.*

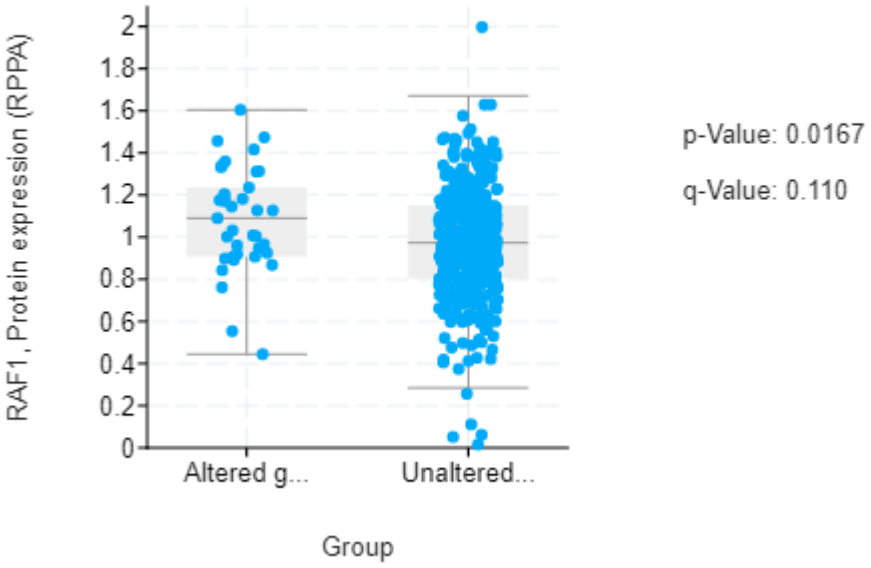


Figure 7.37 JAK alterations and RAF1. *Box plot showing RAF1 protein expression (reverse phase protein array) for JAK mutated and JAK wildtype patient groups (p=0.0167). Analysis was performed in cBioPortal and data is publicly available.*

7.8.2 STAT3 mutations, outcomes, and clinical factors

STAT3 mutation status was not in the custom panel utilised for sequencing the Glasgow cohort however data was available in the TCGA dataset. Presence of STAT3 mutation was observed in 2% of patients and alteration was associated with improved progression-free survival (p=0.0379, figure 7.38) but not cancer-specific or overall survival. Similarly, IL6 was mutated in only 1.6% of patients and alteration was trending towards an association with improved overall survival (p=0.0741) but not progression-free or cancer-specific survival (Figure 7.39). Mutations in none of the other IL6-STAT3 pathway members were associated with outcomes in this cohort.

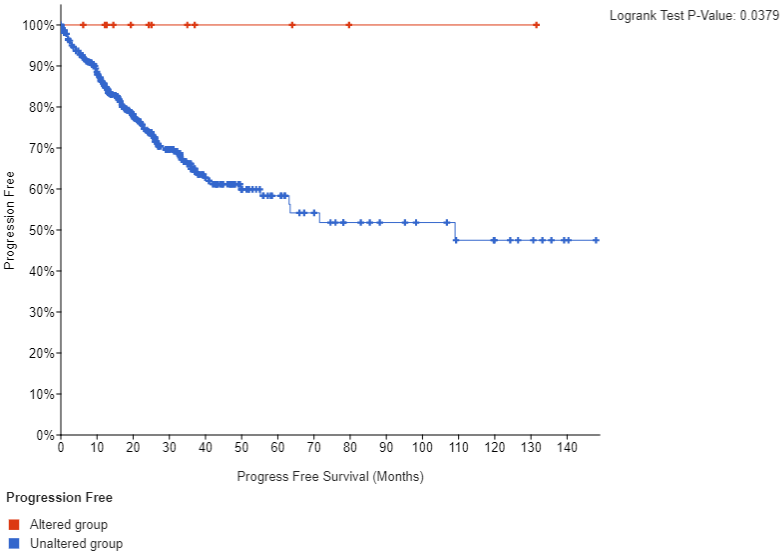


Figure 7.38 STAT3 alterations association with progression-free survival. *Kaplan Meier curve showing association between mutations in STAT3 gene and progression-free survival in TCGA PanCancer Atlas CRC cohort (p=0.0379). Red line represents patients with an alteration in the STAT3 gene and the blue line represents patients with no detectable alterations in the STAT3 gene. Analysis was performed in cBioPortal and is publicly available.*

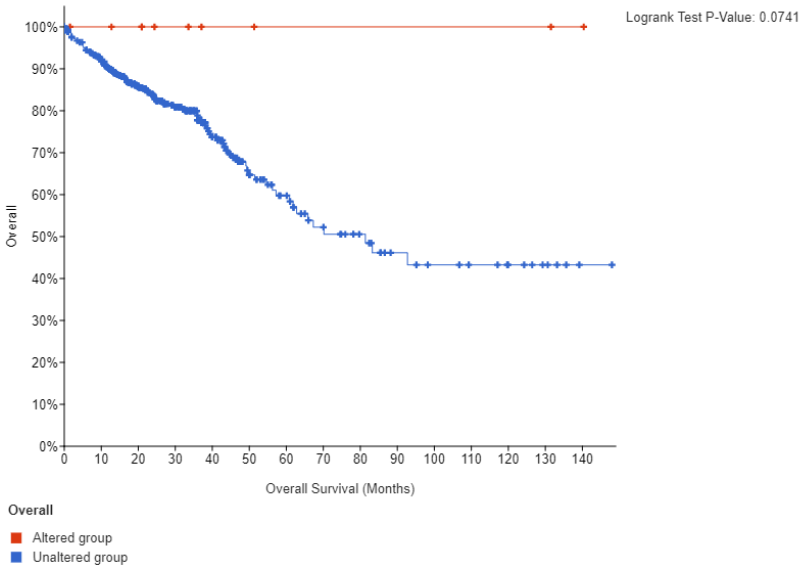


Figure 7.39 IL6 alterations association with overall survival. *Kaplan Meier curve showing association between mutations in IL6 gene and overall survival in TCGA PanCancer Atlas CRC cohort. The red line represents patients with an alteration in the IL6 gene and the blue line represents patients with no alteration in IL6 detected. Analysis was performed in cBioPortal and is publicly available.*

7.8.3 Co-expression of mRNA

Next, the mRNA expression levels of pathway members were analysed for association. Co-expression of JAK1 mRNA was moderately correlated with STAT3 mRNA expression (rho= 0.565, p=3.08e^{-51}, figure 7.40). JAK2 and JAK3 mRNA expression weakly correlated with STAT3 mRNA expression (rho= 0.355 p=5.34e^{-19} and rho=0.406 p=6.69e^{-25}, respectively, figures 6.41-42).

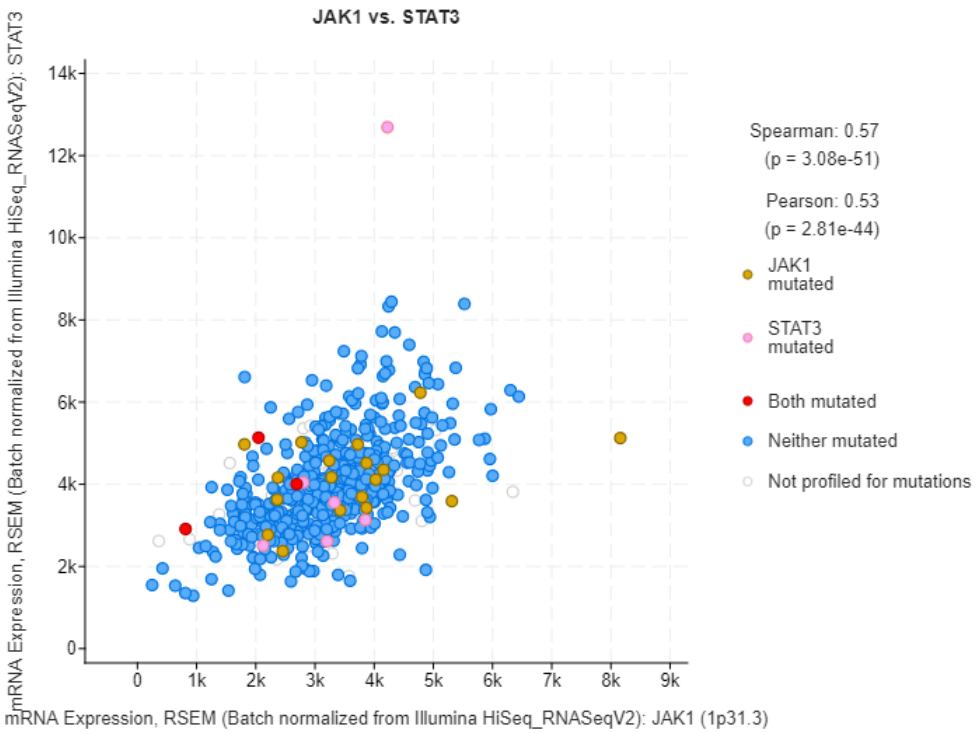


Figure 7.40 Correlation between JAK1 and STAT3 mRNA levels. *Kaplan Meier Scatter plot showing correlation between JAK1 and STAT3 at the mRNA level in TCGA PanCancer Atlas CRC cohort. Blue points correspond to wild type for both JAK1 and STAT3, red are both mutated, yellow only JAK1 and pink only STAT3 mutation/s present. Analysis was performed in cBioPortal and is publicly available.*

Figure 7.41 Correlation between JAK2 and STAT3 mRNA levels. *Kaplan Meier Scatter plot showing correlation between JAK2 and STAT3 at the mRNA level in TCGA PanCancer Atlas CRC cohort. Blue points correspond to wild type for both JAK2 and STAT3, red are both mutated, yellow only JAK2 and pink only STAT3 mutation/s present. Analysis was performed in cBioPortal and is publicly available.*

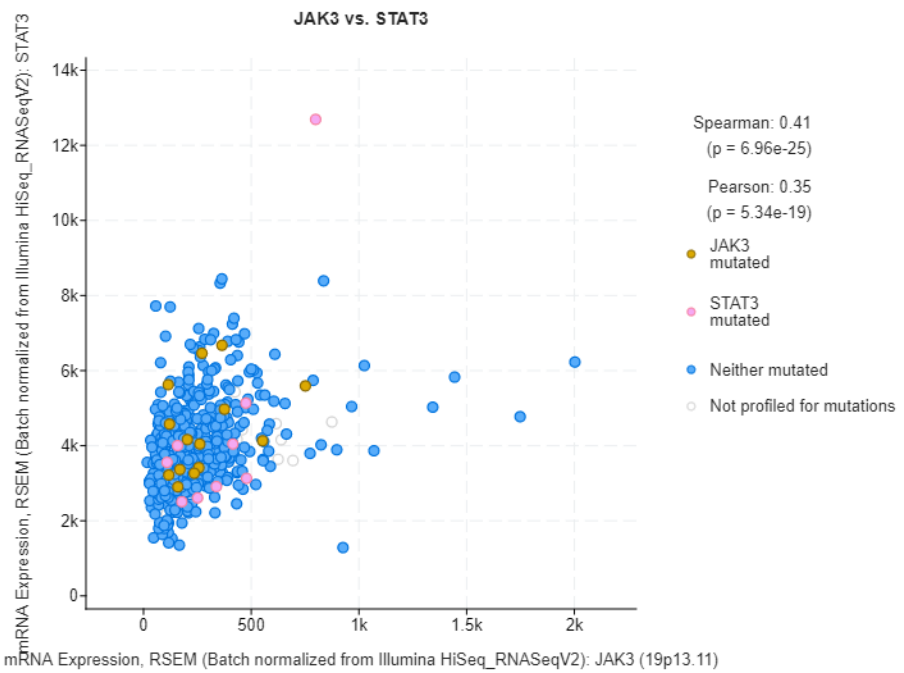


Figure 7.42 Correlation between JAK3 and STAT3 mRNA levels. *Kaplan Meier Scatter plot showing correlation between JAK3 and STAT3 at the mRNA level in TCGA PanCancer Atlas CRC cohort. Blue points correspond to wild type for both JAK3 and STAT3, red are both mutated, yellow only JAK3 and pink only STAT3 mutation/s present. Analysis was performed in cBioPortal and is publicly available.*

7.9 Discussion

Colorectal cancer can be characterised numerous genomic alterations. Mutational profiling is an important tool for understanding tumour heterogeneity and drivers of oncogenesis. Recent advances in the field have identified point mutations and mutational signatures that provide prognostic information. Technologies such as NGS required for genomic profiling are not yet feasible for routine clinical practise due to costs and time-consuming protocols. Therefore, analysing mutation profiles in retrospective clinical specimens and investigating associations with clinicopathological characteristics and measures of tumour phenotype is valuable.

Published studies have reported that presence of prominent CRC-associated mutations does affect tumour phenotype. For example, point mutations in WNT pathway members/β-catenin associate with reduced T cell infiltrates in the tumour microenvironment (TME) in CRC (168). Gain of function mutations in the TP53 gene are thought to modulate the TME in favour of increased stromal infiltrates

(169). KRAS mutations are associated with worse prognosis, and increased tumour proliferation (170).

We firstly investigated if there was a mutational profile underlying the different phenotypes observed with GMS. Several genes were mutated at a higher frequency in each of the GMS groups. Immune GMS0 tumours were characterised by increased alterations in MSI genes, MSH2 and MLH6, and Akt signalling pathway member PTEN. These patterns are consistent with mutations observed in CMS1(171). GMS1 tumours exhibited higher frequencies of mutation in DAXX, HER2 and MTOR. The DAXX gene interacts with WNT pathway member β-catenin and is therefore linked to invasion and metastases(172). Stromally-dense GMS2 tumours showed profoundly high TP53 mutations with over 70% of patients having alterations in the TP53 gene. Interestingly, mutant p53 is known to modulate tumour-stroma crosstalk resulting in a pro-tumour TME via induction of unfavourable cytokines and chemokines(173). Given the differences in the underlying biology of each phenotype it could be hypothesised that GMS classifications could predict response to targeted therapies.

In addition to investigating mutational landscape of each GMS phenotype, mutation patterns associated with STAT3 signalling were investigated. Evidence in the literature for a link between mutations in IL6-STAT3 signal transduction gene and CRC tumour phenotype is limited. JAK1 mutations are strongly associated with MSI in multiple primary cancer sites and loss of function mutations are strongly linked to immune evasion (174). In non-small cell lung cancer (NSCLC), activating JAK2 and JAK3 mutations have been shown to regulate PDL1 expression and ultimately reduce response to chemotherapy (175). However, other studies have suggested mutations in the JAK2 gene may be of little clinical significance in the CRC setting (176). Here, we have shown a strong association between mutations in the pathway of interest (JAK1/2/3) and clinical features such as MMR status ($p<0.001$) and tumour subsite ($p=0.001$). Further research in a larger validation cohort is required, due to the small number of patients exhibiting a mutation in 1 or more JAK genes in this cohort (11.6% of patients). Further evidence to support this was found by analysis of the TCGA PanCancer Atlas CRC cohort with JAK mutated patients more likely to be MSI-H, however in the Glasgow cohort mutation was associated with higher clinical stage and this was not seen in the TCGA cohort. Analysis of the TCGA dataset suggested a stronger link between mRNA expression of STAT3 and JAK1 than JAK2 or JAK3, similar to the observations made from protein (IHC) data from the Glasgow cohort.

To date, only one study has investigated associations between tumour mutational landscape and $pSTAT3^{tyr705}$ protein expression (177). In head and neck cancer patient specimens, no association between mutations, epigenetic modifications or copy number alterations and $pSTAT3^{tyr705}$ protein expression was observed (15). The main aim of this chapter was to try to identify a pattern of mutations associated with aberrant STAT3 signalling in CRC. The $pSTAT3^{tyr705}$ high and low groups

were defined using surv_cutpoint function of survminer in R studio with respect to CSS, which presents some limitations. Patients with a weighted histoscore near the cut off point for high/low expression may have similar underlying tumour biology. To pull out genomic differences underlying STAT3 activation it may be necessary to analyse only the patients at the extreme ends of either group, those observing the highest and lowest expression of pSTAT3^{tyr705}. The continuous weighted histoscore variable could also be used to determine association between mutation landscape and activation of STAT3. Due to the number of patients included in the initial profiling, analysis of the full 2 high/low pSTAT3 ^{tyr705} groups was deemed the most appropriate approach for the purposes of this chapter.

Mutational profiling of stage I-IV CRC patients with high pSTAT3^{tyr705} protein expression showed APC, TP53 and KRAS were the most frequently mutated genes, reflective of patterns seen in the literature for CRC (178). However, a significant association between increased mutations in MAP2K4 gene was observed in the hyper-activated STAT3 phenotypes. A link between MAPK and STAT3 signalling has previously been identified in pancreatic, oral, and gastric cancers. Inhibition of STAT3 in pancreatic ductal adenocarcinoma cell lines activated MAPK, which eventually led to reactivation of STAT3 despite consistent inhibitor treatment (179). Crosstalk between these pathways has been demonstrated via inhibition of ERK1/2 in oral cancer resulting in increased expression of pSTAT3^{tyr705} (180). In gastric cancer, activation of pro-tumour Neutrophils has been shown to occur via a STAT3-Erk1/2 cascade (181). This work coupled with observations from this study suggests dual targeting of STAT3 signalling and MAPK signalling should be investigated for solid tumours. In preclinical studies, dual inhibition utilising BRAF, MAK and ERK inhibitors with a JAK2 inhibitor in melanoma cell lines showed profound effects on inhibiting cancer growth (182).

In addition to MAPK signalling, WNT signalling alterations were observed in a high proportion of patients with activated STAT3 (87.71%), and AMER1 mutations were significantly enriched in these patients ($p=0.01$). There is limited evidence in the literature of cross talk between these pathways and WNT signalling represents a relatively challenging pathway to therapeutically target (183). Upon copy number alteration analysis, WNT signalling pathway member RNF43 was observed to be amplified in pSTAT3^{tyr705} high patients. Mutations in the RNF43 gene associate with poor prognosis in CRC via enhancing tumour growth and promoting disease recurrence (184). Recent data from murine models have shown a reduction in tumour growth through the intestinal epithelium when mice were treated with porcupine inhibitors targeting RNF43 mutations (185). Combination treatment with porcupine inhibitors and JAK inhibitors in preclinical CRC models may represent an interesting therapeutic approach to jointly target WNT and STAT3 signalling.

The results from this chapter show the presence of some differences in mutation patterns for each GMS group, and for high/low STAT3 activation phenotypes. Analysis of NGS mutation profiling in a

validation cohort is needed to corroborate these data. Further investigation is required to understand the significance of mutations in MAP2K4 and amplifications of RNF43 association with pSTAT3^{tyr705} protein expression to determine if there is any clinical utility in dual targeting of both pathways. In addition to mutation patterns underlying each phenotype, analysis of the underlying transcriptome should also be investigated to determine gene expression profiles.

8. Investigating differential gene expression in relation to pSTAT3^{tyr705} protein expression and Glasgow Microenvironment Score in colorectal cancer patients

8.1 Introduction

Given our findings from chapter 3 that GMS phenotyping was highly prognostic and chapter 4/5/6 that IL6/JAK/STAT3 signal transduction was associated with poor outcomes, here we aim to unravel any differences in the underlying transcriptome of these phenotypes. Colorectal cancer is associated with a variety of genetic alterations. Research has identified numerous differentially expressed genes between normal colon and tumour regions (186). Whole transcriptome bulk RNA sequencing was performed on 100 patient samples from the Glasgow combined CRC cohort utilising TempO-Seq profiling (BioSpyder Technologies, Carlsbad, CA, USA). This technique is highly specific and allows for RNA sequencing of formalin-fixed paraffin embedded tissue samples enabling easy profiling of retrospective patient cohorts (187). The aim of performing TempOSeq was to unravel any transcriptomic patterns associated with each GMS to investigate potential drivers of histological phenotype and therefore identify novel therapeutic targets for each GMS classification. Subsequent analysis included investigation of the transcriptome of pSTAT3^{tyr705} high/low protein groups with the aim of understanding mechanisms driving pathway hyperactivation and identification of potential candidates for dual therapeutic targeting. Probes for the full transcriptome were used enabling coverage of ~22000 genes in each sample. It was hypothesised that the underlying transcriptome of patients would be different between GMS phenotypic groups, and between patients with high and low pSTAT3^{tyr705} protein expression.

8.2 Cohort characteristics

A subgroup of the Glasgow combined cohort was included in the study with TempOSeq profiling performed on 100 patients. In this subset of the cohort the median survival time was 84.05 months. In terms of GMS, data were available for 87 patients with 36.7% of patients GMS0 immune, 32.2% GMS1 intermediate and 31.1% were classified as stromal-rich GMS2. The pSTAT3^{tyr705} protein status was available for 85 patients. There were 55.9% of patients categorised as high for pSTAT3^{tyr705} based on IHC data utilising the maxstat cut point of >33. Data were analysed in all 85 patients in the first instance and subsequently patients with the top 25 highest and lowest weighted histoscores were extracted and analysed separately. This was performed to remove any bias occurring due to patients scores falling near the threshold for low and high expression, allowing focus to be on the extremes of pSTAT3^{tyr705} expression.

8.3 Generation of differentially expressed genes files and analytical plan

TempOSeq assays and conversion from FastQ files was outsourced and performed by Bioclavis (Bioclavis Ltd, Glasgow, UK) to generate a raw gene counts file. Mahanolobis distance analysis

identified 9 potential patients' outliers from the 100 patient samples. Raw counts files were normalised in R Studio (RStudio, Boston, MA, USA) by dividing the number of reads per gene by the total number of reads for the sample. Differential gene expression analysis was performed using DESeq2 in R Studio (RStudio, Boston, MA, USA) on scaled counts data to generate tables of differentially expressed genes in relation to GMS groups and pSTAT3^{tyr705} high/low groups. Significance was set to adjusted p (p.adj) <0.05 and log2 fold >1. Gene counts for IL6/JAK/STAT3 pathway members were extracted and assessed for association with GMS and pSTAT3^{tyr705} IHC protein status in SPSS (IBM, NY, USA) via Kruskal Wallis non-parametric tests.

8.4 Transcriptomics analysis of GMS groups

Next, analysis of full transcriptomic data was performed relative to GMS groups to determine any differences in gene expression patterns underlying tumour histological phenotypes. GMS classifications were available for 96 of the 100 patients from the Glasgow combined cohort with TempOSeq gene profiling. After excluding 9 patient outliers, 87 were included in analysis. There were 36 GMS0 immune, 26 GMS1 intermediate and 25 GMS2 stromal patients. Principal component analysis and DESeq2 analyses was performed by Bioclavis (Bioclavis Ltd, Glasgow, UK). Principal component analyses revealed no clustering of gene expression relative to GMS groups (Figure 8.1).

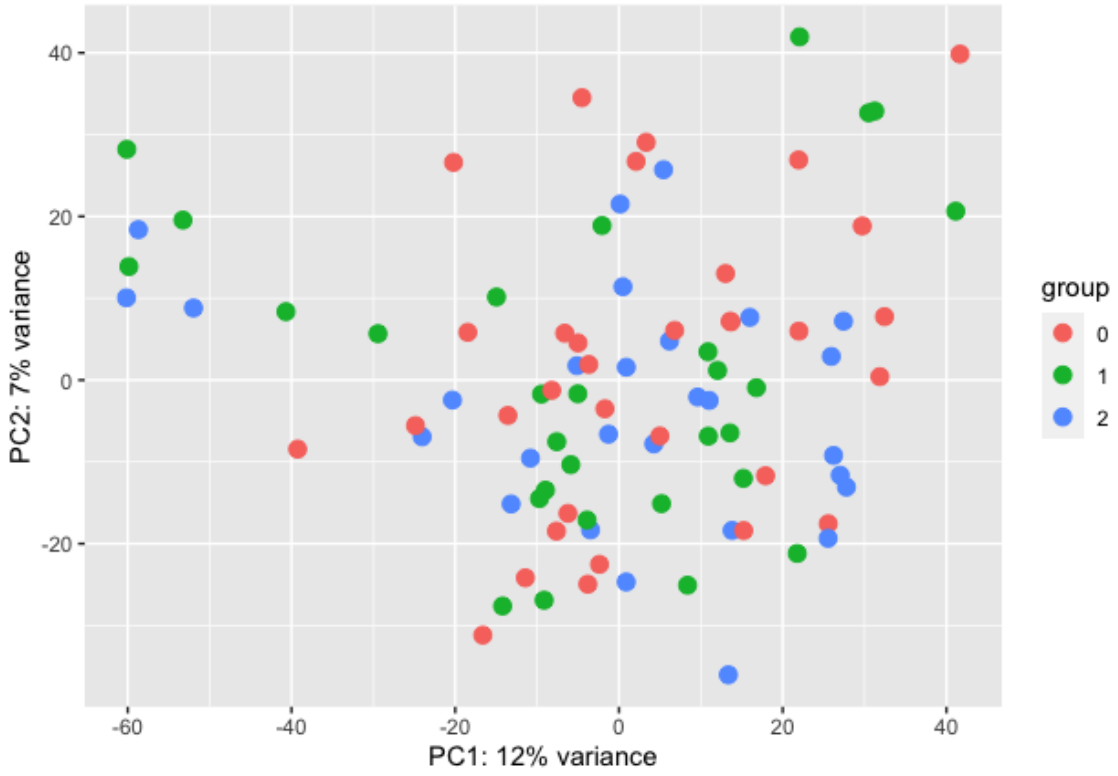


Figure 8.1 PCA plot for GMS classifications. *Unweighted principal component analysis plot showing gene clustering when patients were segregated into GMS0, GMS1 and GMS2 histological groups. GMS0 immune patients are represented by red points, GMS1 by green and GMS2 stromal-rich tumours by blue coloured points.*

When DESeq2 differential gene expression analyses were performed there were no significantly differentially expressed genes when GMS0 was compared to all other patients (GMS1 and GMS2 combined). When GMS1 was compared to all other patients there were 5 significantly differentially expressed genes identified; Free fatty acid receptor 2 (FFAR2), interleukin 1β (IL1β), Variable change X-linked (VCX), GRIA1and MAGEA6 at p.adj <0.05. When GMS2 gene profiles were compared to all other patients one significantly deregulated gene was identified, ADGR5. Raw gene counts for the significantly differentially expressed genes were extracted and added to an SPSS database. Box plots were constructed in SPSS (IBM, NY, USA) to visualise the expression profiles of significant genes relative to GMS groups on log base 10 scale.

FFAR2 is activated by factors produced by the microbiota and signalling is linked to regulation of inflammatory processes and epigenetic modifications. In mouse models of colon cancer loss of FFAR2 is associated with influx of Neutrophils to the TME and promotion of tumourigenesis (188). Patients classified as GMS1 showed reduced expression of FFAR2 compared to GMS0/2 patients as

shown in box plot (Figure 8.2). Nonparametric Kruskal Wallis testing in SPSS confirmed significant difference in gene counts between groups (p=0.004).

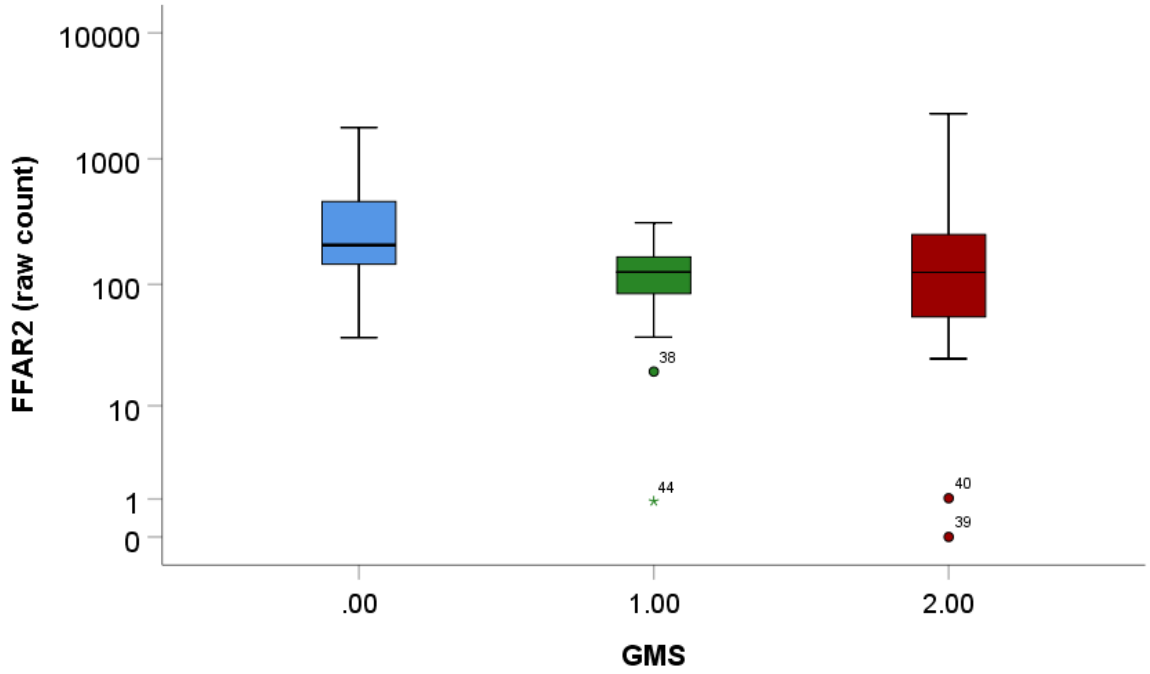


Figure 8.2 Expression of FFAR2 relative to GMS histological subtype. *Box plot showing the raw gene counts of FFAR2 in GMS0, GMS1 and GMS2 histologically subtyped patients. Gene counts were obtained from full transcriptome sequencing performed by TempOSeq in a subset of the Glasgow combined cohort.*

IL1β is an inflammatory cytokine implicated in activating the NF-κB master regulator, and is associated with improved prognosis in CRC (189). Here, IL1β gene counts were higher in GMS0/2 patients compared to GMS1 intermediate patients (Figure 8.3). Nonparametric Kruskal Wallis testing in SPSS confirmed significant difference in gene counts between groups (p=0.045).

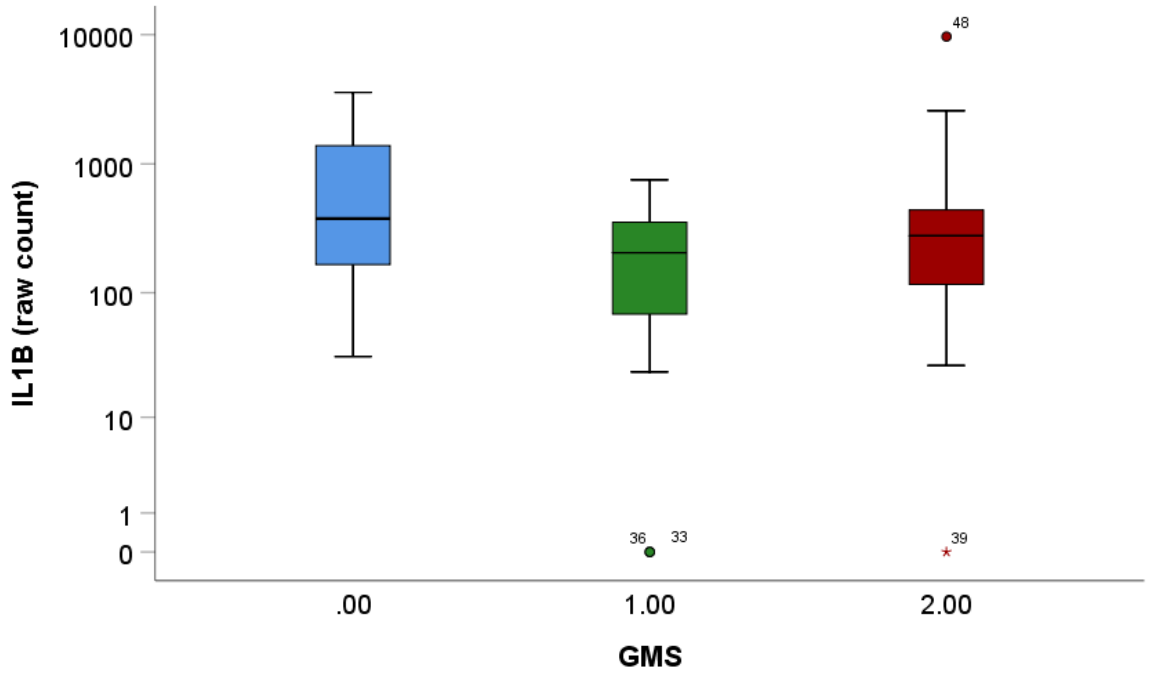


Figure 8.3 Expression of IL1β relative to GMS histological subtype. *Box plot showing the raw gene counts of IL1β in GMS0, GMS1 and GMS2 histologically subtyped patients. Gene counts were obtained from full transcriptome sequencing performed by TempOSeq in a subset of the Glasgow combined cohort.*

VCX is less well-characterised and has been reported to be rarely expressed in cancer with little evidence for its cellular function (190). In this cohort of CRC patients, expression of VCX was higher in GMS0 immune and GMS2 stromal patients than GMS1 patients. A box plot was constructed in SPSS to visualise the data (Figure 8.4). Nonparametric Kruskal Wallis testing in SPSS confirmed significant difference in gene counts between groups (p=0.003).

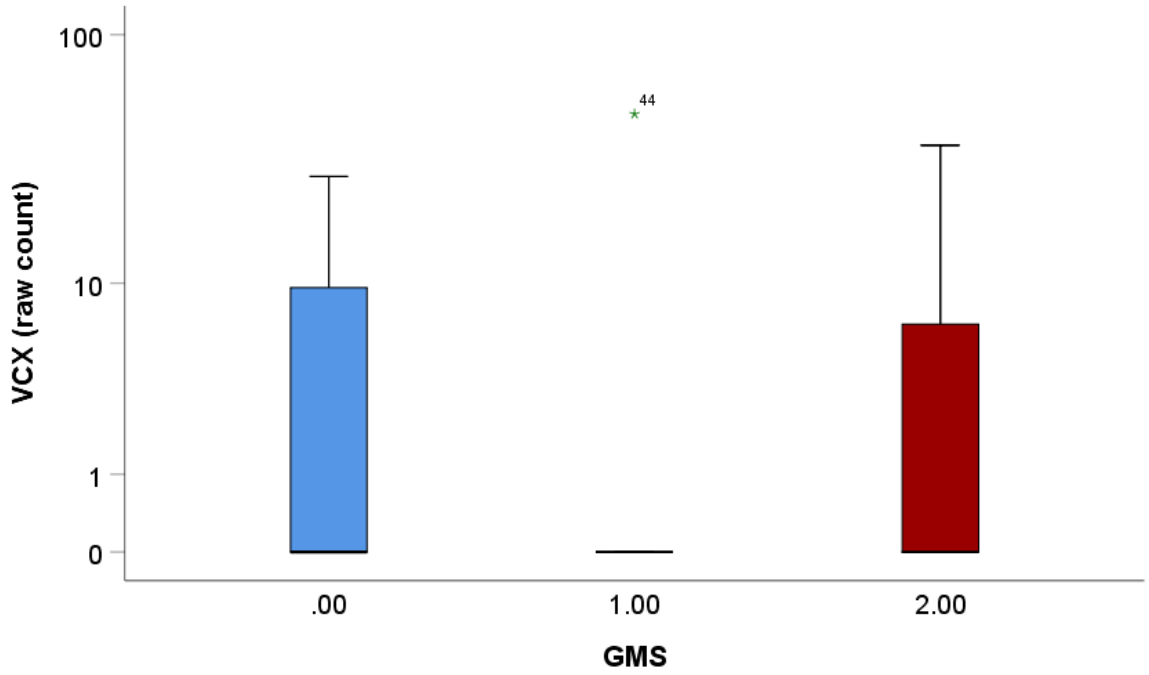


Figure 8.4 Expression of VCX relative to GMS histological subtype. *Box plot showing the raw gene counts of VCX in GMS0, GMS1 and GMS2 histologically subtyped patients. Gene counts were obtained from full transcriptome sequencing performed by TempOSeq in a subset of the Glasgow combined cohort.*

Glutamate ionotropic receptor to AMPA type subunit 1 (GRIA1), is a neurotransmitter and growth factor that can activate cellular signalling pathways such as EGFR, and high mRNA expression correlates with reduced overall survival in basal-like bladder cancer (191). In this cohort of CRC patients, GMS0 immune and GMS2 stromal patients observed higher expression of the GRIA1 gene than GMS1 intermediate classified patients. A box plot was constructed to visualise these results (Figure 8.5). Nonparametric Kruskal Wallis testing in SPSS confirmed significant difference in gene counts between GMS groups in terms of expression of GRIA1 ($p=0.026$).

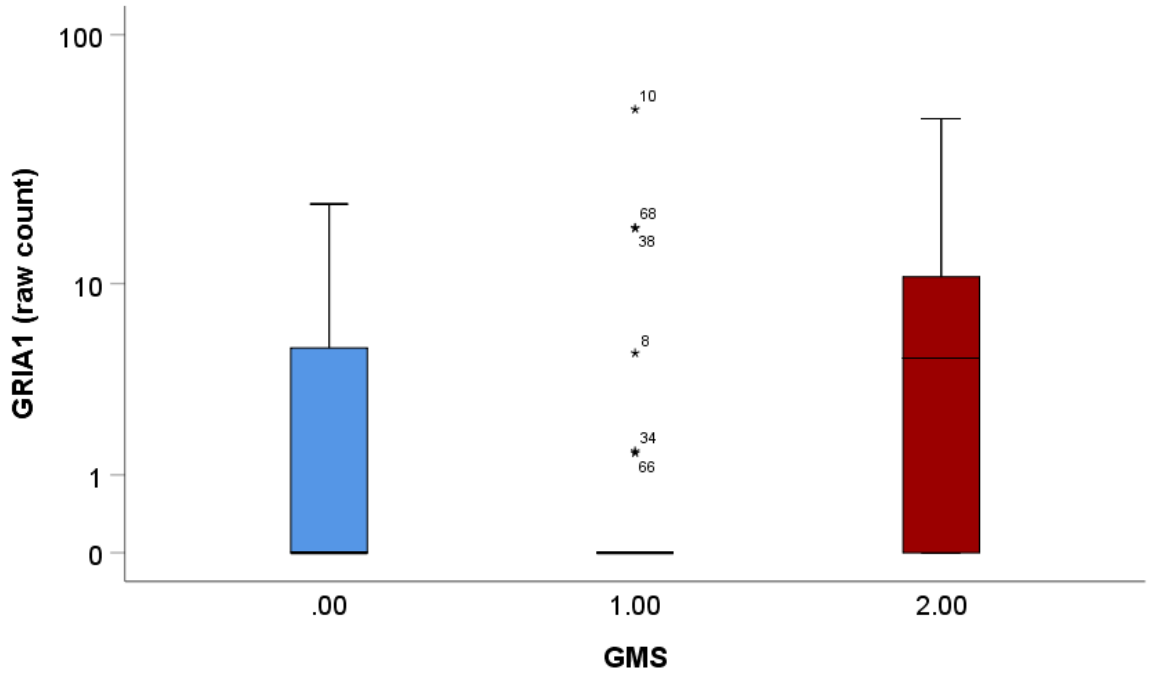


Figure 8.5 Expression of GRIA1 relative to GMS histological subtype. *Box plot showing the raw gene counts of GRIA1 in GMS0, GMS1 and GMS2 histologically subtyped patients. Gene counts were obtained from full transcriptome sequencing performed by TempOSeq in a subset of the Glasgow combined cohort.*

Melanoma-associated antigen family A 6 (MAGEA6) acts as a ubiquitin ligase AMP-activated protein kinase (AMPK) and is overexpressed at both the mRNA and protein level in tumour cells of gastric cancer (192). Raw counts of the MAGEA6 gene were much higher than the other significantly differentially expressed genes. A box plot was constructed in SPSS using a logarithmic scale to visualise the raw gene counts of MAGEA6 for each GMS group (Figure 8.6). When DESeq2 analysis was performed there was a significant difference in expression of the gene between GMS1 and the other patients, however when nonparametric Kruskal Wallis testing was performed in SPSS this significant difference was not validated (p=0.207).

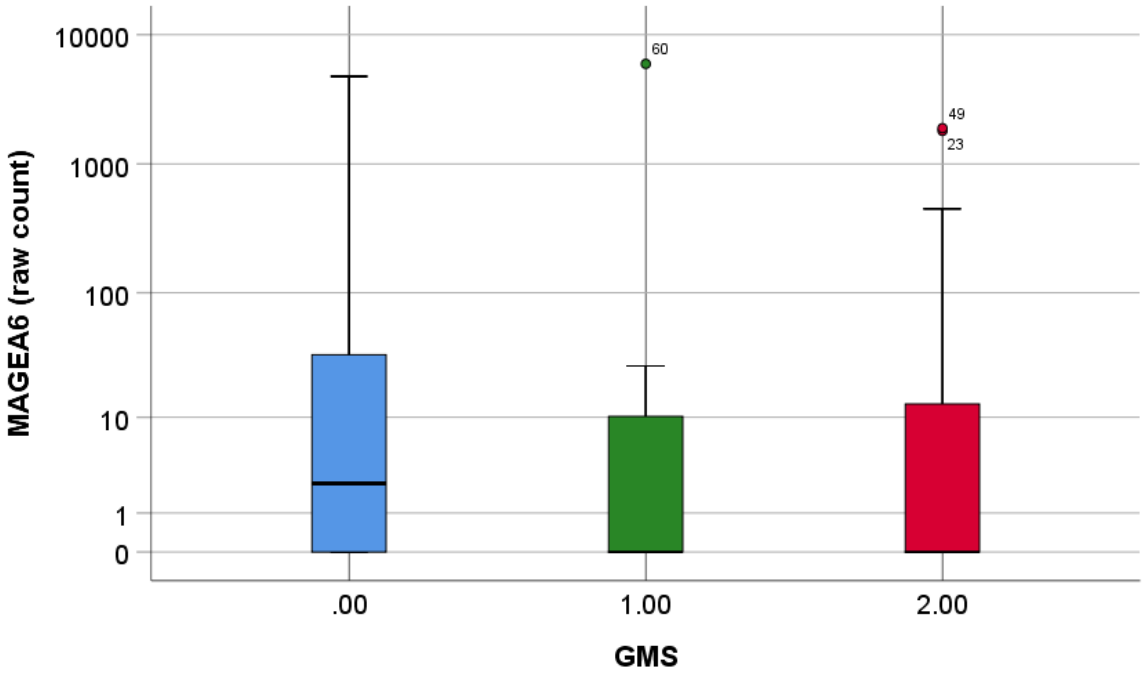


Figure 8.6 Expression of MAGEA6 relative to GMS histological subtype. *Box plot showing the raw gene counts of MAGEA6 in GMS0, GMS1 and GMS2 histologically subtyped patients on a logarithmic scale. Gene counts were obtained from full transcriptome sequencing performed by TempOSeq in a subset of the Glasgow combined cohort.*

When the DE genes were analysed for interactions using STRING analysis, GRIA1 showed association with IL1β, which could be linked to FFAR2 via IL10 (Figure 8.7). There were no protein-protein interactions that could be linked with MAGEA6 and VCX. Overrepresentation analysis (ORA) was performed using the enrichGO function in R studio (IBM, NY, USA). Only 2 of the 5 genes analysed could be linked, and there was a strong inflammatory component which was identified through IL1β and FFAR2 as shown in cnet plot (Figure 8.8) and enrichment dot plot (Figure 8.9) (adj. p = 0.011). This suggests patients with GMS1 tumours have less immunogenic tumours than GMS0/GMS2 tumours.

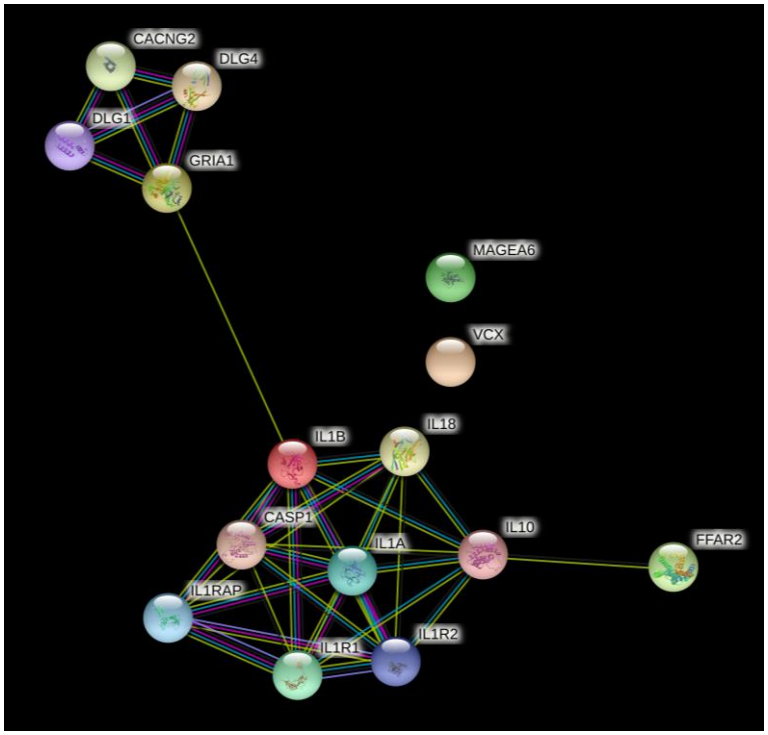


Figure 8.7 String map of DE genes between GMS1 and GMS0/2. *String interaction network diagram showing relationships between the differentially expressed genes between GMS1 compared to GMS0/2 groups from full transcriptional sequencing on a subset of the Glasgow combined cohort.*

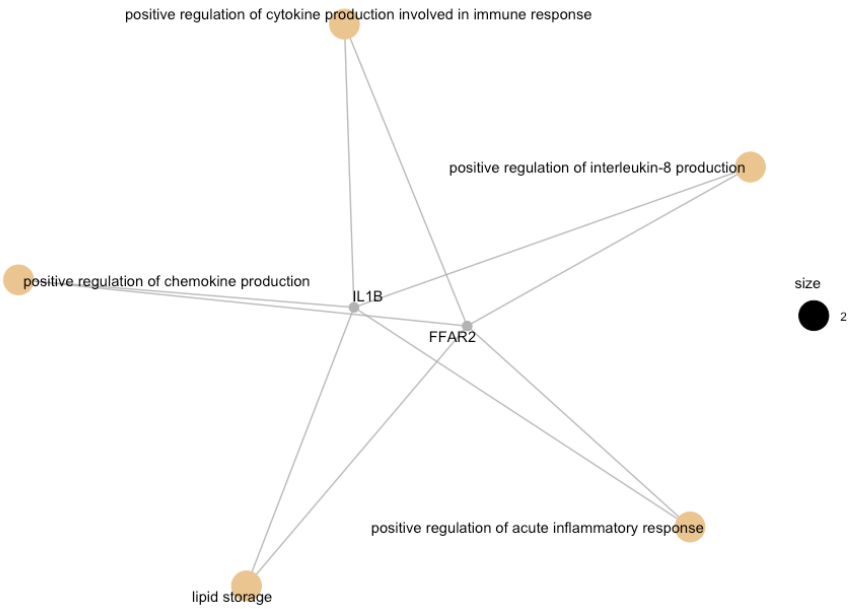


Figure 8.8 Enrichment cnet plot for DE genes. *Enrichment plot showing the network of processes linking the genes identified as differentially expressed between GMS1 and GMS0/2 groups. GMS1 showed differential expression of genes associated with lipid storage, positive regulation of acute inflammatory response, IL8 production, regulation of cytokines involved in the immune response and chemokine production.*

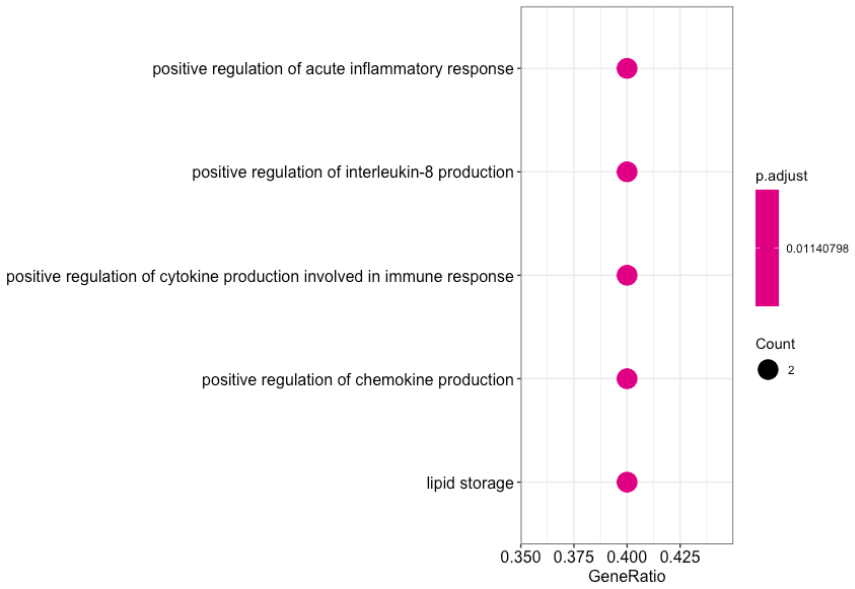


Figure 8.9 Enrichment dot plot for DE genes. *Enrichment dot plot showing the network of processes linking the genes identified as differentially expressed between GMS1 and GMS0/2 groups (p=0.0114).*

8.4 Analysis of IL6/JAK/STAT3 pathway components relative to tumour phenotype

Raw counts for the probes IL6_3341, JAK1_3463, JAK2_3466, STAT3_6829, IL6ST_17219 and IL6R_3341 were extracted from the full transcriptome data and analysed relative to pSTAT3^{tyr705} high/low groups and GMS histological subtypes. A string interaction network plot was constructed using string version 11.0 (STRING consortium) to visually represent the pathway members being investigated (Figure 8.10). Pink lines linking genes show interactions which have been experimentally determined, green lines show interactions determined via textmining and blue lines linking genes show interactions confirmed from curated databases. IL6ST represents the gene that encodes for the gp130, which binds to IL6R to form the signal transduction complex.

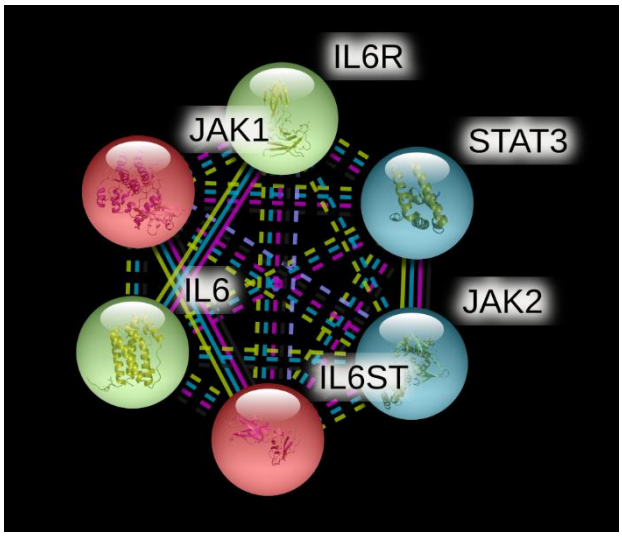


Figure 8.10 String map of pathway members. *String interaction network diagram showing pathway genes in TempOSeq panel.*

8.4.1 IL6/JAK/STAT3 RNA in pSTAT3^{tyr705} protein groups

Patients were divided in into 2 groups based on IHC protein data for pSTAT3^{tyr705} and mean values for each pathway member probe were calculated for high/low groups. A bar chart was plotted in Microsoft Excel (Microsoft, Albuquerque, NM, USA) to visualise mean counts (Figure 8.11). Non-parametric K independent Kruskal-Wallis tests were performed in SPSS version 22 (IBM, NY, USA) to determine any statistical differences between patients classified as high and low for pSTAT3^{tyr705} tumour protein. Significance was set to $\alpha<0.05$. No statistically significant differences were observed between counts from any gene and pSTAT3^{tyr705} groups (Figure 8.11). Independent T tests were also performed and no significant association between gene counts and pSTAT3^{tyr705} status were identified.

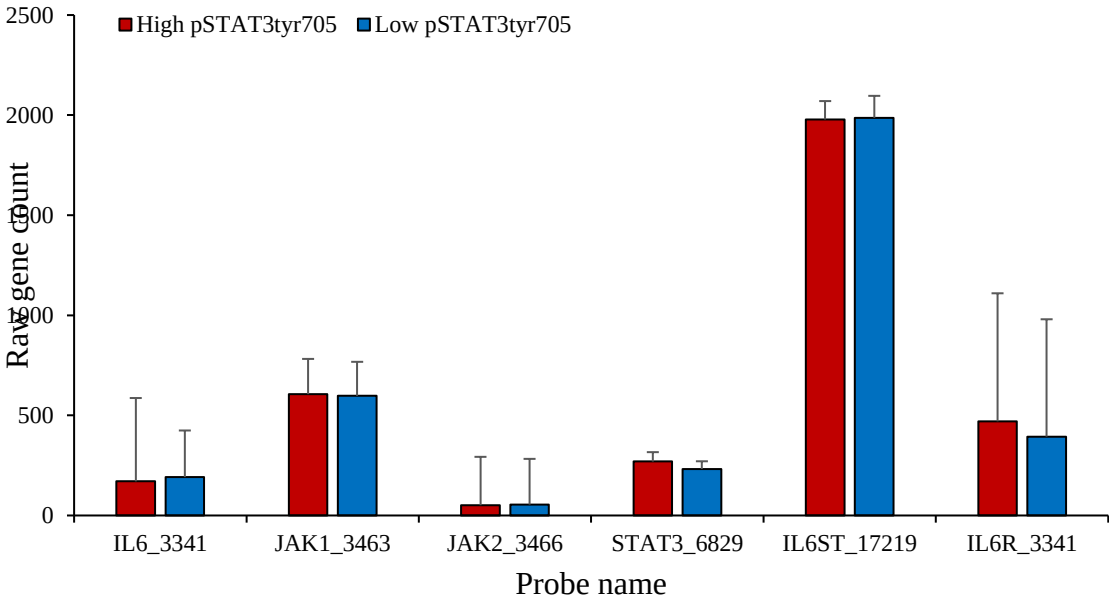


Figure 8.11 Pathway gene counts relative to pSTAT3^{tyr705} status. *Bar chart showing mean raw gene counts for members of the IL6/JAK/STAT3 pathway for high and low pSTAT3^{tyr705} protein groups. Gene counts were obtained from full transcriptome sequencing performed by TempOSeq in a subset of the Glasgow combined cohort.*

8.4.2 IL6/JAK/STAT3 RNA in GMS histological groups

Patients were then segregated into Glasgow Microenvironment Score histological groups. Kruskal Wallis non-parametric testing was performed in SPSS version 22 (IBM, NY, USA) and no significant differences in raw counts of any pathway probe were identified across GMS classifications. A bar chart was plotted in Microsoft Excel (Microsoft, Albuquerque, NM, USA) to visualise the data (Figure 8.12). One-way ANOVAs were performed in SPSS version 22 (IBM, NY, USA) and no significant differences were observed for any genes counts relative to GMS histological subtype.

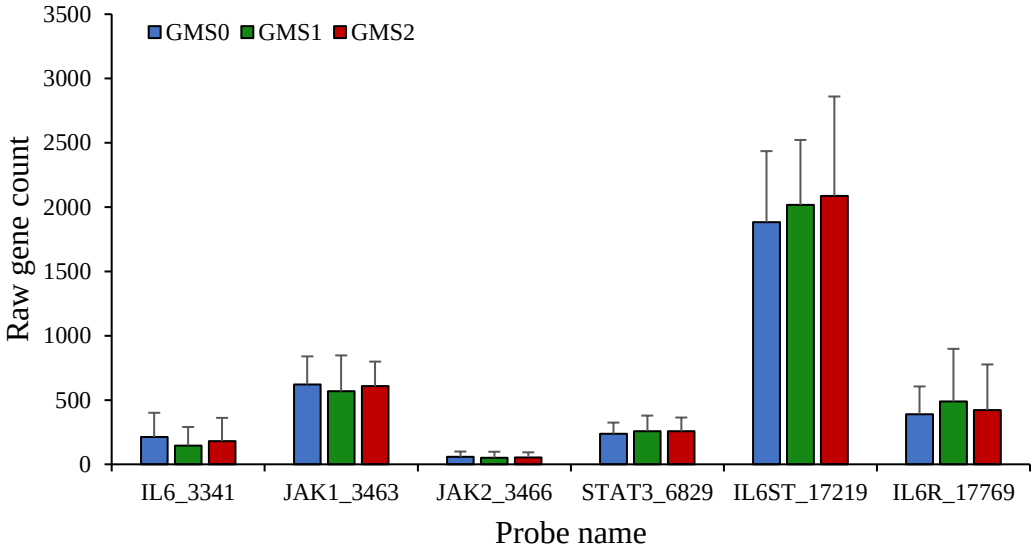


Figure 8.12 Pathway gene counts relative to GMS. *Bar chart showing mean raw gene counts for members of the IL6/JAK/STAT3 pathway for Glasgow Microenvironment Score groups. Gene counts were obtained from full transcriptome sequencing performed by TempOSeq in a subset of the Glasgow combined cohort.*

8.6 Transcriptomic profiles of pSTAT3^{tyr705} groups

8.6.1 Full cohort of pSTAT3^{tyr705} patients

There were 85 patients included in the TempOseq assay with pSTAT3^{tyr705} protein expression from IHC data in chapter 6. These 85 patients were extracted from the full cohort raw counts file and a sample sheet was built to align each patient with respective pSTAT3^{tyr705} high/low status. The counts and sample sheet files were loaded into R Studio (RStudio, Boston, MA, USA) and DESeq2 analysis was performed. There was some potential clustering of gene expression observed between pSTAT3^{tyr705} groups when principal component analysis was performed by Bioclavis (Bioclavis Ltd, Glasgow, UK) (Figure 8.13). Overall, there were no significantly differentially expressed genes between groups at p.adj <0.05 as shown in volcano plot constructed by Dr Gerard Lynch using ggplot in R Studio (RStudio, Boston, MA, USA) (Figure 8.14).

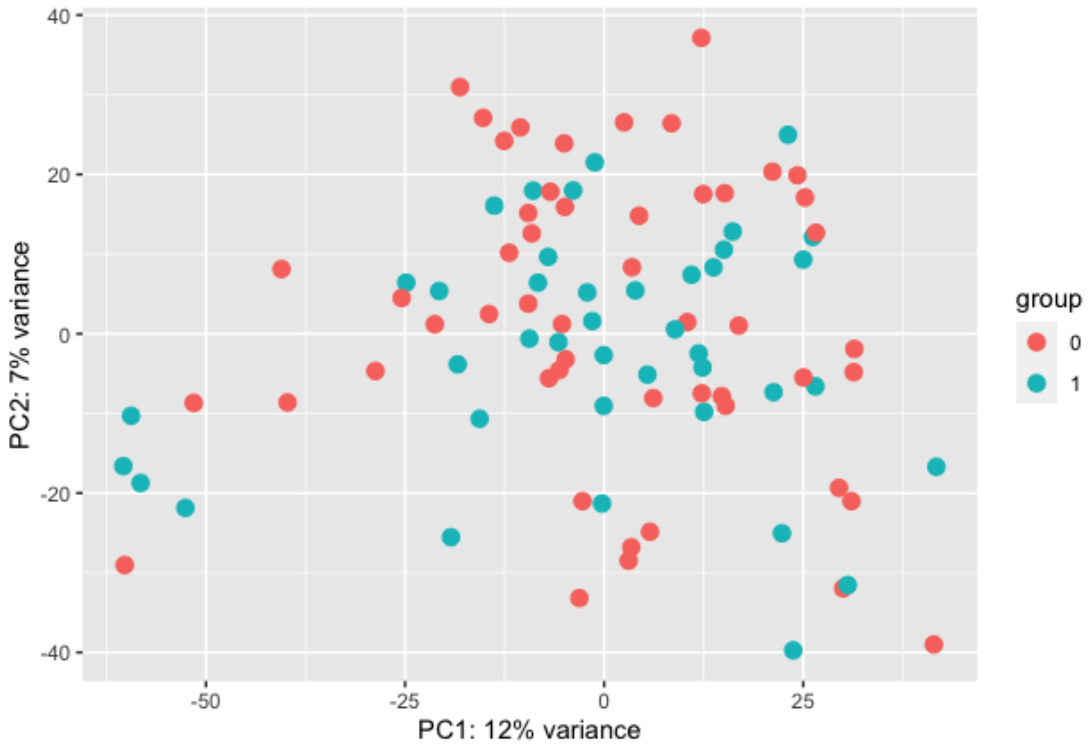


Figure 8.13 PCA plot for pSTAT3^{tyr705} classifications. *Principal component analysis plot showing gene clustering when patients were segregated into high and low for protein expression of pSTAT3^{tyr705} groups. Gene counts were obtained from full transcriptome sequencing performed by TempOSeq in a subset of the Glasgow combined cohort.*

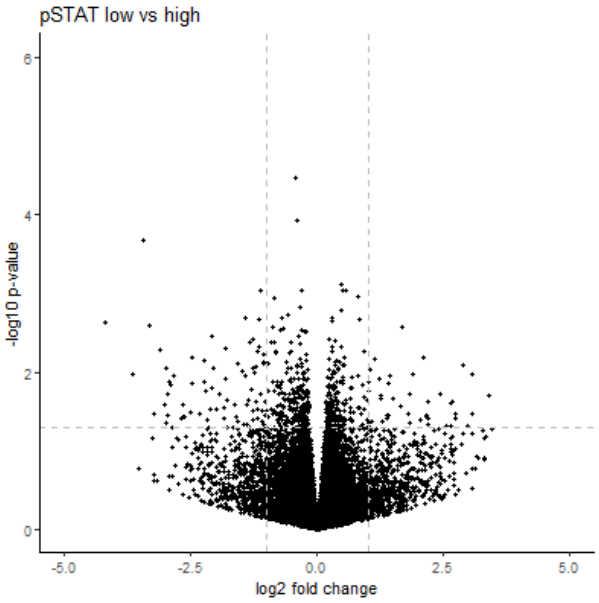


Figure 8.14 Volcano plot for full cohort based on pSTAT3^{tyr705} protein status. *Differential gene expression analysis on the full cohort relative to pSTAT3^{tyr705} group shown in a Volcano plot. Gene counts were obtained from full transcriptome sequencing performed by TempOSeq in a subset of the Glasgow combined cohort.*

8.6.2 Extremes of pSTAT3^{tyr705} patients

The 85 patients included in the first step of pSTAT3^{tyr705} analyses contained some patients with protein scores close to the cut point of ≥33 for high/low expression. Given that no significantly differentially expressed genes were identified from analysis of the full 85 patients, analysis was conducted again focused on the top 25 highest and lowest IHC-based scores for pSTAT3^{tyr705}, referred to as the cohort 'extremes'. This results in 49 patients being included due to 3 patients having the same weighted histoscore, and 24 was deemed more appropriate than selecting 27 patients for one group. However again, there were no significantly differentially expressed gene identified between pSTAT3^{tyr705} groups as shown in volcano plot constructed by Dr Gerard Lynch (Figure 8.15).

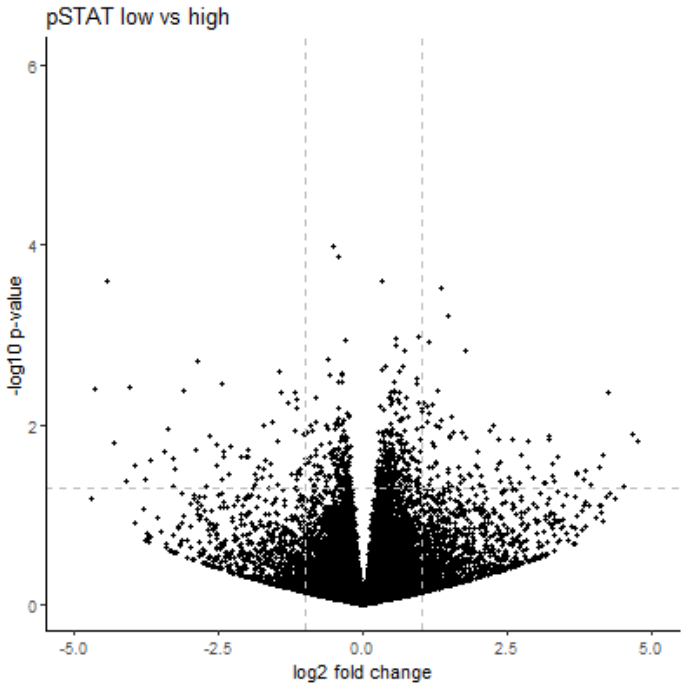


Figure 8.15 Volcano plot for full cohort based on extremes of pSTAT3^{tyr705} protein status. *Differential gene expression analysis on the extremes (top 25 highest and lowest weighted histoscore) of pSTAT3^{tyr705} groups shown in a Volcano plot. Gene counts were obtained from full transcriptome sequencing performed by TempOSeq in a subset of the Glasgow combined cohort.*

8.7 Single sample Gene Set Enrichment Analysis

8.7.1 Full cohort ssGSEA

Analysis of differences in expression of single genes between pSTAT3^{tyr705} groups yielded no differential expression. Data were then analysed to determine if common cancer-associated pathways of genes were dysregulated specific to high/low pSTAT3^{tyr705} groups in the full cohort. Single sample

gene set enrichment (ssGSEA) was performed by Sidhir Malla and Dr Philip Dunne. The normalised gene counts file (by DESeq2) was transformed using the variance stabilising method in RStudio (RStudio, Boston, MA, USA) to obtain a vst matrix from the count's matrix. A heatmap was constructed to visualise the top dysregulated pathways amongst the full 100 patients using heatmap.2 in R Studio (RStudio, Boston, MA, USA) (Figure 8.16). Further downstream analysis was performed in the form of microenvironment cell populations counter (MCP) analysis to gain a measure of different types of infiltrating cell populations to the TME. A heatmap was constructed in R studio using heatmap.2 in RStudio (RStudio, Boston, MA, USA) to visualise the different cell populations abundance in the full 100 patients (Figure 8.17). These data are similar to what would be expected from a CRC cohort highlighting heterogeneity of gene expression in cancer-associated pathways and in the types of tumour-infiltrating cells.

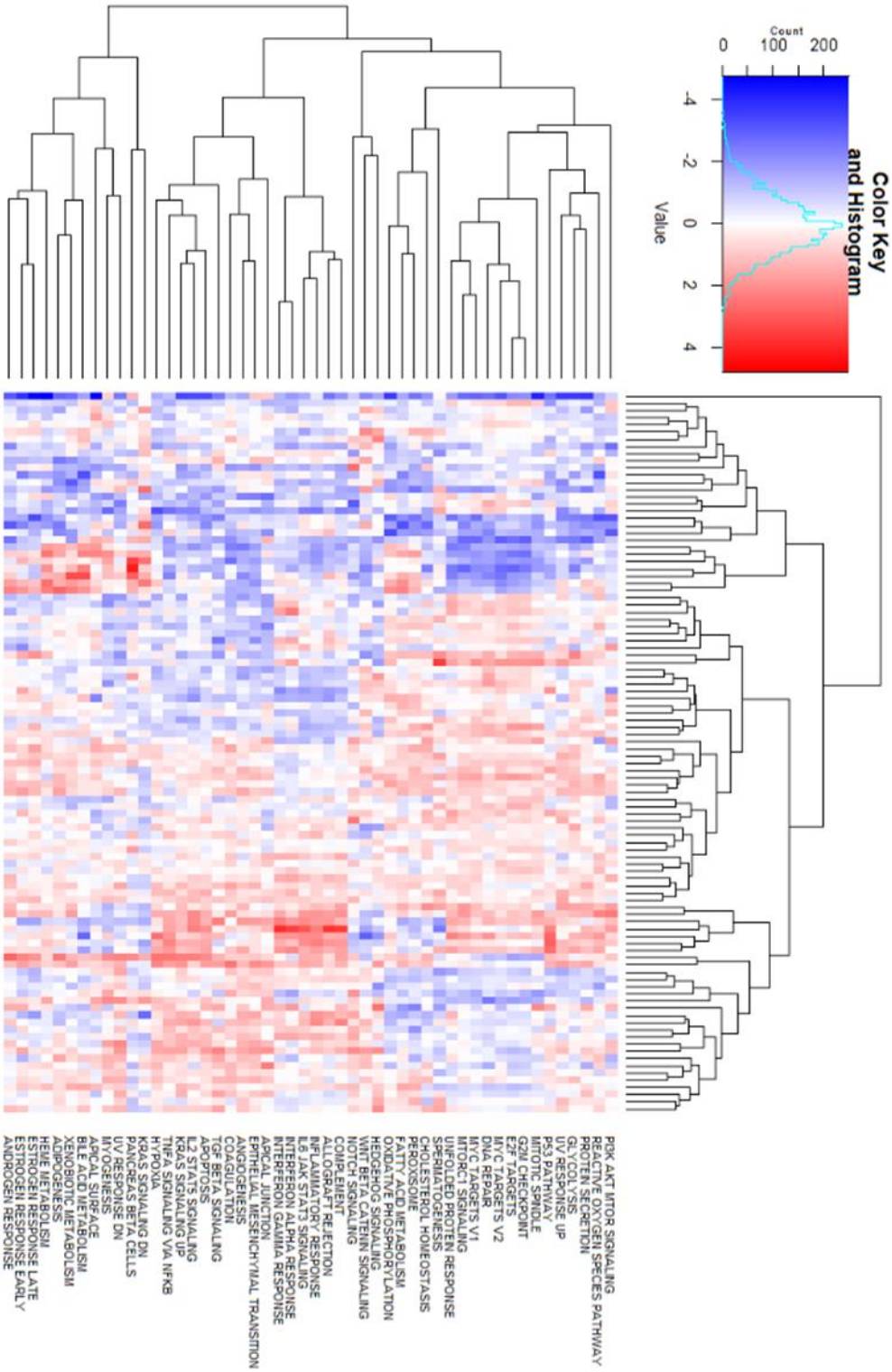


Figure 8.16 Full cohort single cell gene set enrichment. *Heatmap showing the patterns of gene expression of common cancer-associated pathways as measured by single cell gene set enrichment analysis for full the full cohort. Plot was constructed by Sidhir Malla.*

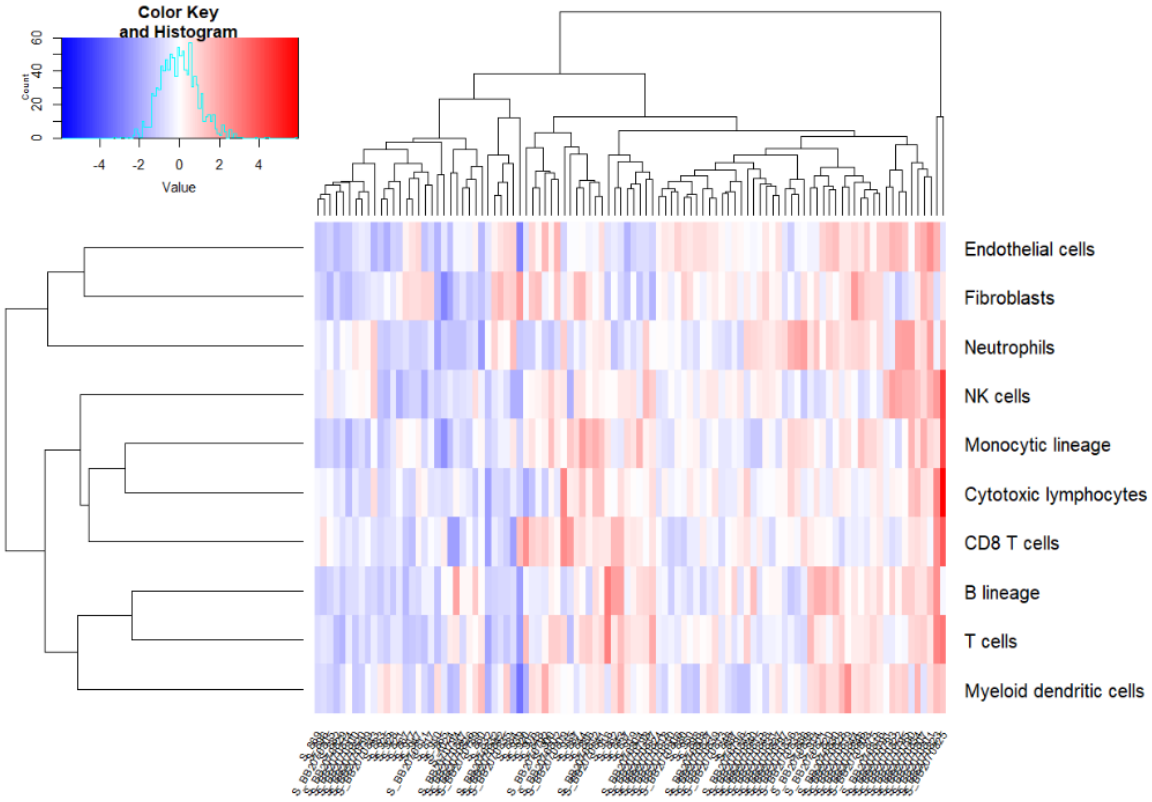


Figure 8.17 Microenvironment cell populations counter analysis of the full cohort. *Heatmap showing the patterns of gene expression used to determine infiltrating cell populations by MCP analysis in the full cohort. Plot was constructed by Sidhir Malla.*

Next, analysis was performed to determine if pathway gene set enrichment was associated with pSTAT3^{tyr705} protein groups. High and low pSTAT3^{tyr705} groups were overlayed onto a ssGSEA heatmap using ComplexHeatmap in R studio. No patterns of pathway dysregulation were associated with pSTAT3 status (Figure 8.18). Similarly, when MCP analysis was related to pSTAT3^{tyr705} protein status there were no patterns observed in terms of infiltrating cell population between groups (Figure 8.19).

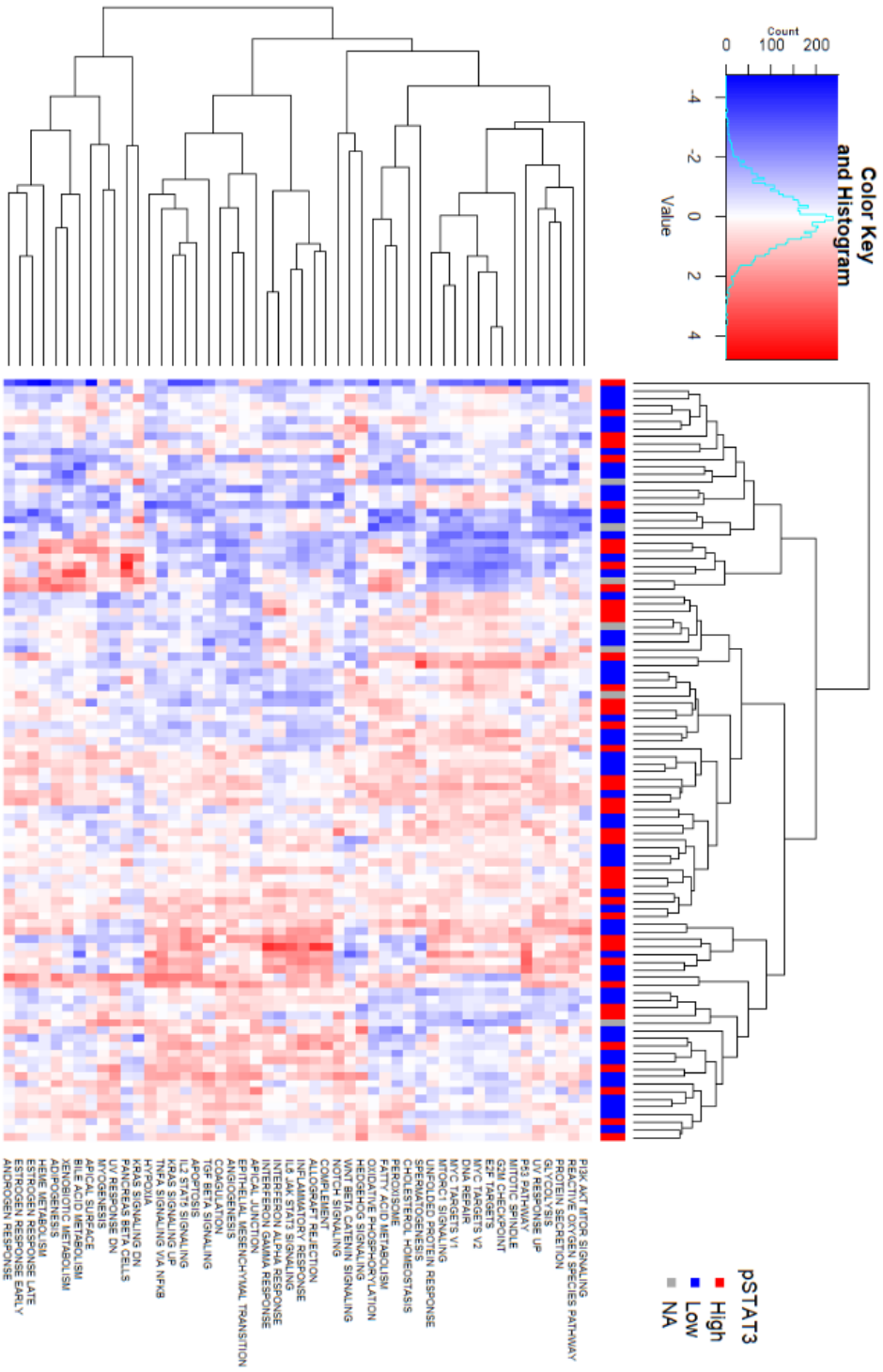


Figure 8.18 Full cohort single cell gene set enrichment overlayed with pSTAT3^{tyr705} status.

Heatmap showing the patterns of gene expression of common cancer-associated pathways as measured by single cell gene set enrichment analysis for full the full cohort with each patients pSTAT3^{tyr705} protein status overlayed. Plot was constructed by Sidhir Malla.

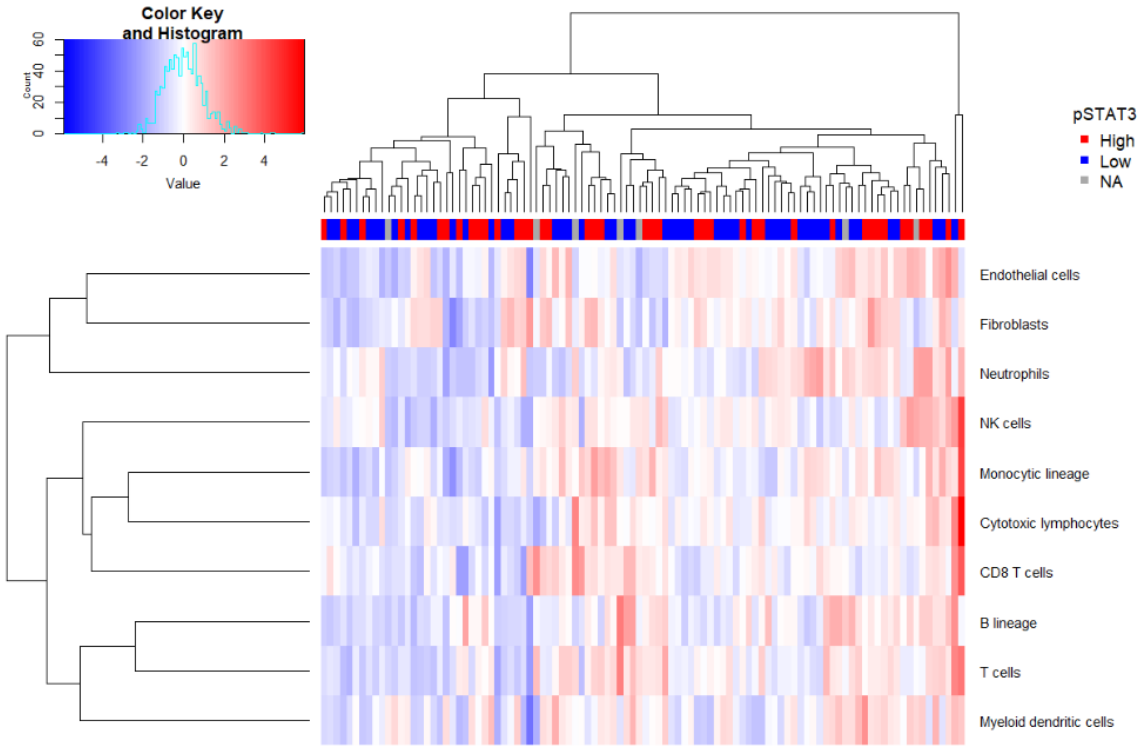


Figure 8.19 Microenvironment cell populations counter analysis of the full cohort with pSTAT3^{tyr705} status overlayed. *Heatmap showing the patterns of gene expression used to determine infiltrating cell populations by MCP analysis in the full cohort with pSTAT3^{tyr705} protein status overlayed. Plot was constructed by Sidhir Malla.*

8.6 Discussion

Results from previous chapters have implicated a strong role for IL6/JAK/STAT3 signalling in CRC at the protein level particularly in GMS2 patients, and a profound association between GMS and cancer-specific survival. We therefore aimed to investigate IL6/JAK/STAT3 at the transcriptomic level in a sub cohort of CRC patients, and determine any patterns of underlying biology associated with GMS histological subtypes and STAT3 activation at the protein level.

Genes included in the probe panel associated with IL6/JAK/STAT3 signal transduction included IL6_3341, JAK1_3463, JAK2_3466, STAT3_6829, IL6ST_17219 and IL6R_334. When raw counts were assessed, there was no association between gene counts of pathway members and GMS or pSTAT3^{tyr705} protein status. There were not enough patients included in the study to perform survival

analysis. We hypothesised that distinct gene expression patterns would be associated with driving tumour phenotype, specifically IL6/JAK/STAT3 genes and stromal-rich GMS2, however this was not observed in the subset of patients from the Glasgow combined cohort assessed. Interestingly, of the pathway genes assessed, IL6ST showed the highest expression amongst all patients, and the protein this gene encodes for (gp130) is currently under investigation for therapeutic targeting in preclinical models of CRC (193).

When the full transcriptome was assessed for gene enrichment relative to GMS classification, there were 5 differentially expressed genes identified between GMS1 versus GMS0/2. Gene set ontology enrichment analysis suggested a strong inflammatory response link between the five genes, due to higher expression of IL1β and FFAR2 in GMS0/2 patients. These data suggest GMS1 patients may harbour an immune-related anomaly whereby there is dysfunction in production of cytokines, chemokines, and the acute response. However, this study is unpowered and future work should look to validate these findings in a larger cohort. Another factor to consider is the overlap in some phenotypic measures between GMS groups. For example, patients with a high KM grade are automatically assigned to the GMS0 group, however these patients may have a high or low stromal component to their tumours. Future analysis could include analysing the RNAseq data focused on the component histological scores of GMS (TSP and KM grade) in a larger cohort of patients.

There are few studies in the literature which have looked to determine transcriptomic patterns associated with other measures of CRC tumour histology. Tumour budding is a histological score assessed from an H&E full section, which can be used to determine patient prognosis in CRC. Studies have investigated the underlying biology of budding phenotypes and found a strong correlation with KRAS mutations, β-catenin expression, and microsatellite stability (194, 195). Interestingly, when tumour budding was assessed in >1300 patient samples, budding was enriched in CMS4 patients (194). This study again emphasises the need for a larger cohort to be used when analysing the transcriptome of GMS groups in future work. In breast cancer, a recent novel study investigated if tumour morphology from an H&E section of breast cancer could be utilised to predict gene expression profiles (196). Of the 17, 695 genes assessed, 9334 were predicted by morphological assessment of H&E, as validated by matched RNA-sequencing(196). In multiple cancer types it is possible to infer MSI status from a routine H&E section based on a deep learning approach (197). Transcriptomic datasets from lung cancer patients have revealed genes involved in cell cycle regulation and nucleotide binding by KEGG pathway analysis can predict histological grade in terms of tumour differentiation(198). In the same study it was found that an algorithm built on features of histology such as cell size, shape and intensity distribution was able to significantly predict TP53 mutation status (198). A study by Kather et al published in 2020 collected data on over 5000 tumour types found that with they could predict 9 oncogenic drivers (APC, KRAS, TP53, BRAF, PIK3CA, KMT2B, KMT2D, BRCA1 and RNF43) and significantly predict CMS classification using an H&E

tumour section (199). Analysis of a significantly larger number of GMS-graded tumours transcriptome may tease out differences in the transcriptomic factors driving histological differences in each GMS, as previous studies have reported links between certain phenotypic measures and gene expression.

After focussing on GMS histological subtypes, the full transcriptome was assessed relative to pSTAT3^{tyr705} protein groups. In the full cohort (n=85), there were no differentially expressed genes identified, and subsequently when the extremes of pSTAT3^{tyr705} high/low expression were selected there were still no significantly differentially expressed genes at padj <0.05. This could be due to the study being underpowered, and ultimately due to the heterogeneity of cancer a significantly larger cohort would be required to identify/validate any meaningful differentially expressed genes. The ratio of gene probes to samples may have also decreased the likelihood of identifying differential expression, therefore future work could include analysing the data using a more focused panel of CRC-associated genes to dilute out the lowly-expressed genes. For examples, gene profiles of pSTAT3^{tyr705} groups could be reanalysed after extracting the ~600 gene panel used to determine CMS. Principal component analysis highlighted some potential gene clustering between pSTAT3^{tyr705} groups, however more patients should be included in further analyses to confirm this. Overall data from transcriptomics and pSTAT3^{tyr705} groups in this cohort preliminarily suggests that post-translational modifications such as phosphorylation of STAT3 are of greater importance in driving cancer progression than the underlying transcriptome of phenotypes. This hypothesis is validated by data from a breast cancer study which found that the tumour transcriptome was not predicative of the tumour proteome, and in patients where there was concordance, there was a decrease in survival outcomes (200). Similarly, in CRC a study which assessed STAT3 and pSTAT3 at the protein and mRNA levels, there was no correlation between the measures, and only protein expression was significantly upregulated in tumour versus normal samples (201).

Data from this chapter has identified 5 differentially expressed genes between GMS1 and GMS0/2 and no differences in the underlying transcriptome of pSTAT3^{tyr705} high and low protein expressing patients. There are limitations to these data due to the limited number of patients included in the study and the low mean expression of genes, and therefore further research is needed to validate the findings. It appears that in the case of GMS and pSTAT3^{tyr705} protein status, post-translational modifications such as phosphorylation may be of more importance clinically. From prior chapters, STAT3 activation at the protein level associated with reduced CSS in GMS2 patients and inhibition represents a promising therapeutic target for CRC. This transcriptomics analysis has not identified any obvious pathways for dual targeting with IL6/JAK/STAT3 inhibition, however IL1β and FFAR2 could be further assessed for any therapeutic potential given the association with GMS.

Chapter 9: Inhibition of STAT3 activation in 2-D colorectal cancer cell lines and 3-D tumour models

9.1 Introduction

Small molecule inhibitors that block specific components of dysregulated cellular signalling pathways are increasingly used to treat inflammatory and autoimmune disorders. A promising therapeutic approach for CRC is repurposing of these inhibitors to target tumour progression.

Data from the previous chapters could lead us to hypothesise that patients with GMS2 stromally-dense tumours may benefit from blockade of IL6/JAK/STAT3 signal transduction. The aim of this chapter was to investigate the therapeutic potential of repurposed JAK family inhibitors; Ruxolitinib (JAK1/2) and Tofacitinib (JAK2/3). These small molecule inhibitors were FDA approved in 2012 and 2011, respectively, to treat myelofibrosis and ulcerative colitis/rheumatoid arthritis(202, 203), and therefore this project will determine if they could potentially be repurposed for treatment of colorectal cancer.

To determine the effect of JAK inhibition on cell viability of 7 CRC tumour cell lines, drug screening was performed and evaluated using WST-1 cell viability assays. An ELISA probing for $pSTAT3^{tyr705}$ was performed on cell lysates to ensure both inhibitors were on target for reducing STAT3 pathway activation. It is now widely accepted that 2D cell lines are relatively poor models of disease, therefore, to better recapitulate CRC in the clinic, patient-derived organoids (PDOs) were developed. This was particularly relevant for this study as the source of the ligand which activates the pathway (IL6) likely arises from components of the TME not present in 2D lines. The 3D tumour models were treated with Ruxolitinib and Tofacitinib and assessed for response via WST-assays and IF staining for makers of apoptosis (Caspase 8) and proliferation (Ki67). The original tumour block for each patient from which the organoids were derived were obtained from tissue archives and stained with H&E to determine histological subtype and ultimately map response to phenotype. Matched sections were also stained for Ki67 to determine phenotypic subtype and $pSTAT3^{tyr705}$ to assess constitutive levels of pathway activation within the tumour.

These experiments aimed to provide a step towards understanding the therapeutic potential of JAK inhibitors, for use in combination with standard of care chemotherapy. A secondary aim was to investigate if tumour phenotype could predict response to targeted inhibition of STAT3 signalling. It was hypothesised that inhibiting JAK2 would result in STAT3 blockade and cell death in CRC cell lines and additionally a reduction in proliferation and apoptosis in patients derived organoids.

9.2 Targeting IL6/JAK/STAT3 signalling in colorectal tumour cell lines

9.2.1 Inhibitors of JAK1/2 and JAK2/3

Two inhibitors of JAK/STAT3 signalling were selected for inclusion. Ruxolitinib is a potent inhibitor of JAK1 and JAK2 and was FDA approved in 2012 for the treatment of myelofibrosis (204). Tofacitinib inhibits JAK2 and JAK3 and gained FDA approval in 2011 for the treatment of inflammatory disorders including rheumatoid arthritis and ulcerative colitis. Repurposing of these drugs for the treatment of colorectal cancer in combination with standard of care chemotherapy, may represent a promising therapeutic option for a subset of patients. To confirm that inhibition of JAK1/2 and JAK2/3 impact STAT3 activation, an ELISA was performed on SW620 cell lysates treated with 0.01% DMSO vehicle control, 1mM Ruxolitinib (Abcam, Cambridge, UK), or 1mM Tofacitinib (Abcam, Cambridge, UK) for 72 hours to an n=3. Optical density was measured at 450nM using a TECAN plate reader (TECAN, AG, Switzerland). Data were analysed in Microsoft Excel (Microsoft, Washington, USA). A standard curve was plotted and a linear trendline equation utilised to determine the concentration of experimental samples in ng/ml. A box plot was constructed to visualise data, with 10 times standard deviation error bars (Figure 9.1). Treatment with Ruxolitinib (p=0.077) and Tofacitinib (p= 0.062) trended towards reduced $pSTAT3^{tyr705}$ protein compared to vehicle control. The mean concentration of $pSTAT3^{tyr705}$ in vehicle control samples was 13.0ng/mL (+/-0.036), compared to 7.69ng/mL (+/-0.086) in Ruxolitinib-treated samples and 11.12ng/mL (+/-0.023) in Tofacitinib-treated cells.

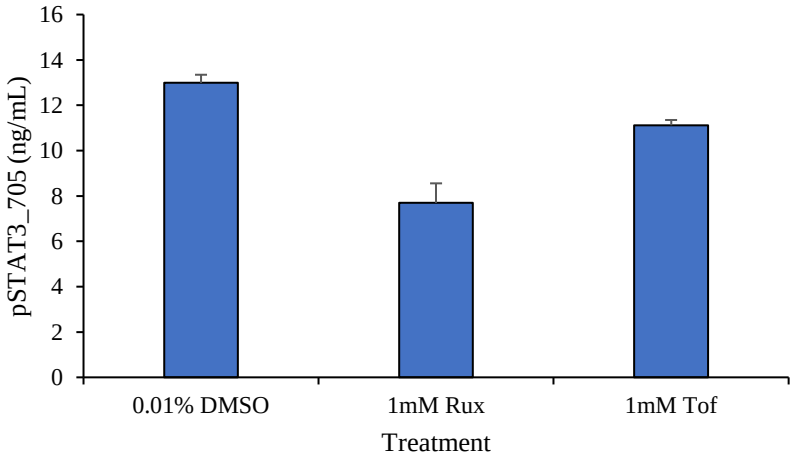


Figure 9.1 $pSTAT3^{tyr705}$ following JAK1/2/3 Inhibition in CRC lines. *Box plot showing cell levels of $pSTAT3^{tyr705}$ from cell lysates treated with 0.01% DMSO, 1mM Ruxolitinib or 1mM Tofacitinib over 72 hours with 10xSD error bars. Concentration of $pSTAT3^{tyr705}$ was measured via ELISA and units are ng/ml.*

Further testing for on target effects was performed by staining patient-derived normal and tumour explants for pSTAT3^{tyr705} via IHC. Representative images are shown in 9.2. Nuclear expression of pSTAT3^{tyr705} protein was weaker in both normal and tumour samples treated with Ruxolitlinib (C, F) and Tofacitinib (B, E) compared to vehicle control (A, D) (Figure 9.2).

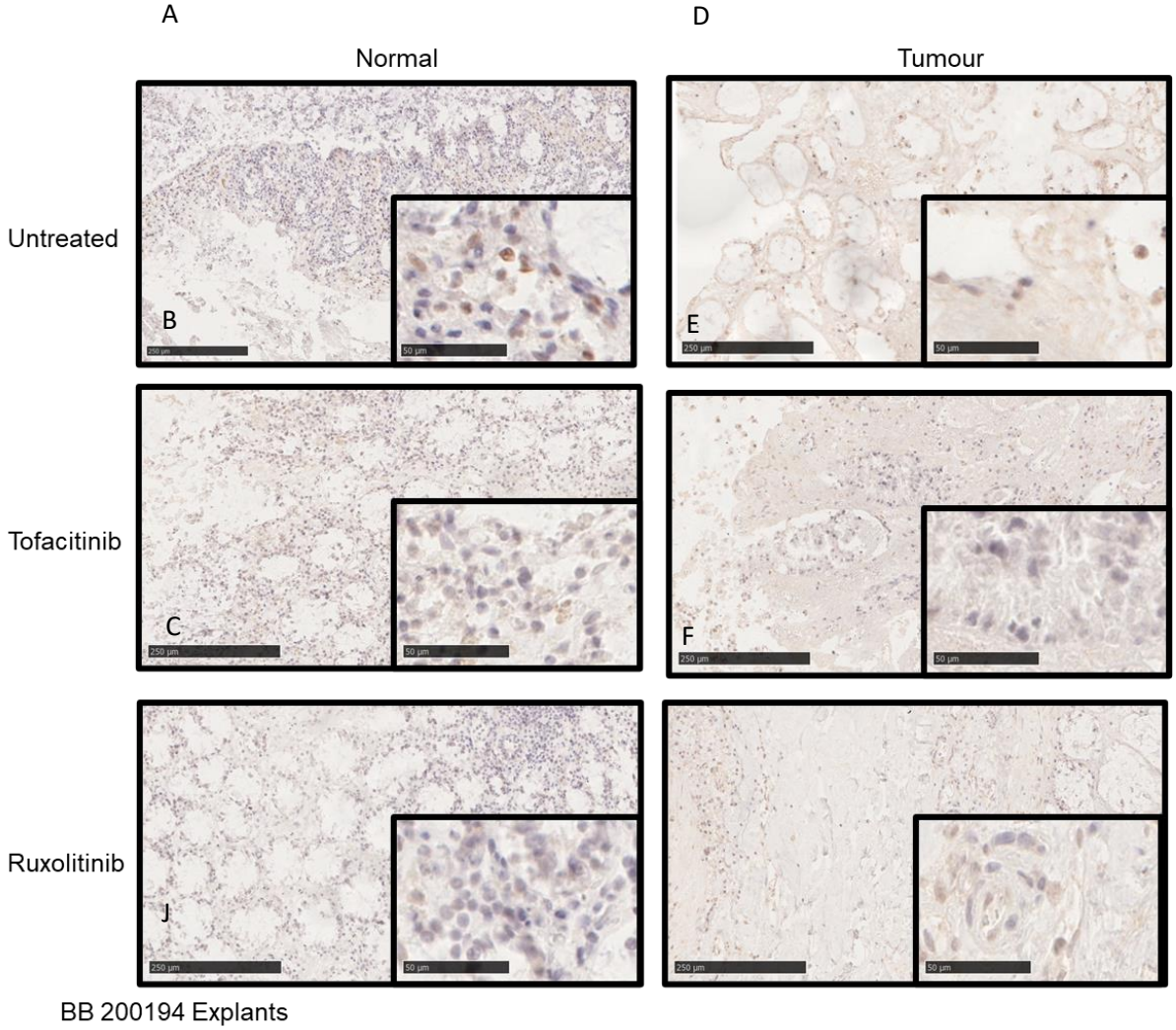


Figure 9.2 Representative images of tumour explants stained for pSTAT3^{tyr705} via IHC.

Explants were grown from surplus CRC adjacent normal and tumour resection tissue cultured for 48 hours in complete medium with 0.01% DMSO, 1mM Tofacitinib or 1mM Ruxolitinib. Explants were then fixed in formalin and embedded in paraffin blocks, cut and stained for pSTAT3^{tyr705} via IHC. Representative images of adjacent normal colon untreated (A), Tofacitinib treated (B) and Ruxolitinib treated (C) explants stained via IHC for pSTAT3^{tyr705}. Representative images of tumour colon untreated (D), Tofacitinib treated (E) and Ruxolitinib treated explants (F) stained via IHC for pSTAT3^{tyr705}.

9.2.2 Inhibiting JAK1/2 in CRC cell lines

CRC cell lines were available through the American Type Culture Collection (ATCC) or provided by other research groups within the University of Glasgow and the CRUK Beatson institute. Cells were grown according to standard aseptic culture conditions provided by ATCC. Mutational profiles, MSI status and CMS classification based on transcriptomic profiles of each cell line are shown in Table 9.1(11, 205). There were 3 cell lines characterised as CMS1: COLO205, DLD1 and SW48. COLO205s have APC, TP53 and BRAF mutations and are microsatellite stable. DLD-1s have microsatellite instability and mutations in APC, TP53, KRAS, BRCA2 and PIK3CA. SW48s are MSI and observe mutations in APC, BRCA2 and PIK3CA genes. T84 cell lines are categorised as CMS2, MSI and have mutations APC, TP53 and KRAS. HT29 cells are generally MSS, CMS3 and have mutations in APC, TP53 and BRAF. There were 3 cell lines available categorised as CMS4: SW620, SW480 and HCT116. SW620 and SW480 cells are MSS with mutations in APC, TP53 and KRAS. HCT116 cells are MSI with mutations in KRAS, BRCA2 and PIK3CA. Utilising these cell lines allowed for investigation of response to JAK inhibitors in a wide variety of genotypes and cellular phenotypes.

Cell line	CMS	Mutation profile	MSI Status
COLO205	1	APC, TP53, BRAF	MSS
DLD-1	1	APC, TP53, KRAS, BRCA2, PIK3CA	MSI
SW48	1	APC, BRAC2, PIK3CA	MSI
T84	2	APC, TP53, KRAS	MSS
HT29	3	APC, TP53, BRAF	MSS
SW620	4	APC, TP53, KRAS	MSS
SW480	4	APC, TP53, KRAS	MSS
HCT116	4	KRAS, BRCA2, PIK3CA	MSI

Table 9.1 Mutational profile and characteristics of CRC cell lines. *Table showing mutational, CMS and MSI status of CRC cell lines based on characterisation from previous studies.*

As a step towards understanding if inhibiting JAK1/2 or JAK2/3 could be therapeutically beneficial WST-1 assays were performed to determine the percentage of viable cells following treatment with Ruxolitinib and Tofacitinib relative to controls. Standard of care chemotherapeutic agent 5FU was used a positive control. Cells were seeded in triplicate and an n=3 for each condition in 96 well plates according to appropriate densities calculated during optimisation experiments. Treatments were applied 24 hours after seeding and WST-1 cell viability assays were performed at 48-, 72-, or 96-hours post-treatment. Optical density was measured at 450nM using a TECAN plate reader (TECAN, AG, Switzerland). Data were normalised to vehicle control in Microsoft Excel (Microsoft,

Washington, USA) and statistically analysed by paired T-tests comparing treatment group with vehicle control. Significance was set to $\alpha \leq 0.05$. Representative light microscopy images of a subset of lines; COLO205 (A), HT29 (B) and HCT116 cells (C) treated with media control, vehicle control and 1mM Ruxolitinib and 1mM Tofacitinib, are shown in Figure 9.3.

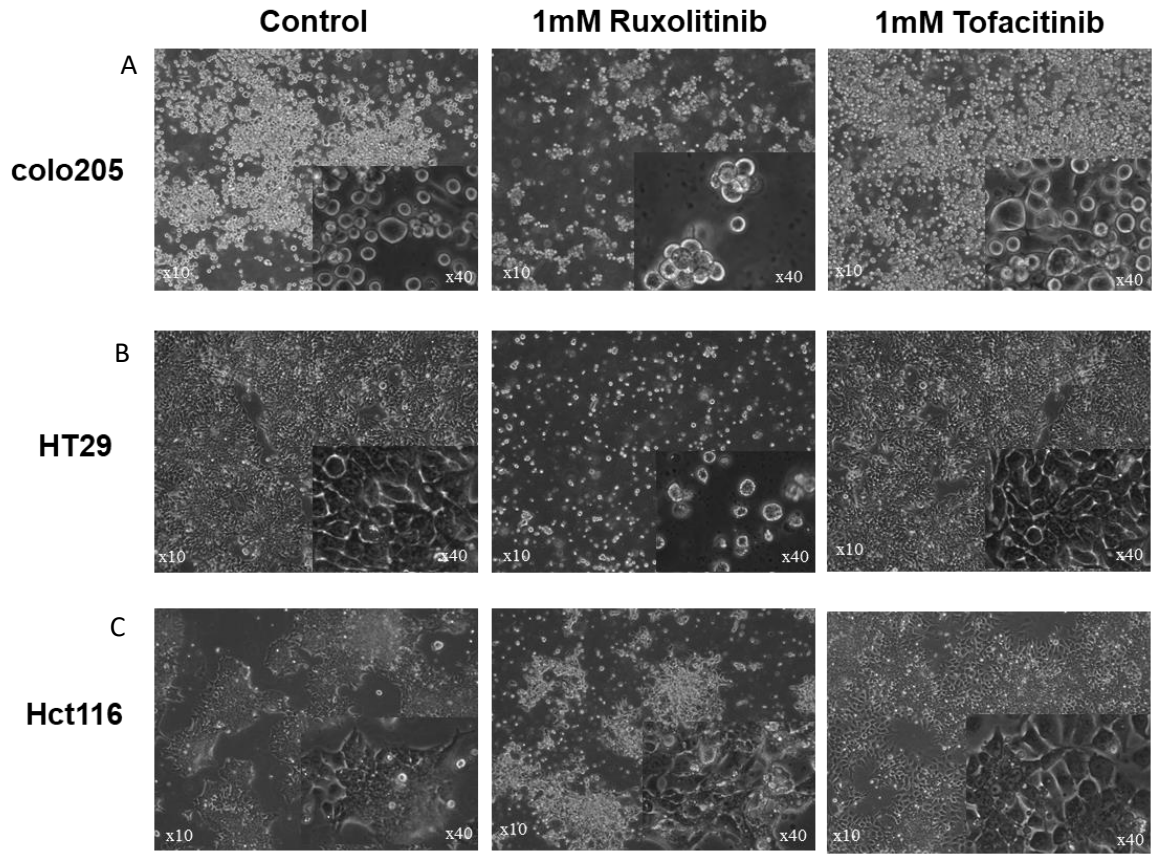


Figure 9.3 CRC Cell lines treated with JAK inhibitors. *Representative light microscopy images from a selection of CRC cell lines utilised in this study including COLO205 (A) , HT29 (B), and HCT116 (C) cells treated with 0.1% DMSO vehicle control, 1mM Ruxolitinib or 1mM Tofacitinib for 48 hours.*

Analysis of WST-1 assay data from cells treated with Ruxolitinib data was performed and a box plot was constructed to visualise the results (Figure 9.4). The percentage of viable DLD-1 cells was significantly reduced when treated with 1mM Ruxolitinib versus vehicle control at 48- ($p=0.006$), 72- ($p=0.005$) and 96 hours post-treatment ($p=0.023$) (Figure 9.4). There was no significant difference in DLD-1 cell viability between vehicle control and any other Ruxolitinib doses at any time point (Figure 9.4). The percentage of viable COLO205 cells was significantly reduced following 1mM

Ruxolitinib treatment compared to vehicle control at 48- (p=0.011) and 96 hours post treatment (p=0.034). HCT116 cells treated with 1mM Ruxolinitib showed significant reduction in viability compared to vehicle control at 96- hours post-treatment (p=0.009) (Figure 9.4). A significant reduction in HCT116 cell viability was also observed at 100μM (p=0.011) and 10μM (p=0.018) at 96 hours post treatment. There was a significant reduction in cell viability of HT29 cells when treated with Ruxolitinib at 1mM for 48 hours (p=0.037) and this was potentiated at 96 hours post-treatment (p=0.0009) (Figure 9.4). In terms of SW620 cells, treatment with 1mM Ruxolitinib significantly reduced cell viability compared to vehicle control 48 hours post-treatment (p=0.0007) (Figure 9.4). A significant difference between vehicle and drug treatment was also observed in SW620 lines for 10μM (p=0.05) and 1mM Ruxolitinib (p=0.0055) 96 hours after treatment (Figure 9.4). The vehicle control sample from the SW620s yielded low cell viability and this assay should be repeated in future work. There was a significant reduction in cell viability of T84 cells when treated with Ruxolitinib at 1mM (p=0.0052) and 100μM (p=0.0075) 96 hours post-treatment (Figure 9.4). SW480 cells treated with 1mM Ruxolitinib observed a significant reduction in percentage of viable cells compared to vehicle control at the 48-hour time point (p=0.009) (Figure 9.4). Finally, SW48 cells also showed significantly reduced viability at 48 hours post treatment with 1mM Ruxolitinib compared to vehicle control (p=0.037) (Figure 9.4). Collectively these data suggest that Ruxolitinib is effective at killing most CRC tumour cell lines, but at a dose not clinically relevant (1mM). The only cell lines to respond to Ruxolitinib at a lower, more promising concentration wereCMS4 lines SW620 and HCT116, which supports our hypothesis that inhibiting IL6/JAK/STAT3 signalling would be most beneficial in patients with stromal-rich tumours. Data from the 2 cell lines (HCT116, SW620) which showed the most promising response following treatment with Ruxolitinib were extracted and plotted on another bar chart (Figure 9.5).

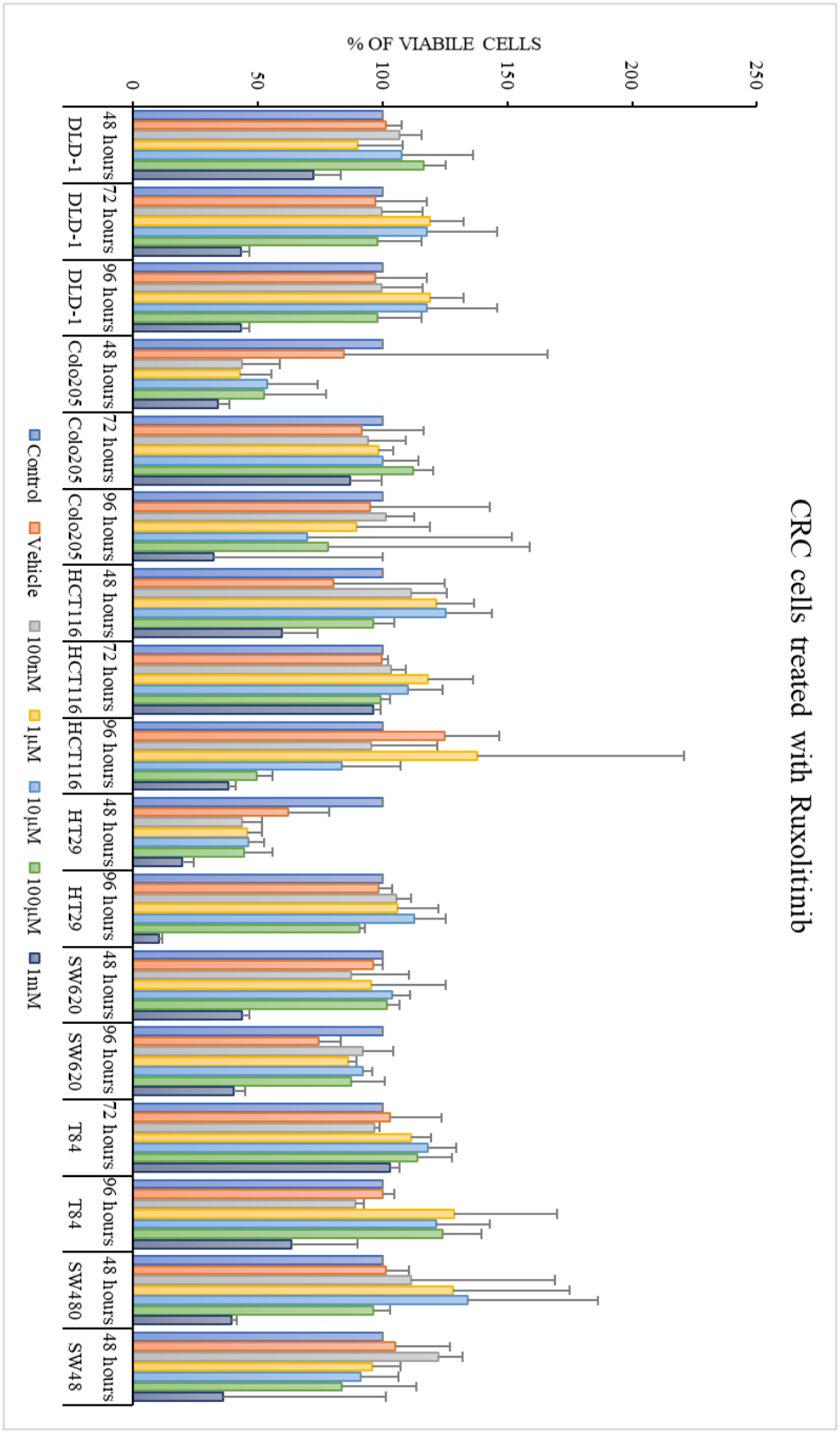


Figure 9.4 Cell Viability following JAK1/2 Inhibition in CRC lines. *Box plot showing cell viability normalised to media control of CRC cell lines treated with 0.01% DMSO, 100nM, 1µM, 10µM, 100µM and 1mM Ruxolitinib over at least two time points from 48-, 72- and 96- hours.*

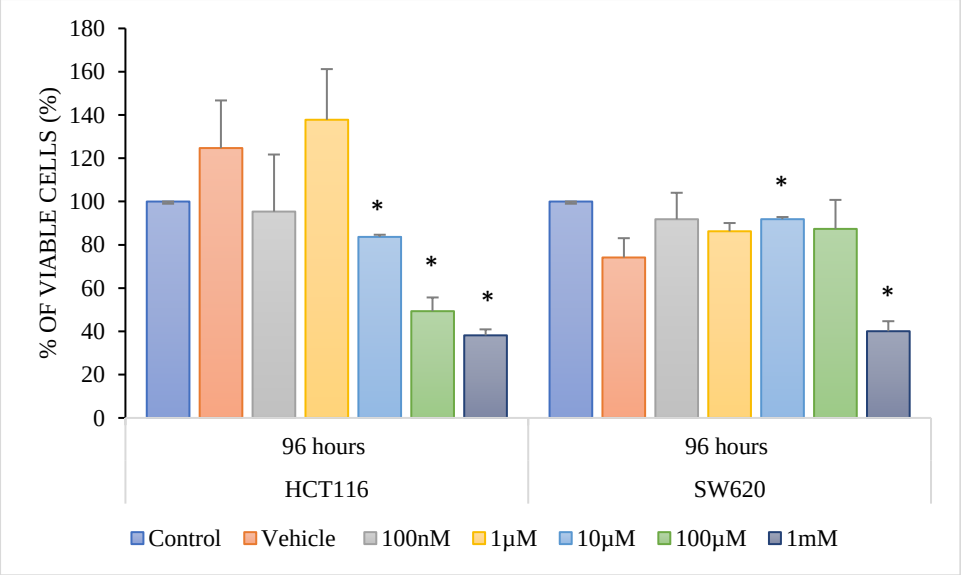


Figure 9.5 Cell Viability following JAK1/2 Inhibition in CMS4 CRC lines which showed a response. *Box plot showing cell viability normalised to media control of CRC cell lines which showed a response to treatment with Ruxolitinib at 96- hours post-treatment. Paired t-tests were utilised to determine any significant differences between vehicle control and treated samples with $p<0.05$ set as the threshold for significance.*

Next, data from WST-1 assays of CRC cell lines treated with JAK2/3 inhibitor Tofacitinib were analysed via paired T-tests and visualised in a bar chart (Figure 9.6). DLD-1s observed a significant reduction in cell viability following Tofacitinib treatment at 96 hours at 1mM (p=0.024) and 1µM (p=0.04). There was no significant effect on COLO205 cell viability following Tofacitinib treatments. In HCT116 cells, 48 hours post Tofacitinib treatment no significant effects on cell viability were observed. However, at 72 hours post treatment HCT116 cells treated with 1mM Tofacitinib had a significant reduction in viability compared to vehicle control (p=0.04). This was potentiated at 96 hours when Tofacitinib was successful at reducing the percentage of viable cells at 1µM (p=0.0096), 100µM (p=0.011) and 10µM (p=0.018) compared to vehicle control. HT29 cells treated with 1mM Tofacitinib had significant reduction in percentage of viable cells at 48 hours (p<0.0001) and 96 hours (p=0.02) post treatment compared to vehicle control. There were no significant effects of Tofacitinib treatment on percentage of viable SW620 cells at any time point. Cell viability of T84 cells was significantly reduced compared to vehicle control treated cells at 96 hours but only at 100nM Tofacitinib (p=0.03). Interestingly, the viability of SW480 cells was significantly reduced compared to vehicle control at 48 hours post treatment with 1mM (p=0.014), 100µM (p=0.041), 1µM (p=0.023) and 100nM (p=0.001) Tofacitinib. There was no significant difference in percentage of viable SW48 cells between any dose of Tofacitinib and vehicle control at 48 hours. Overall, akin to the observations with JAK1/2 inhibition, CMS4 cell lines (SW480 and HCT116) showed the most

promising response to treatment with Tofacitinib. HCT116 and DLD-1 cell viability data at 96 hours post Tofacitinib treatment, and SW480 at 48 hours were plotted on bar chart to enable easier visualisation of the significant responses to drug treatment (Figure 9.7).

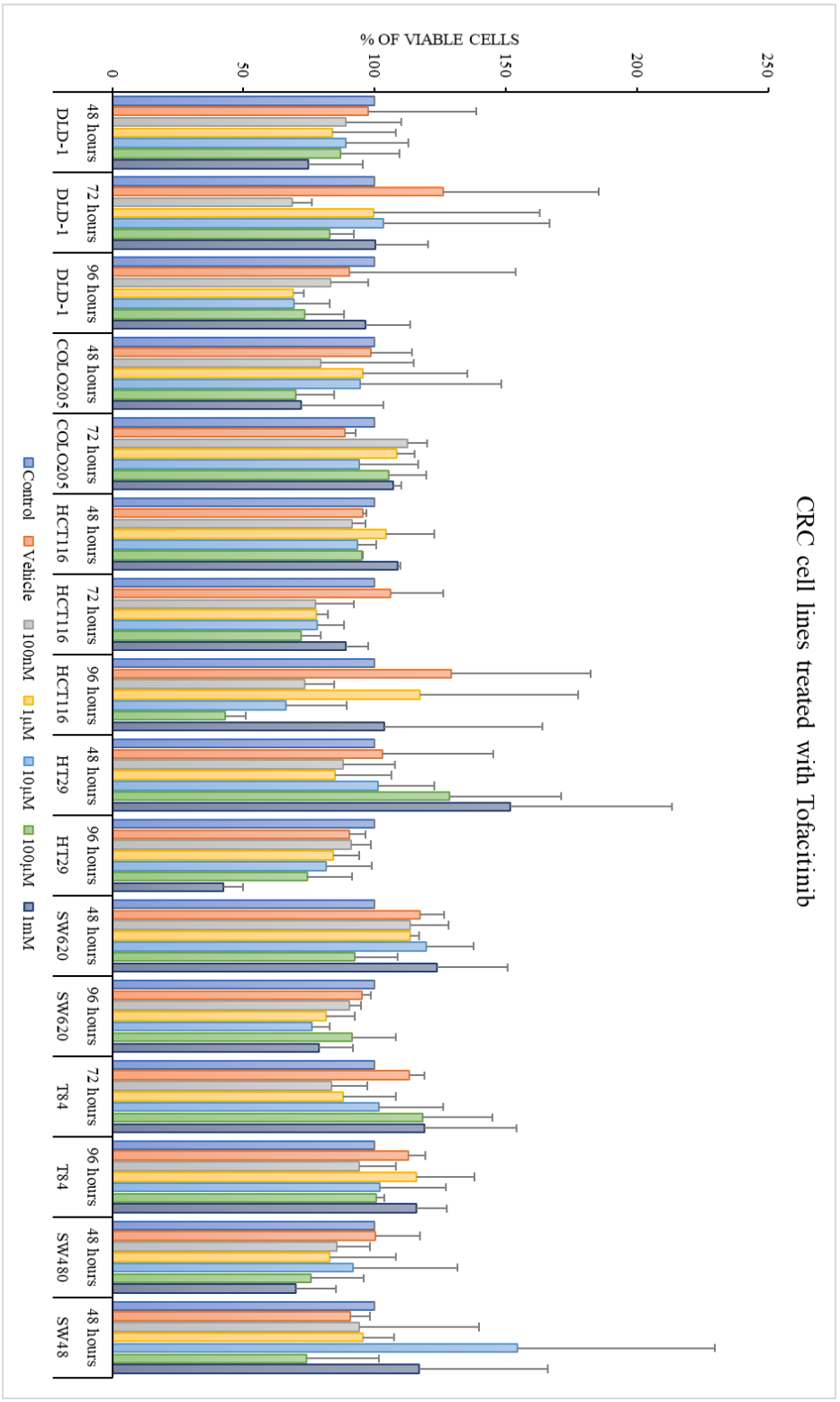


Figure 9.6 Cell Viability following JAK2/3 Inhibition in CRC lines. *Box plot showing cell viability normalised to media control of CRC cell lines treated with 0.01% DMSO, 100nM, 1µM, 10µM, 100µM and 1mM Tofacitinib over at least two time points from 48-, 72- and 96- hours.*

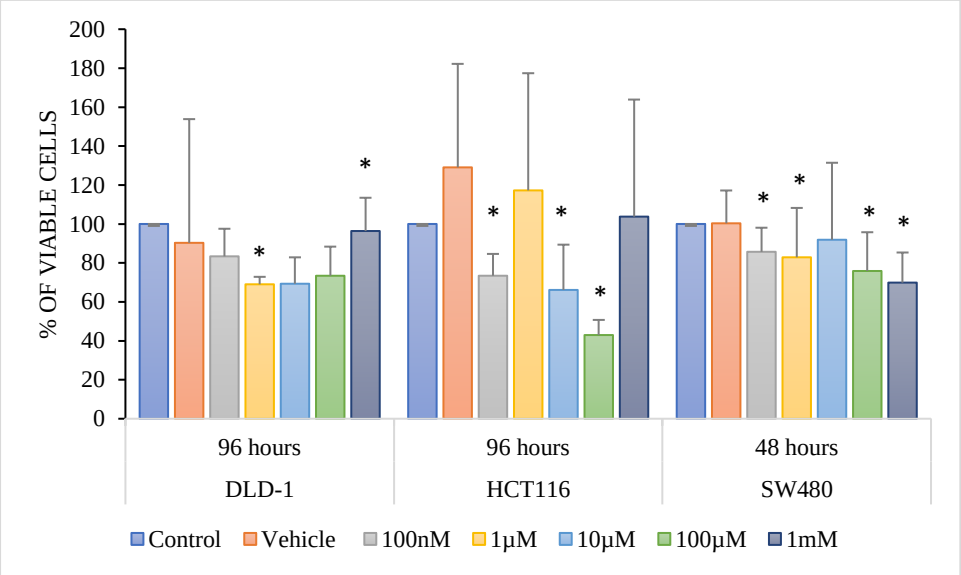


Figure 9.7 Cell Viability following JAK1/2 Inhibition in CMS4 CRC lines which showed a response. *Box plot showing cell viability normalised to media control of CRC cell lines which showed a response to treatment with Tofacitinib at 96- hours/48 hours post-treatment. Paired t-tests were utilised to determine any significant differences between vehicle control and treated samples with $p<0.05$ set as the threshold for significance.*

9.3 Targeting IL6/JAK/STAT3 signalling in patient-derived tumour organoids

9.3.1 Derivation and characterisation of patient-derived organoids

Due to the importance of the tumour microenvironment being increasingly recognised, it is now accepted within the field of cancer research that 2D cell lines are not the gold standard for modelling patient disease, and there is a move towards 3D tumour models which better recapitulate CRC. Patient-derived organoids (PDOs) are spheres of mixed populations of cells derived from tumour resections and/or biopsies. The 3D nature of PDOs and heterogeneity results in a more representative model of patient disease than 2D cell lines. Therefore, PDOs were utilised in this project to build on the 2D cell line experiments and test the efficacy of JAK inhibitors in a more robust, clinically relevant way.

For the purposes of this project a method to grow CRC PDOs directly from Greater Glasgow and Clyde surplus patient tissue was developed and optimised. Collaboration with Dr Chao Wu and Mr

Josh Smith at Memorial Sloan Kettering, New York and Dr Kevin Myant at the University of Edinburgh enabled translation of protocols to the University of Glasgow. This resulted in the establishment of 7 organoid lines. Additionally, there were 5 PDO lines available from the Sanger centre array, provided by Professor Sansoms laboratory at the Cancer Research UK Beatson Institute.

Given the hypothesis that patient with GMS2 tumours is the group with the most potential for therapeutic inhibition of IL6/JAK/STAT3 signalling, the original donor tissue resection blocks were requested and pulled by the Glasgow Tissue Research Facility. An H&E section was utilised to determine GMS of each of the Sanger and Glasgow lines. The tumour resection block from the cases Sanger organoids were derived from were requested from NHS GGC and provided by Glasgow Tissue Research Facility. H&E staining was performed by Glasgow Tissue Research Facility. Ki67 IHC staining was performed by KP. Resections were scored for KM grade, TSP and Ki67 proliferation index. There were 2 GMS0, 2 GMS1 and 1 GMS2 patient in the Sanger cases (Table 9.2). In terms of phenotypic subtype, there were 2 immune cases, 1 canonical, 1 latent and 1 stromal patient (Table 9.2). Sections were also stained for $pSTAT3^{tyr705}$ via IHC with representative images shown in figure 9.8. Sanger31, 37 and 25 showed some positive staining for $pSTAT3^{tyr705}$, whereas Sanger13 and 41 showed no positive staining for $pSTAT3^{tyr705}$.

Sanger ID	GMS	Phenotypic Subtype
13	1	Latent
25	1	Canonical
31	2	Stromal
37	0	Immune
41	0	Immune

Table 9.2 Histological subtypes of Sanger organoid patients. *Table showing GMS classification and phenotypic subtypes assigned to Sanger case organoids as determined by histological scoring of a matched full tumour section from the same patient.*

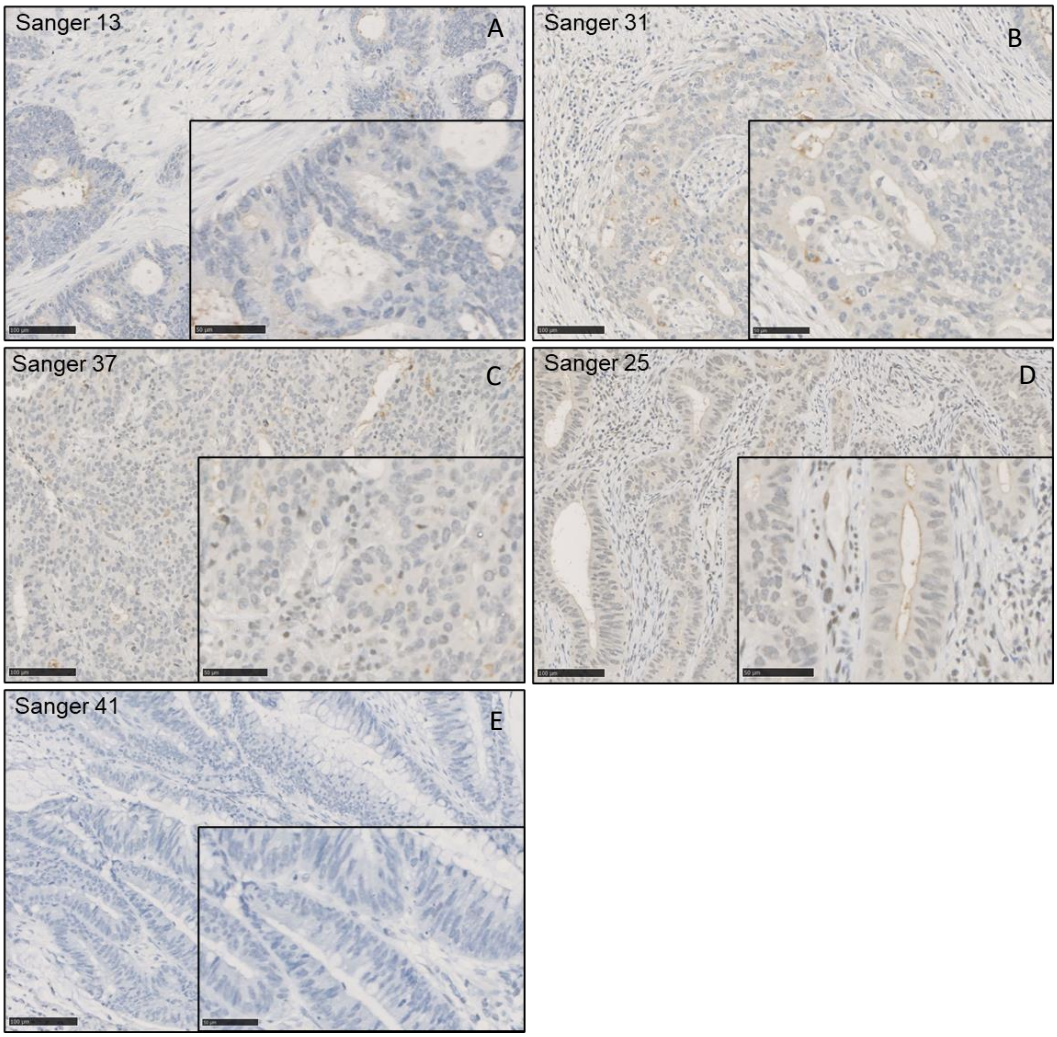


Figure 9.8 IHC staining of Sanger cases for pSTAT3^{tyr705}. *Representative images of IHC staining for pSTAT3^{tyr705} of Sanger case tumour resections, with Sanger 13 (A), Sanger 31 (B), Sanger 37 (C), Sanger 25 (D) and Sanger 41 (E).*

The tumour resection blocks for patients from which Glasgow organoids were derived were pulled by Glasgow tissue research facility and stained with H&E and for Ki67 via IHC. GMS and phenotypic subtyping were determined with data shown in Table 9.3. There were 2 patients classed as GMS0, 4 patients as GMS1 and 1 patient was classed as GMS2. In terms of phenotypic subtype, there were 2 immune patients, 2 canonical, 2 latent and 1 patient assigned to the stromal subtype. Tumour resections were also stained for pSTAT3^{tyr705} via IHC and representative images for each patient are shown in figure 9.9. The highest expression of pSTAT3^{tyr705} was observed in patient BB200040. Positive staining was observed in the other cases except for BB200534.

BB ID	GMS	Phenotypic Subtype
BB200534	1	Canonical
BB200154	1	Canonical
BB210022	0	Immune
BB190614	1	Latent
BB200099	2	Stromal
BB191240	0	Immune
BB200040	1	Latent

Table 9.3 Histological subtypes of Glasgow organoid patients. *Table showing GMS classification and phenotypic subtypes assigned to Glasgow case organoids. Histological scoring was performed on full face sections from the patients from which the organoids were derived.*

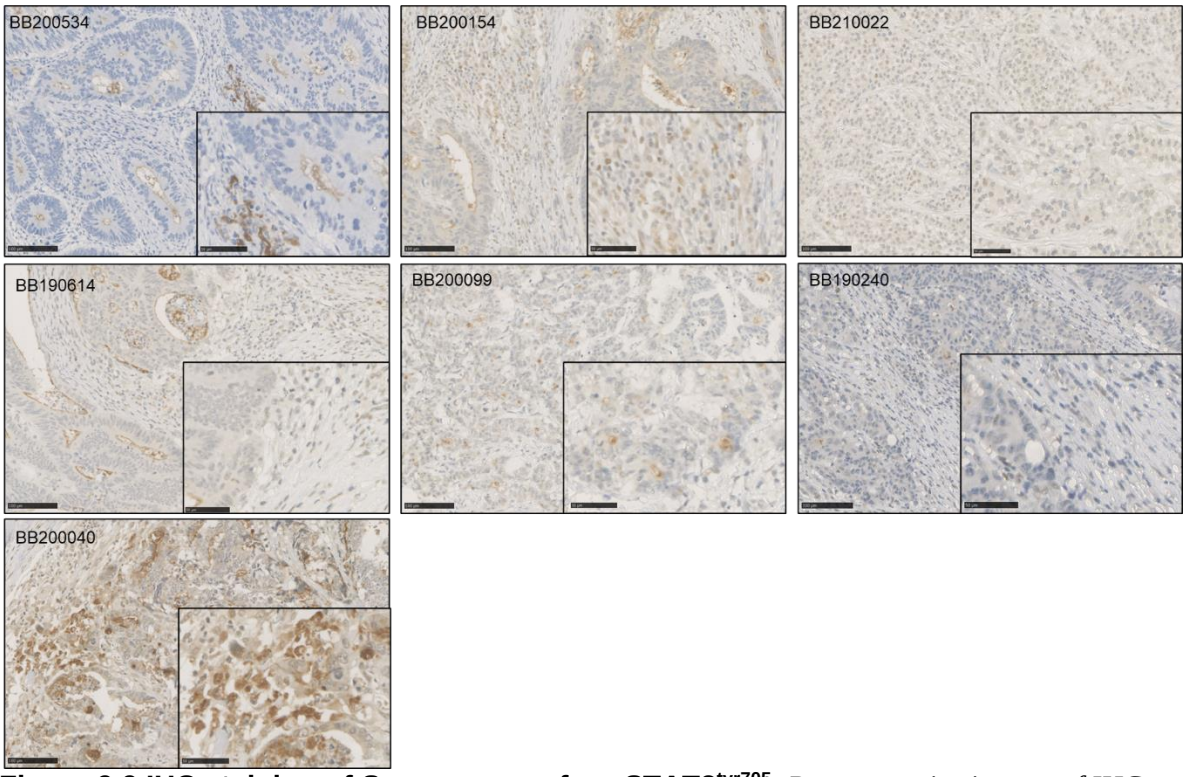


Figure 9.9 IHC staining of Sanger cases for pSTAT3^{tyr705}. *Representative images of IHC staining of original Glasgow cases tumour resections for pSTAT3^{tyr705}.*

CMS classification and mutational profiles were available for Sanger samples, as outlined in table 9.4. Patients 13 and 37 were CMS2, 25 and 41 were CMS3 and 31 was CMS4 (Table 9.4) giving a range

of molecular subtypes. All patients observed mutation in APC and all patients except Sanger41 also had mutations in the TP53 gene (Table 9.4). For the Glasgow patient-derived samples transcriptomics and mutational profiles were not yet available.

Sanger ID	CMS	Driver mutations
13	2	APC, TP53, KRAS, FBXW7, PCBP1, CHD4, LRP1B
25	3	APC, TP53, KRAS
31	4	APC, TP53, BCL9L
37	2	APC, TP53, FBXW7, CHD4, MAPK1, CCDC169, SOHLH2
41	3	APC, SOX9, SMAD2, MAP2K1, KMT2B, STRN, PIK3CB

Table 9.4 Sanger organoid mutational profiles and CMS classifications. *Table outlining the CMS classifications assigned to each case of the Sanger organoids included in the stud as previously performed by the Sanger centre.*

9.3.2 Inhibiting JAK proteins in CRC tumouroids

Prior to drug screening assays, organoids were analysed for the presence of a mixed population of cell types. IF staining for epithelial cell adhesion molecule-1 (EpCAM1) was utilised to confirm the presence of an epithelial tumour component and staining for α-smooth muscle actin (α-SMA) was utilised to confirm a stromal component to the PDOs.

Subsequently, drug screening of JAK inhibitors Ruxolitinib and Tofacitinib was performed in a similar manner to the WST-1 assays in 2D CRC cell lines. PDOs were plated in 5uL domes in 96 well plates, with treatment conditions set up in duplicate or triplicate based on availability of cells. Images were taken at 72-hours post-treatment with a Zeiss light microscope at x20 magnification. WST-1 reagent was added 72 hours post treatment and optical density was measured at 450nM using a TECAN Infinite PRO plate reader (TECAN, AG, Switzerland). Absorbance data for each treatment were averaged and normalised to vehicle control. Paired T-tests were performed in Microsoft Excel (Microsoft, Washington, USA) to compare each treatment to vehicle control. Statistical significance was set to α=0.05.

Representative images of Sanger 31 organoids stained via IF for EpCam1 and α-SMA are shown in Figure 9.10 (A) indicating the presence of tumour epithelial cells and stromal cells. This provided evidence that the Sanger 31 PDOs were good models of disease and therefore drug screening of JAK inhibitors was proceeded with. Light microscope representative images taken 72 hours post-treatment shows healthy Sanger31 organoids in the vehicle control sample (Figure 9.10 (B)). PDOs treated with 10μM Ruxolitinib or 10μM 5FU show necrotic cores and dead cells (Figure 9.6 (B)). The morphology

of organoids treated with 1mM Tofacitinib was similar to that of the control samples (Figure 9.6 (B)). Immediately after light microscope images were taken, WST-1 assays were performed to assess the percentage of viable cells in each condition. A bar chart was plotted in Microsoft Excel to visualise absorbance data from WST-1 data (Figure 9.10 (C)). There was a significant reduction in the percentage of viable cells following treatment with 10µM Ruxolitinib ($p=0.015$) and 10µM 5FU ($p=0.003$) (Figure 9.10 (C)). No significant difference in the percentage of viable cells was observed between vehicle control and cells treated with 10µM Tofacitinib. IF staining was utilised to investigate the effects of Ruxolitinib on apoptosis and proliferation of PDOs. The percentage nuclei positive for Ki67 was reduced following treatment with 1mM Ruxolitinib compared to samples treated with 0.01% DMSO, as seen in representative images (Figure 9.10 (D)). Ruxolitinib-treated samples also showed a profound increase in expression of mid-phase apoptosis marker caspase-8 compared to vehicle control as shown in representative images (Figure 9.10 (D)). These data were interesting as Sanger 31 was graded GMS2, stromal phenotypic subtype and CMS4, which supported the hypothesis that stromal-rich tumour would respond to JAK inhibition.

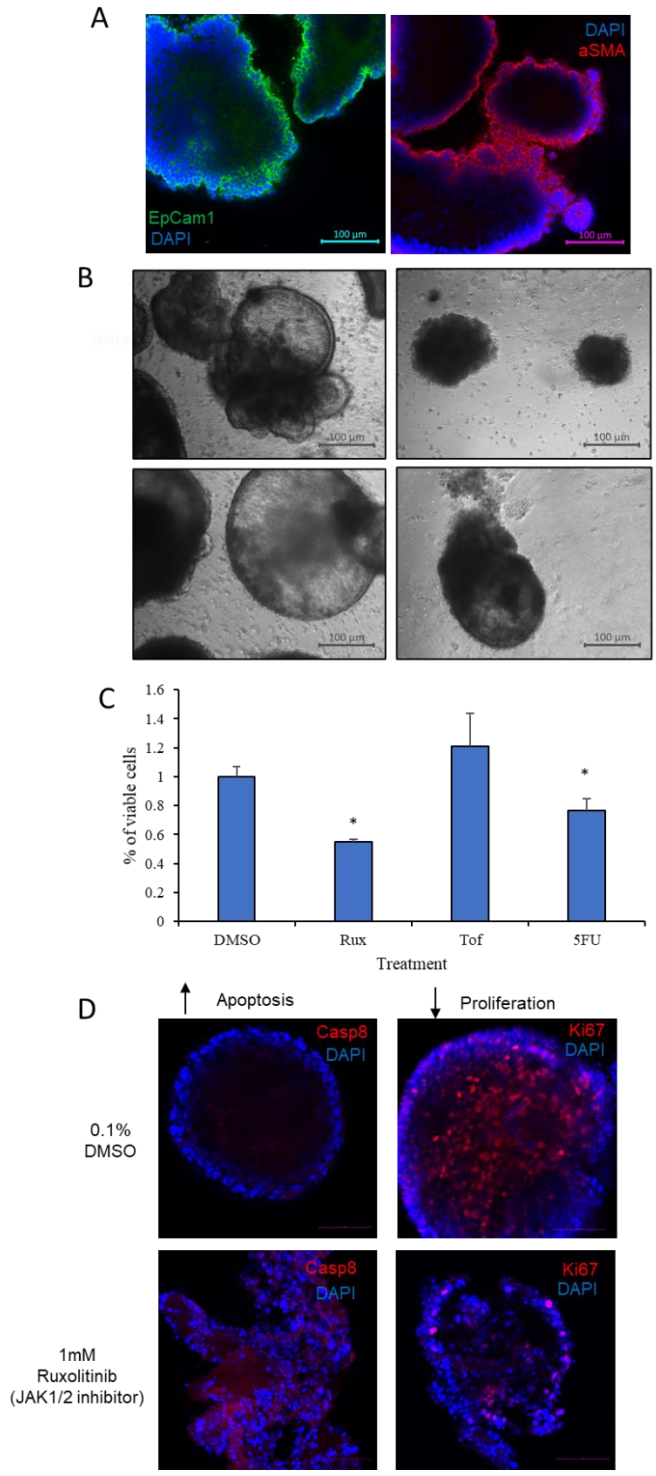


Figure 9.10 Drug screening of JAK inhibitors in Sanger 31. *Representative images showing IF staining for characterisation of tumour cells (EpCam1) and stromal cells (α-SMA) within Sanger31 PDOs (A). Representative light microscopy images showing Sanger13 organoids 72 hours post media control, 0.01% DMSO, 1mM Ruxolitinib, 1mM Tofacitinib and 5FU treatment (B). Box plot showing % of viable cells normalised to media control of CRC cell lines treated with 0.01% DMSO (C). Representative images of IF staining for Ki67 proliferation marker and Caspase8 apoptosis marker in Sanger31 PDOs treated with vehicle control or 1mM Ruxolitinib (D).*

Next, Sanger 37 PDOs were utilised for investigation into therapeutic potential of JAK inhibition in CRC. Representative images of tumour and stromal cell characterisation of Sanger37 organoids are shown in Figure 9.11 (A). This IF staining confirmed the presence of a mixed population of cells and therefore drug screening and WST-1 cell viability assays were subsequently performed. A bar chart was plotted in Microsoft Excel (Microsoft, Washington, USA) to visualise the effects of each treatment on the percentage of viable cells (Figure 9.11 (B)). There was a significant reduction in cell viability following addition of 1mM Ruxolitinib to cultures at 72 hours post treatment ($p=0.049$). There was no significant difference between vehicle control sample viability compared to Tofacitinib or 5FU-treated specimens, however there was a trend towards reduced viability in the 5FU-treated PDOs ($p=0.08$) (Figure 9.11 (B)). To investigate the effects of Ruxolitinib treatment further, PDOs were stained for Ki67 to assess proliferation and Caspase 8 to look for signs of apoptosis. Vehicle control samples were negative for caspase 8 and high for nuclear Ki67 expression (Figure 9.11 (C)). In contrast, PDOs treated with 1mM Ruxolitinib had high expression of caspase-8 and lower expression of nuclear Ki67 (Figure 9.11 (D)). Sanger 37 was classified as GMS0, Immune, so it would have been hypothesised that this patient would likely not benefit most from treatment with Jakinibs.

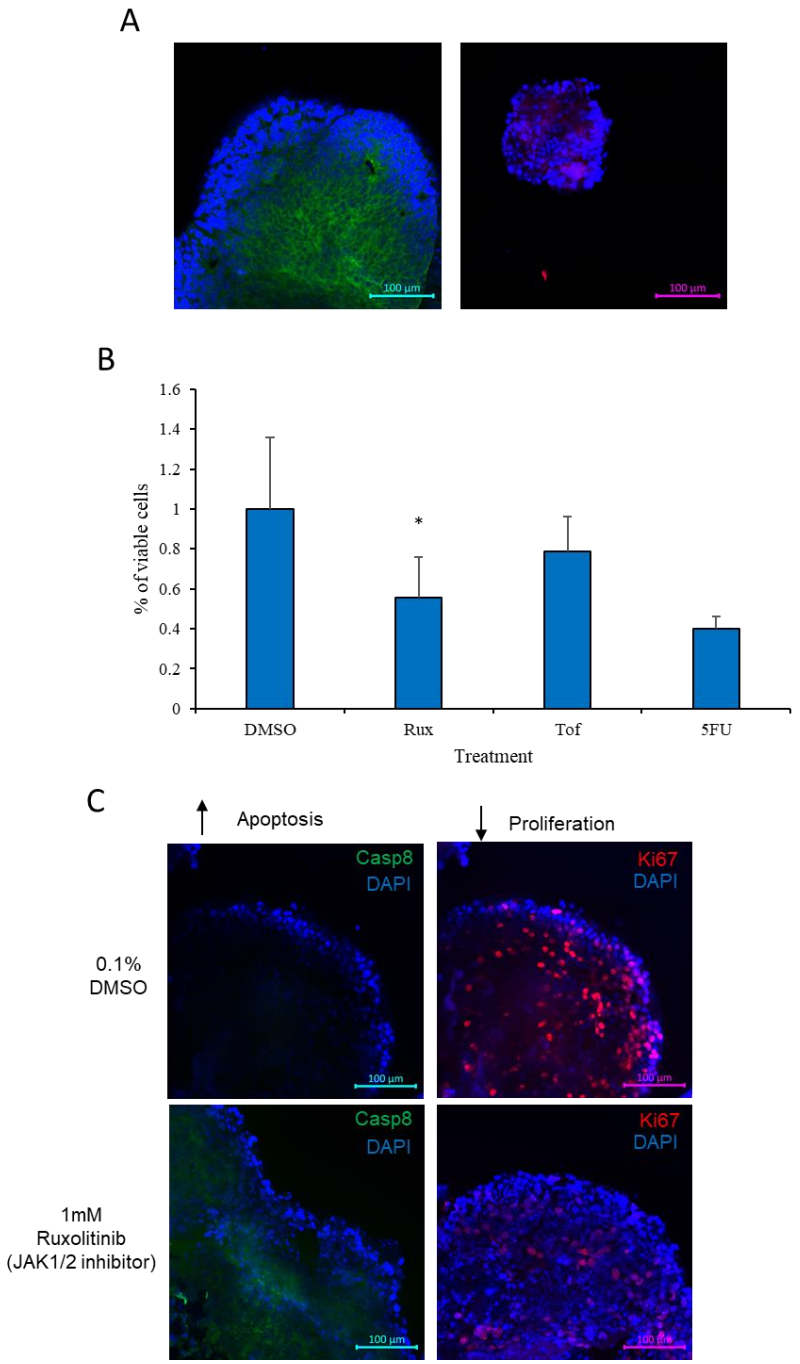


Figure 9.11 Drug screening of JAK inhibitors in Sanger 37. *Representative images of IF-stained Sanger37 PDOs for EpCam1 and α-SMA (A). Bar chart showing cell viability following treatment with 0.01% DMSO, 1mM Ruxolitinib, 1mM Tofacitinib and 1mM 5FU (B). Representative images of IF staining for Ki67 proliferation marker and apoptosis marker Caspase 8 in Sanger 37 PDOs treated with vehicle control or Ruxolitinib.*

Sanger 25 PDOs highly expressed both EpCam1 and α-SMA when stained via IF (Figure 9.12 (A)). A bar chart was plotted to visualise data from cell viability assays investigating the effects of inhibiting JAK1/2 and JAK2/3 (Figure 9.12 (B)). Paired t-tests were performed to analyse data from the WST-1

assay, which indicated a significant reduction in the percentage of viable cells following treatment with 1mM Ruxolitinib (p=0.034) compared to vehicle control (Figure 9.12 (B)). There was no significant effect on cell viability following Tofacitinib or 5FU treatment, however there was a trend towards reduced viability in the 5FU PDOs (p=0.095). To investigate the effects of Ruxolitinib in more detail, treated PDOs were stained via IF for markers of proliferation (Ki67) and apoptosis (caspase 8). PDOs treated with Ruxolitinib showed increased expression of caspase 8 compared to vehicle control (Figure 9.12 (C)). There was no tangible difference in expression of proliferation marker Ki67 between PDOs treated with vehicle control and Ruxolitinib, with both moderate nuclear positively observed for both conditions, suggesting inhibiting JAK1/2 in this PDO line did not affect cell proliferation (Figure 9.12 (C)). When graded histologically Sanger 25 was assigned to the GMS1 and canonical phenotypic subtype groups, and therefore a strong response to JAK inhibition would not have been predicted as the tumour was not deeply infiltrated with stroma.

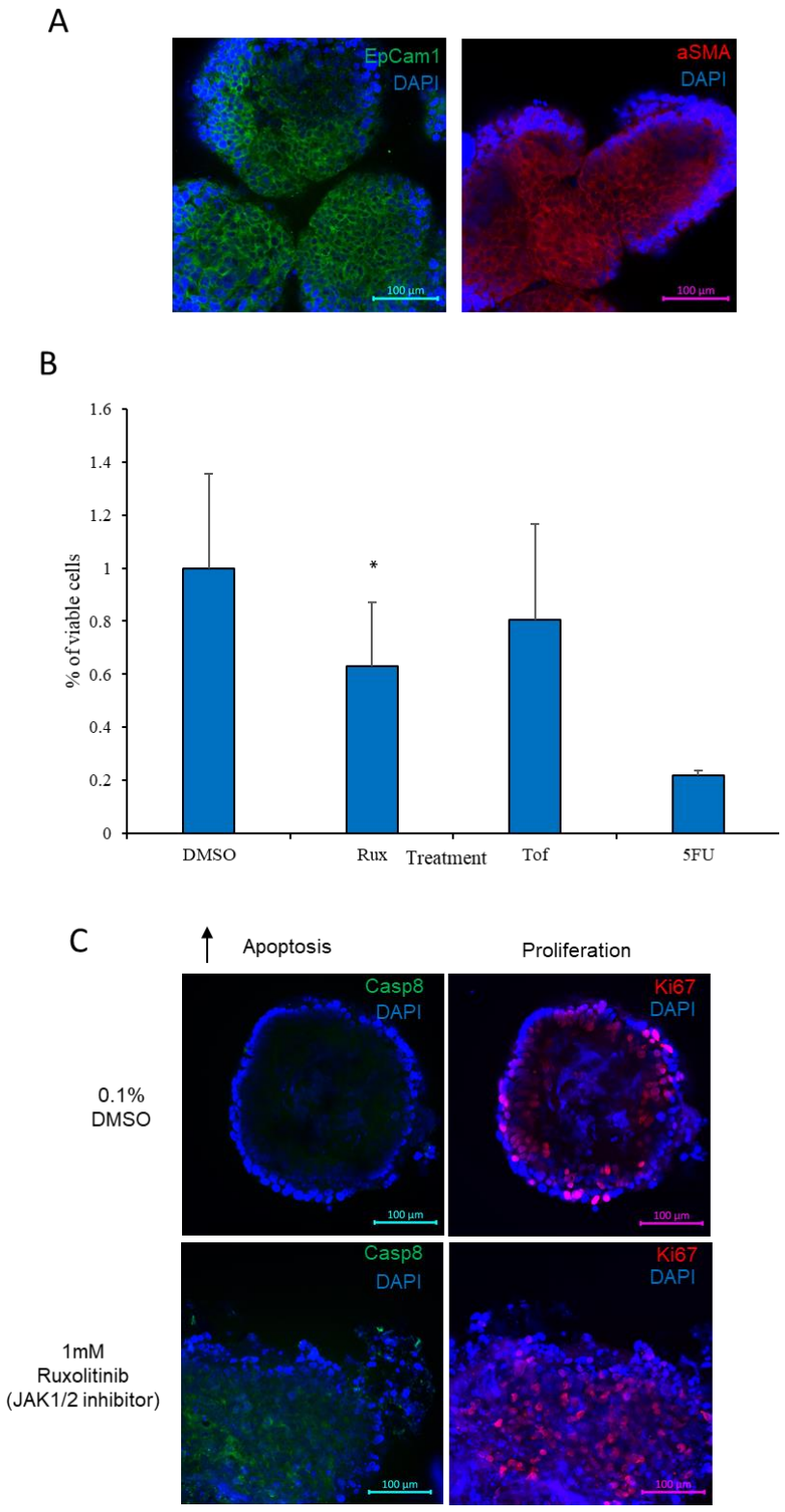


Figure 9.12 Drug screening of JAK inhibitors in Sanger 25. *Representative images of IF-stained Sanger25 PDOs for EpCam1 and α-SMA (A). Bar chart showing cell viability following treatment with 0.01% DMSO, 1mM Ruxolitinib, 1mM Tofacitinib and 1mM 5FU (B). Representative images of IF staining for Ki67 proliferation marker and apoptosis marker Caspase 8 in Sanger 25 PDOs treated with vehicle control or Ruxolitinib.*

Next Sanger 13 was assessed via IF for expression of EpCam1 and α-SMA. PDOs showed moderate expression of both markers confirming the presence of a mixed population of tumour and stromal cells (Figure 9.13 (A)). To assess the percentage of viable cells following treatment of PDOs with 1mM Ruxolitinib, 1mM Tofacitinib and 1mM 5FU a WST-1 assay was performed to n=3. A bar chart was plotted to visualise the percentage of viable cells in each treatment group (Figure 9.13 (B)). There was no significant difference in cell viability of PDOs treated with Ruxolitinib, Tofacitinib or 5FU compared to 0.01% DMSO vehicle control (Figure 9.13, B).

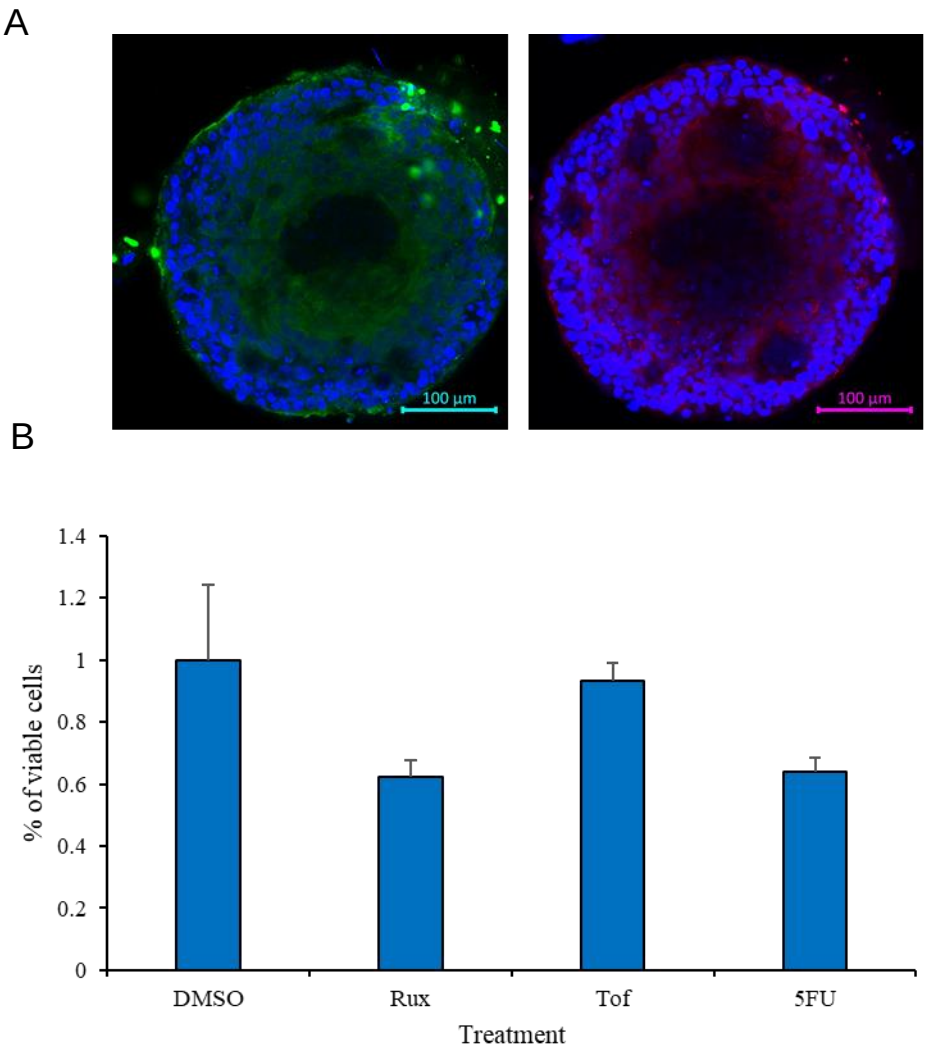


Figure 9.13 Drug screening of JAK inhibitors in Sanger 13. *Representative images of IF-stained Sanger13 PDOs for EpCam1 and α-SMA (A). Bar chart showing cell viability of Sanger 13s following treatment with 0.01% DMSO, 1mM Ruxolitinib, 1mM Tofacitinib and 1mM 5FU (B).*

Sanger 41 PDOs were stained via IF for EpCam1 and α-SMA to investigate the cellular populations present. Representative images of tumour (EpCam1) and stromal (α-SMA) cells are shown in Figure 9.14 (A). Compared to other Sanger cases, expression of both markers was weaker, however some

positive staining was observed. A WST-1 cell viability assay was performed to assess response to 1mM Ruxolitinib, 1mM Tofacitinib and 1mM 5FU treatments. A bar chart was plotted to visualise the percentage of viable cells in each treatment group (Figure 9.14 (B)). Paired T tests were performed to compare vehicle control viability to treatment groups and no statistically significant differences were observed at α=0.05.

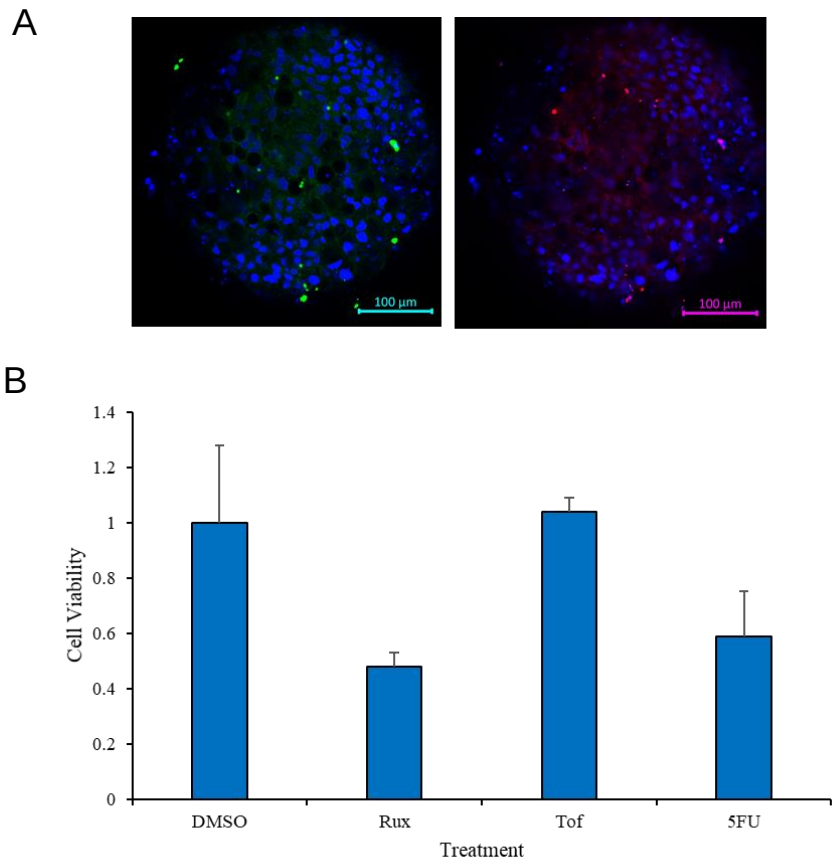


Figure 9.14 Drug screening of JAK inhibitors in Sanger 41. *Representative images of IF-stained Sanger41 PDOs for EpCam1 and α-SMA (A). Bar chart showing cell viability of Sanger41s following treatment with 0.01% DMSO, 1mM Ruxolitinib, 1mM Tofacitinib and 1mM 5FU (B).*

Following experiments conducted on of Sanger organoid lines, PDOs derived from Glasgow patient surplus tissue were assessed. Staining for EpCam1 and α-SMA via IF confirmed organoids derived from patient resection tissue BB200040 consisted of a mixed population of tumour and stromal cells. Representative images are shown in Figure 9.15 (A). Subsequently cell viability was assessed via WST-1 assays following treatment with 10µM Ruxolitinib, 10µM Tofacitinib and 10µM 5FU. Light microscopy images were taken 72 hours post treatment and representative images are shown in Figure 9.15 (B). PDOs treated with complete media and 0.01% DMSO vehicle control observed a healthy morphology at 72 hours. Ruxolitinib and 5FU treated samples appeared dead and necrotic with dark centres (Figure 9.15 (B)). The morphology of Tofacitinib treated samples was similar to that of the

media control (Figure 9.15 (B)). Interestingly, at n=3, there was no difference in cell viability following WST-1 assay between media control and Ruxolitinib, Tofacitinib or 5FU treated PDOs via paired T-tests. A bar chart was plotted using Microsoft Excel (Microsoft, Washington, USA) to visualise the data (Figure 9.15 (C)).

Organoids derived from Glasgow patient tissue BB200099 were established and stained for Epcam1 and α-SMA. Representative images are shown in Figure 9.16 (A), with positive staining observed for both markers indicating the presence of tumour and stromal calls (Figure 9.16 (A)). Drug screening with JAK inhibitors Ruxolitinib and Tofacitinib was performed only to an n=1 at a concentration of 10μM, therefore no statistical testing was performed. Light microscopy images were taken 72 hours post-treatment. Representative images of control and vehicle treated organoids show healthy cells (Figure 9.16 (B)). Ruxolitinib and Tofacitinib treated PDOs had similar morphology to control samples (Figure 9.16 (B)). The PDOs treated with 5FU appeared to be dead and much smaller in size (Figure 9.16 (B)). A bar chart was plotted to visualise the WST-1 data, 5FU-treated PDOs had a lower percentage of viable cells compared to vehicle control, and the same was observed for Ruxolitinib treatment, albeit to a lower extent (Figure 9.16 (C)).

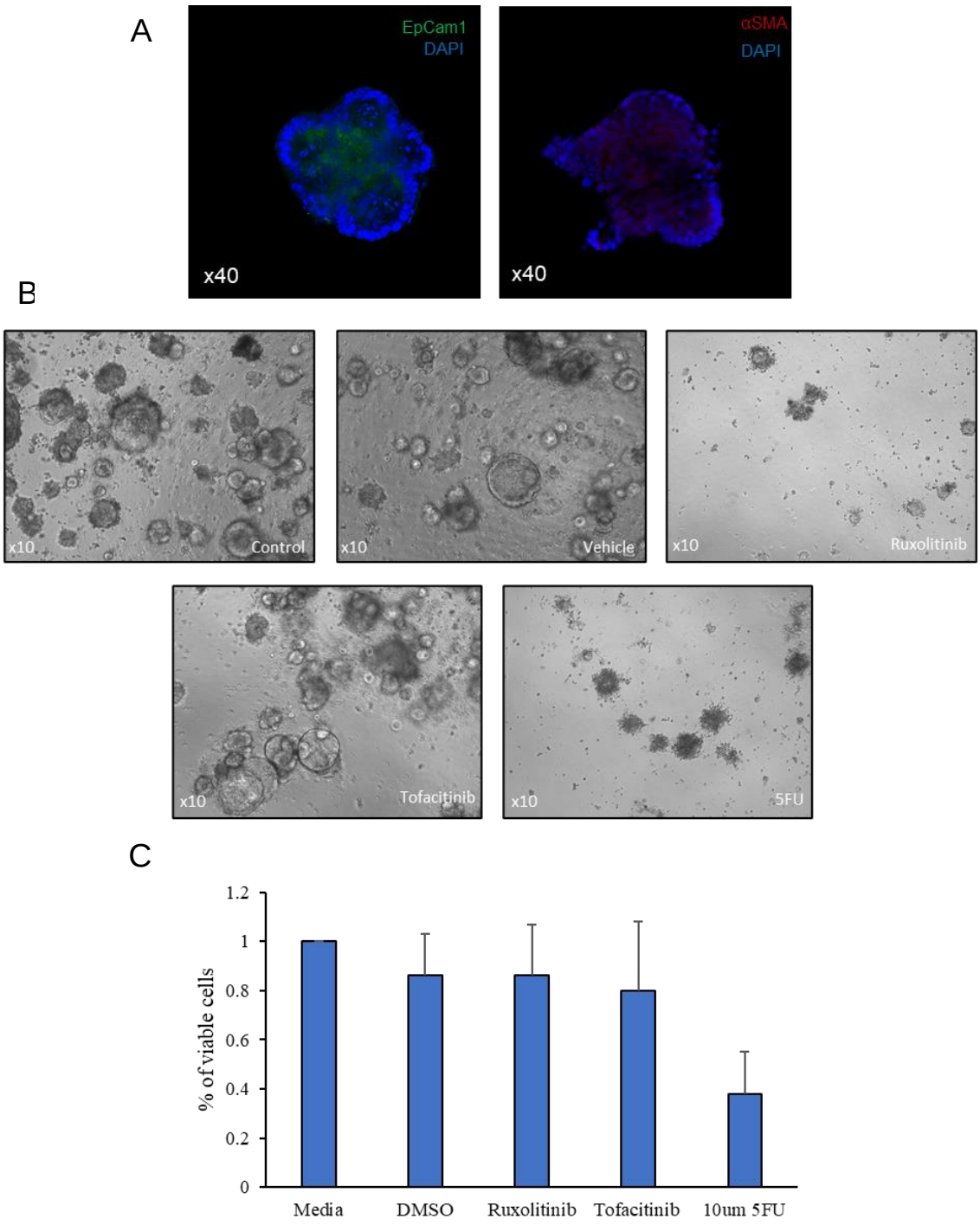


Figure 9.15 Drug screening of JAK inhibitors in Glasgow BB200040. *Representative images of IF-stained BB200040 PDOs for EpCam1 and α-SMA (A). Bar chart showing cell viability of BB200040 following treatment with 0.01% DMSO, 1mM Ruxolitinib, 1mM Tofacitinib and 1mM 5FU (B).*

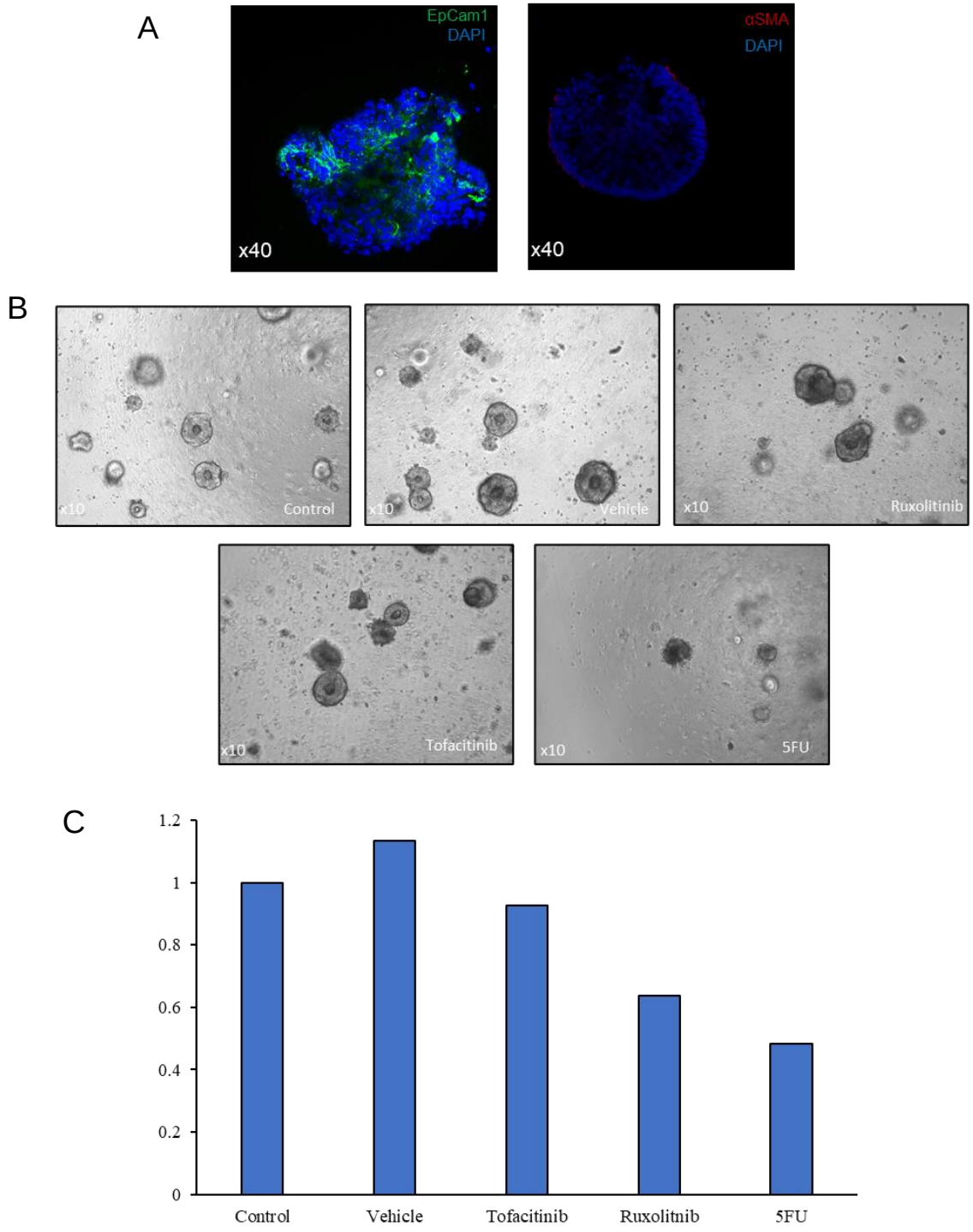


Figure 9.16 Drug screening of JAK inhibitors in Glasgow BB200099. *Representative images of IF-stained BB200099 PDOs for EpCam1 and α-SMA (A). Bar chart showing cell viability of BB200099 following treatment with 0.01% DMSO, 1mM Ruxolitinib, 1mM Tofacitinib and 1mM 5FU (B).*

The PDOs established from BB190240 were grown and then stained via IF for EpCam1 and α-SMA to assess tumour and stromal compartments, respectively. Representative images of staining are shown in Figure 9.17 (A). Drug screening was performed to investigate the effect of JAK inhibitors on PDO cell viability. Representative images of light microscope images are shown in Figure 9.17 (B). From these images, control, and vehicle control PDOs appear healthy and had similar morphologies 72 hours post-treatment. However, 10µM Ruxolitinib and 10µM 5FU samples had dark cores and membrane blebbing indicating cell death. The 10µM Tofacitinib-treated PDOs appeared similar in morphology to both control samples. A WST-1 assay was used to measure cell viability 72 hours post-treatment. A bar chart was plotted in Microsoft Excel (Microsoft, Washington, USA) to visualise the percentage of viable cells in each treatment group (Figure 9.17 (C)). Paired T-tests were performed to determine any statistical difference between vehicle control and drug-treated groups and no difference were observed at $\alpha=0.05$. The dose of JAK inhibitors used in this experiment was only 10µM as opposed to the 1mM utilised when treated Sanger organoids. This was used as it was felt to be a more clinically relevant dose, however response in terms of reduction in cell viability was not as evident at this lower concentration.

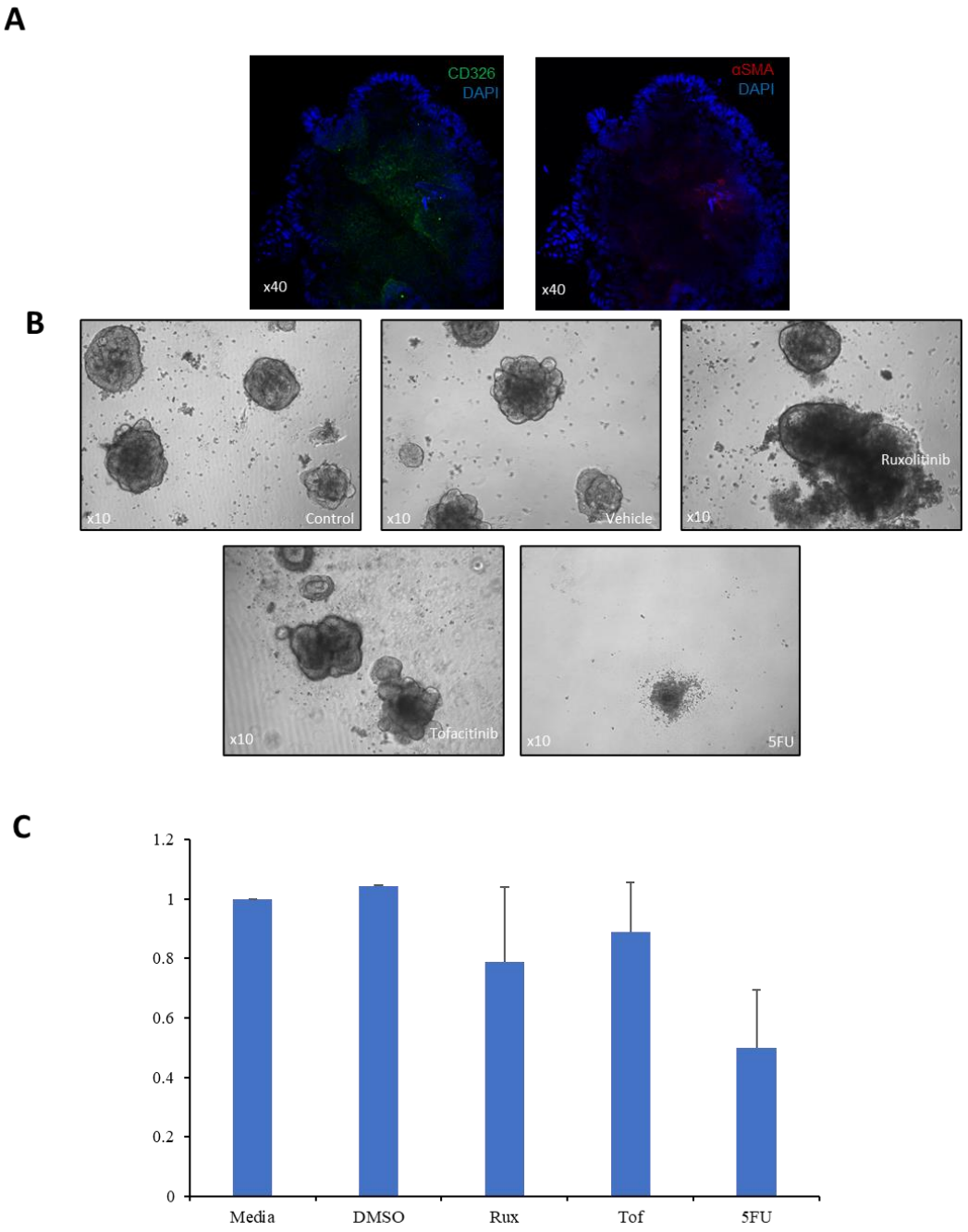
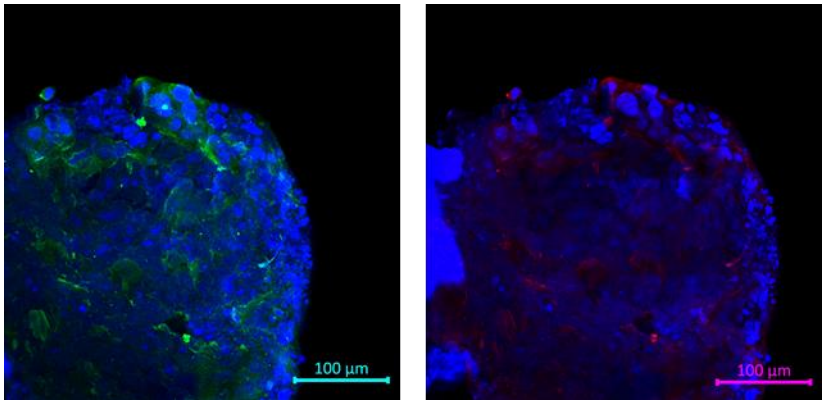


Figure 9.17 Drug screening of JAK inhibitors in Glasgow BB190240. *Representative images of IF-stained BB190240 PDOs for EpCam1 and α-SMA (A). Bar chart showing cell viability of BB190240 following treatment with 0.01% DMSO, 1mM Ruxolitinib, 1mM Tofacitinib and 1mM 5FU (B).*

A



B

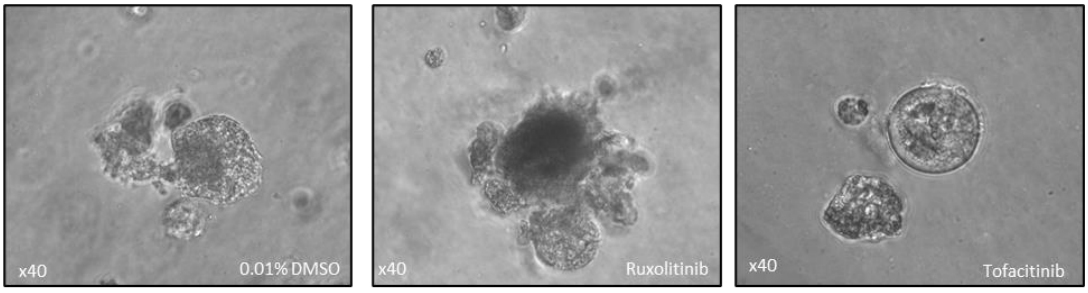


Figure 9.18 Characterisation and pilot screening of JAK inhibitors in Glasgow BB200534. *Representative images of IF-stained BB200534 PDOs for EpCam1 and α-SMA (A). Light microscope images of PDOs treated with 0.01% DMSO, 10µM Ruxolitinib, or 10µM Tofacitinib taken at x40 objective magnification.*

Developing the method to grow patient-derived organoids directly from patient tissue was a challenging and time-consuming process, and some pilot experiments were not successful through to full cell viability assay. PDOs from BB200534 were characterised and showed positive staining for tumour cells via EpCam1 and stromal cells via α-SMA (Figure 9.18 (A)). Although for this patient sample WST-1 assay data was not available, light microscope images revealed distinct difference in morphology between PDOs treated with 0.01%DMSO vehicle control and 10µM Ruxolitinib (Figure 9.18 (B)). Inhibiting JAK1/2 visually appeared to kill cells as PDOs formed a dark necrotic core. There was no morphological difference observed between Tofacitinib treated cells and vehicle control (Figure 9.18 (B)). Similarly, organoids derived from BB190614 were characterised for α-SMA expression via IF and light microscope images were taken to visually assess response to the JAK1/2 inhibitor Ruxolitinib (Figure 9.19 (A-B)). PDOs treated with control media and 0.01% DMSO appeared healthy at 72 hours but those treated with 10µM Ruxolitinib or 10µM 5FU appeared fragmented and as if undergoing cell death (Figure 9.19 (B)). Two PDO samples were characterised but did not proceed to drug screening assays. BB200154 and BB200011 showed modest expression of EpCam1 and α-SMA when stained via IF and representative images are shown in Figure 9.20.

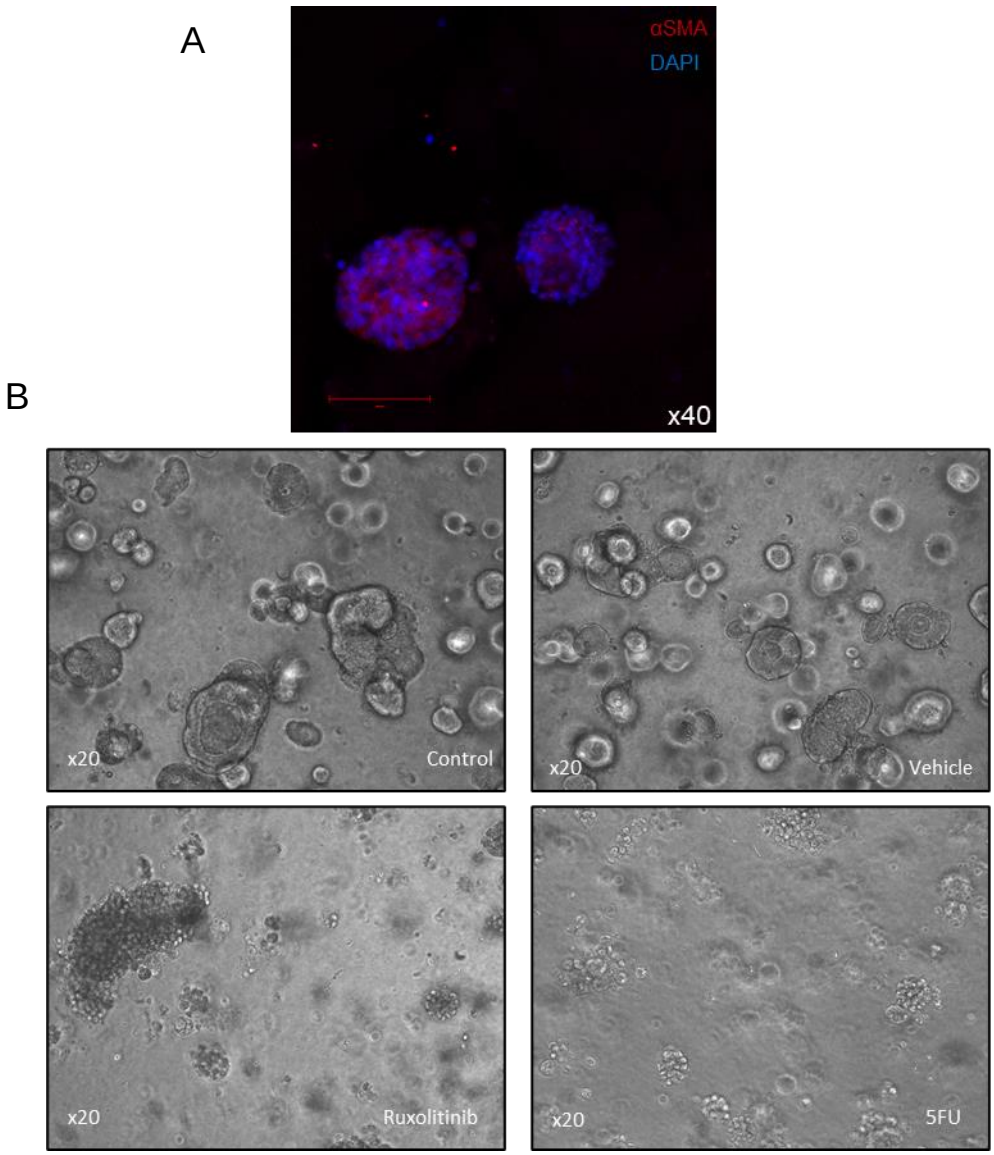


Figure 9.19 Characterisation and pilot screening of JAK inhibitor in Glasgow BB190614. *Representative images of IF-stained BB190614 PDOs for α-SMA (A). Light microscope images of PDOs treated with media control, 0.01% DMSO, 10μM Ruxolitinib, or 10μM Tofacitinib taken at x40 objective magnification.*

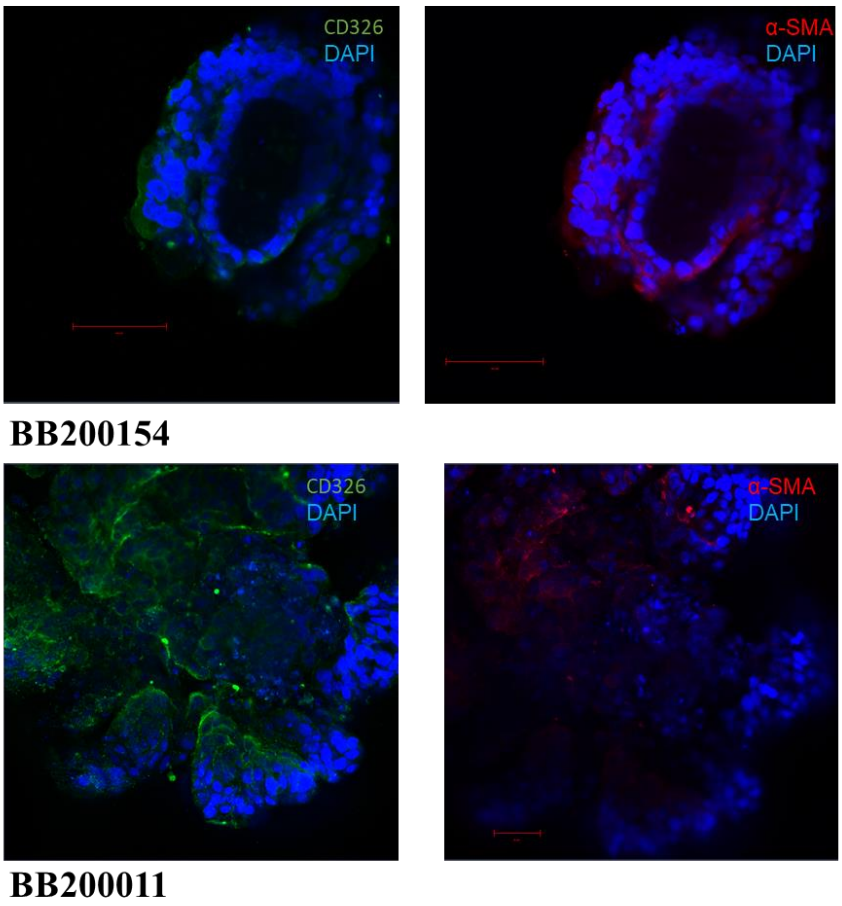


Figure 9.20 Characterisation of Glasgow PDOs. *Representative images of IF-stained BB200154 and BB200011 PDOs for EpCam1 and α-SMA.*

9.5 Discussion

IL6/JAK/STAT3 signalling is dysregulated in CRC. Inhibiting signal transduction represents a promising therapeutic strategy for improving CRC survival outcomes in a subset of patients. Protein data from chapter 4 highlighted a strong association between activated STAT3 and poorer outcomes specifically in patients with GMS2 stromal-rich tumours. Here, we investigated inhibiting upstream pathway components of pSTAT3^{tyr705} using two repurposed JAK inhibitors in 2D and 3D tumour models.

Ruxolitinib is a potent JAK1/2 inhibitor FDA approved for the treatment of myelofibrosis (204). Tofacitinib targets JAK2 and preferentially JAK3 and was approved in 2011 for the treatment of the autoimmune disorder rheumatoid arthritis (10). Repurposing of these agents would be easily translatable to CRC, if efficacy was proven, as they are already known to be well-tolerated and safe in patients with other diseases.

There is accumulating evidence that supports inhibiting STAT3 signalling in preclinical models slows cancer growth and differentiation. For example, in head and neck cancer xenograft models, inhibiting JAK1/2 with AZ1480 significantly reduced tumour proliferation and expression of $pSTAT3^{tyr705}$(206). Similarly in ovarian cell lines, Ruxolitinib synergises with paclitaxel, cisplatin and other chemotherapeutic agents to reduce tumour cell viability, promote apoptosis and reduce tumour volume in xenograft models (207). Meanwhile, JAK2/3 inhibitor Tofacitinib has shown promise for myeloma through preventing cancer cell line proliferation (203). Focusing more specifically on CRC, inhibiting JAK2 in HCT116, SW480 and HT29 cell lines promotes apoptosis and inhibits proliferation (78) In another study JAK1/2 inhibition in HT29 and SW1116 cell lines reduced cell growth, increased apoptosis, induced cell cycle arrest and inhibited tumour cell invasion (208).

In this chapter, a reduction in the percentage of viable cells following treatment with Ruxolitinib and Tofacitinib was observed in a subset of available CRC 2D cell lines. The most profound responses were seen in CRC lines classified as CMS4 by their transcriptomic profiles and histologically subtyped as stromally-dense. This would fit well with the hypothesis that GMS2 patients are most likely to benefit from therapeutic targeting of the pathway based on aforementioned IHC protein data. The dose of Jakinib required to elucidate a response was potentially not clinically relevant to the dose that would ultimately be tolerated in humans. Therefore, further research is required to determine if combining JAK inhibition with standard of care chemotherapy would synergistically promote tumour destruction. The proposal would be that similar to novel immunotherapeutic drugs such as anti-PDL1, which have recently been approved for use in some CRC cases, Jakinibs would be given alongside standard of care chemotherapy. In terms of the cell lines experiments, data could be enhanced by stimulating cells with IL6 pre-treatment to better recapitulate human disease and ensure pathway activation. Additionally, ELISAs could be performed again using a positive control to relate constitutive $pSTAT3^{tyr705}$ levels to a proven high-expressor, for example fibroblast lines treated with IL6. It would also appear from the data and literature that JAK2 is the main JAK family member involved in driving CRC progression and therefore identifying a repurposable drug which specifically targets only JAK2 should be investigated.

Although the results from JAK inhibition in 2D lines could have been more promising, there are now more advanced models of disease to perform drug screening on, which are indictors of patient response. Patient-derived organoids are a 3D tumour model with recapitulate CRC clinical disease better than 2D immortalised cell lines(202). PDOs can be utilised to study the potential therapeutic effects of novel/repurposed drugs and predict patient response to therapy (209). Development of tumouroid systems was a complex process with very specific culture conditions required for establishment and optimal growth. In this chapter PDOs derived from Sanger centre patients and surplus tissue from Glasgow patients were utilised to study the effects of JAK inhibitors on cell morphology, viability and in some cases proliferation and apoptosis. Treatment with 1mM Ruxolitinib

resulted in a reduction in the percentage of viable cells in most cases. This was not observed following treatment with 1mM Tofacitinib. The dose of 1mM is not clinically relevant and further work is needed to elucidate if lower doses of Ruxolitinib could still be therapeutically beneficial in a subset of patients. Further work is also required to investigate if combination of 5FU with Ruxolitinib could produce a synergistic effect resulting in enhanced reduction in cell viability.

Although PDOs represent a better model of disease than 2D cell lines there are still limitations associated. The TME consistent of a plethora of cell types broadly categorised into the immune, stromal and tumour compartments. In terms of PDOs, the only compartment available to study is the tumour cells. Given that STAT3 signalling is known to be present in stroma and immune populations testing the drug only on tumour cells means that results should be interpreted with caution. Utilising a mouse model to test Jak inhibitors would provide a more comprehensive model of disease, although this does also have some limitations including the issues with translatability and ethical considerations. A coculture system where PDOs are grown alongside immune cells or fibroblasts could also be a good model for studying mechanisms and providing a better model than tumouroids alone, however this system is still artificial and no considered the gold standard of modelling patient disease.

Another limitation of this chapter was measuring response solely using cell viability assays. In future work, further measures of response should be considered including apoptosis assays, cell cycle assays.

Given that these inhibitors are already used clinically, the need to investigate mechanisms of action was overlooked somewhat. Future work should seek to repeat ELISA assays using a positive control for $pSTAT3^{tyr705}$, stimulate matched plates with IL6 to try and observe an enhanced response to inhibition and silence STAT3 in matched cell lines to investigate whether the effect of the inhibitor on cell viability in these cells is lost.

In addition to chemotherapy and JAK inhibitor combination therapy, there is evidence that JAK inhibitor action may be enhanced by combination therapy with chemo-radiotherapy (cRT). JAK2 is the predominant JAK family oncokinase involved in driving CRC and JAK2/STAT3 signalling may be involved in resistance to radiotherapy(130). Pilot data from a proof-of-concept study lead by Mr Campbell Roxburgh show at the transcriptomic level, IL6/JAK/STAT3 gene signatures are upregulated following radiotherapy in serial pre- and post-treatment rectal biopsies. These data suggest combination therapy of JAK2 inhibitors with cRT may be therapeutically beneficial for a subset of rectal cancer patients. Future work should include investigating this in 2D lines, PDOs and xenograft models.

In conclusion the data from this chapter show that stromal CMS4 CRC cell lines observed the most promising response to inhibition of JAK family proteins. Furthermore, in 3D tumouroid models, which recapitulate patient disease to a greater extent than 2D models, JAK1/2 inhibitor Ruxolitinib

successfully reduced cell viability compared to vehicle controls in a subset of patients, particularly those with stromal-rich GMS2 tumours. Further work is necessary to understand the mechanisms underlying drug responses observed, to perform screening in a higher throughput system and to unpick the clinical relevance of combination therapy of JAK inhibitors with standard of care chemotherapy or cRT.

Chapter 10: General Discussion

10: General Discussion and future work

Colorectal cancer remains the third leading cause of cancer-related death worldwide and it is therefore pertinent that scientific research identifies new therapeutic options for patients. CRC is characterised by all seven of the hallmarks of cancer (126). CRC is a heterogenous disease and treatment regimens need to evolve to target factors driving tumorigenesis on a patient-by-patient basis with a precision medicine-based approach. This thesis firstly aimed to validate clinically translatable histological subtyping methods GMS and phenotypic subtyping, and subsequently investigate the role of one inflammatory signalling pathway, IL6/JAK/STAT3, in a retrospective stage I-IV CRC cohort through immunohistochemistry, RNAscope®, mutational profiling and bulk RNA-sequencing. The secondary aim was to determine if pathway inhibition could be a novel therapeutic target for a histologically classified subset of CRC, which was investigated through drug screening in 2D CRC cell lines and patient-derived 3D tumour models.

Data from previous in vivo mouse model studies have implicated IL6/JAK/STAT3 in CRC development and progression. For example IL6-/- mice injected with azoxymethane and dextran sodium sulphate (DSS) had significantly reduced tumour number and size compared to wildtype controls (113) This data built on work from Becker at al in 2004 which showed a reduction in tumour growth following IL6 inhibition(145, 210). Loss of STAT3 in DSS treated mice resulted in increased apoptosis compared to IL6 ablation, which itself caused an increase in apoptosis compared to wildtype (113). In mouse xenograft models, injecting HT29 cell lines expressing dominant negative STAT3 resulted in the formation of smaller tumours than wild type HT29 cells (145). HT29 cells with constitutively active STAT3, or those stimulated with IL6 showed increased proliferation compared to dominant negative STAT3 lines (145). In other studies which aimed to block IL6 signal transduction in mouse models, inhibiting gp130 in colitis associated premalignant mice caused a reduction in pSTAT3 expression and using short hairpin RNA to block IL6/STAT3 inhibited tumour growth (211, 212). These data from earlier mouse models support the hypothesis that inhibition of STAT3 signalling could be a promising therapeutic strategy for CRC. The work in this thesis aimed to investigate the translatability of IL6/JAK/STAT3 as a target in human tissue and relative to a novel, prognostic histologically subtyping method.

Tissue-based studies- validation of histological subtypes

A vast body of research over the past decade has focused on finding ways to classify CRC disease to predict prognosis and optimal treatment regimes. The development of the consensus molecular subtypes (CMS) in 2015 and CRIS subtypes in 2017 were steps towards segregating disease for prognostic information (2, 3). More recently, CMS subtyping has been analysed for its ability to predict response to chemotherapy regimens. A study by Okita et al in 2018, highlighted a progression-free survival benefit in CMS4 patients receiving first-line irinotecan-based therapy over oxaliplatin-

based therapy (213). In the same study it was observed that CMS2 patients responded well to anti-EGFR therapy, and CMS1 patients who received EGFR inhibitors exhibited significantly worse outcome (213). However, another study focussed on the MOSAIC clinical trial cohort of stage III CRC found no survival benefit of oxaliplatin-based therapy for any subtype (214). Evidence for using CMS to predict optimal therapy is still limited and further research is required.

The transcriptomic techniques required to determine a patients CMS/ CRIS subtype are time-consuming, expensive, and not currently feasible for routine clinical testing. This led to the development of GMS and phenotypic subtyping, which are determined by pathological assessment of H&E and Ki67 stained sections (103, 104). From chapter 3, GMS was independently prognostic in the stage I-IV Glasgow combined cohort. Patient with a low inflammatory infiltrate and high stroma percentage (GMS2) observed the worst prognosis and those with a high inflammatory infiltrate (GMS0) show markedly better outcome in terms of CSS. This was validated in the TransSCOT clinical trial high risk stage II-III CRC cohort and the stage IV synchronous resection cohort. In the TransSCOT cohort GMS0 immune patients observed a significantly better clinical response in terms of disease-free survival to FOLFOX chemotherapy(47). Interestingly, in the stage IV mCRC cohort, GMS of primary tumour was significantly predicative of GMS of matched liver metastases GMS. This is potentially very important as if GMS is used to predict optimal therapeutic regimen in the future, there would need to be phenotypic concordance between sites for this approach to work in mCRC cases. GMS showed no association with clinicopathological characteristics in T stage I/II screen-detected DM-CRC-TMA cohort, suggesting it should be confined to >stage II disease. There was also no association between GMS and CSS in the Australian TMA stage II-III cohort when KM grading and TSP were performed on tissue cores as opposed to full tumour resections. These data suggest for performing GMS accurately, full tumour sections must be assessed. This is likely due to discrepancies with KM grading, as differentiating between patchy and full bands may not be possible on a TMA even when invasive edge cores are present (215). One method to combat this could be to replace KM grading with CD3+ T cell IHC staining, which has previously been proven to work in tumour biopsies by Park et al (215). Neoadjuvant therapy can alter the TME, therefore utilising preoperative pre-treatment biopsies rather than tumour resections would be the optimal method for determining tumour phenotype (215).

GMS can be further segregated into 4 phenotypic subtypes as devised by Roseweir et al (104). GMS0 is referred to as immune phenotypic subtype and GMS2 is the stromal subtype. GMS1 is split into 2 group based on Ki67 proliferation index (104). Patients with a high proliferation rate are classed as canonical and those with a low proliferation rate are assigned to the latent phenotypic subtype (216). Phenotypic subtype was associated with CSS in the Glasgow combined cohort, TransSCOT clinical trial cohort and primary tumours of the synchronous resection cohort. From the TransSCOT cohort, patients in the immune subtype had improved disease-free survival when treated with FOLFOX

chemotherapy, and conversely the latent group trended towards a better response to CAPOX chemotherapy. Further research is required to validate these findings in other retrospective CRC cohorts with chemotherapy data. Concordant with GMS, phenotypic subtype was not able to stratify patient survival in the Australian TMA cohort or the screen-detected stage I-II DM-CRC-TMA cohort. In the liver metastases of the synchronous cohort, there was a trend towards an association of phenotypic subtype with CSS and a larger cohort may yield significant results.

This thesis has shown that in stage II-IV CRC both histological scoring methods are independently prognostic when determined using full tumour resections in stage II-IV primary CRC. However, perhaps the real challenge is identifying how we can use these phenotypes to inform therapeutic decision making in the clinic. Thus far the only study that has investigated how histological subtyping could be used to guide treatment, found that patients in the immune GMS0 group responded better to CAPOX chemotherapy than FOLFOX chemotherapy, was published in 2020 (47). Further work is needed to validate these findings and determine if GMS/phenotypic subtyping could be used to predict response to novel/repurposed targeted biologic therapies.

Tissue-based studies- IL6/JAK/STAT3

Stromal subtype patients have the poorest prognosis, whether you look at the transcriptomic level of CMS (CMS4) or histologically via GMS (GMS2) and phenotypic subtyping (stromal)(2, 103, 104). Therefore, this group represents a key subset of patients who need better treatment options, through novel/repurpose-able drugs. IL6/JAK/STAT3 signalling has been implicated in CRC development and progression, and several inhibitors of STAT3 activation are already approved for other diseases, which makes this pathway a promising target for CRC therapeutics (9). In this thesis we found that key members of IL6/JAK/STAT3 signalling were associated with worse prognosis, particularly in patients assigned to GMS2 stromal subtype in a stage I-IV retrospective CRC cohort.

Data from chapter 4 showed that high IL6 mRNA copies within the tumour stroma were associated with reduced CSS in the full Glasgow combined cohort. There was no association between tumoural IL6 mRNA copies and CSS in the full cohort or when disease was segregated by GMS, MMR status or tumour subsite. Previous studies have shown the importance of circulating IL6 in CRC, however there is limited data focused on IL6 in tissue-based studies (217). High expression of the cognate receptor of IL6, IL6R, was also associated with poor prognosis, particularly in GMS2 patients in the Glasgow combined cohort. Interestingly, expression of IL6R was mostly noted in the tumour cytoplasm suggesting soluble IL6R may be more important than receptor bound IL6R in the CRC setting. sIL6R allows for signalling in a variety of cells due to ubiquitous expression of gp130 tails, which forms part of the receptor needed for signal transduction (218). Previous studies have suggested that the switch from classical signalling via membrane bound IL6R to *trans*-signalling via sIL6R is

associated with tumorigenesis, inflammatory processes and is implicated in driving metastases (218, 219).

For IL6 to activate STAT3, intermediate pathway components, JAK family members must be phosphorylated. JAK1 and JAK2 are thought to be the main signal transducers for this pathway, however there is also evidence that TYK2 and to a lesser extent JAK3 can be involved in STAT3 activation (220). Results from chapter 5 suggest that JAK2 is perhaps the key pathway member implicated in CRC. High expression of cytoplasmic JAK2 was significantly associated with reduced CSS in GMS2 stromal patients from the Glasgow combined cohort. A combined score of JAK1 and JAK2 within the tumour cell membrane was also able to stratify survival in the Glasgow combined array, with high expression predicative of worse cancer-specific survival time. High protein expression of JAK2 has recently shown to associate with reduced overall survival in pancreatic ductal adenocarcinoma and ovarian cancer (221, 222). However, in breast cancer there is evidence that high JAK2 expression is in fact associated with better clinical outcomes (223, 224). In a study of 62 patients, high JAK1 expression was associated with increased stage and independently associated with reduced prognosis in colon cancer (225). In the same study they found a similar role for total STAT3 expression (225). Further mechanistic research is required to understand JAK family proteins driving STAT3 activation in CRC and to validate whether JAK2 is the preferential kinase for signal transduction of this pathway.

In this thesis no association between total STAT3 in the nucleus or cytoplasm was associated with outcome from chapter 6 data. High nuclear expression of pSTAT3^{tyr705} trended towards an association with reduced CSS in the full cohort, and this was potentiated in the GMS2 stromal subtype. It was hypothesised that nuclear STAT3 would be significantly associated with reduced prognosis as nuclear localisation is a surrogate marker of activation and STAT3 is held inactive in the cytoplasm. In the literature, several studies have assessed the prognostic significance of tumoural phosphorylated STAT3 levels in CRC, with conflicting results. One study in 2010 found that higher pSTAT3^{tyr705} specifically amongst advanced rectal cancer patients from the EORTC 22921 clinical trial was indicative of significantly better outcomes (226). However, in colorectal cancer retrospective studies, Kusaba et al and Park et al, observed a survival benefit in patients with low pSTAT3^{tyr705} protein expression (143, 148). This may highlight the underlying difference in colon versus rectal cancer cases. However, in chapter 6 when patients were stratified by tumour subsite pSTAT3^{tyr705} expression the greatest survival benefit of low expression was observed in rectal cases (although p=0.065). Most studies focus on pSTAT3^{tyr705} due to the main role it plays in pathway activation. There is more limited evidence on involvement of pSTAT3^{ser727}, however for maximal pathway activation at serine 727 must occur in addition to initial activation at tyrosine 705 (71). In the Glasgow combined array there was no association between pSTAT3^{ser727} expression and survival in the full cohort, or when patients were segregated by GMS, MMR status of tumour subsite. However, when the two STAT3

phosphorylation sites were combined to form a new score of 0 (low for both), 1 (high for one) and 2 (high for both markers), high expression of both proteins was significantly associated with reduced CSS. Another interesting observation was the expression pattern of pSTAT3^{ser727}, which was mainly observed in the cytoplasm of tumour cells in the Glasgow combined array contrary to the expectation of nuclear localisation. It's not clear why the pattern was observed, and previous studies focus on expression in the nucleus. In glioblastoma high expression of nuclear pSTAT3^{ser727} associated with reduced progression-free survival in a retrospective cohort of 88 patients assessed via IHC (227). In prostate cancer, activation at serine 727 has been shown to drive cancer progression independently of tyrosine 705 activation (228). Further research is required to elucidate the implication of STAT3 activation sites in CRC tumorigenesis and validate expression relative to histological subtyping. It would be of interest to perform IF staining of both phosphorylation sites to investigate levels of colocalization. It would then be possible to determine if patients that showed high expression of both pSTAT3 proteins in the same cells observe worse prognosis than patients with high expression of each marker in neighbouring/different cells. This could provide evidence that both pSTAT3 markers need to be expressed in the same cell for maximal pathway activation.

One potentially important factor is the location of STAT3 within the tumour microenvironment. This thesis focused solely on tumoral expression of each pathway member (aside from IL6 measured in the stroma). Potentially of equal importance is activation of the pathway amongst the immune infiltrate and cancer associated stromal cells. Histological scoring of expression of each pathway member should be assessed in these other components of the TME in future work. STAT3 is known to be expressed in a number of tumour-infiltrating immune cells, for example regulatory T cells, CD8+ T cells, and macrophages. In vitro studies have demonstrated a pro-tumour role for STAT3 in cytotoxic T cells, with high expression associated with downregulation of chemokine receptor CXCR3 which blocks their recruitment to the TME (229). Conversely, deletion of STAT3 in macrophages enhances tumour progression in breast cancer models (230). One of the hypotheses to come out of this thesis is that the source of IL6 is mainly of stromal origin, which is why GMS2 patients with stromal-rich tumours show the worst prognosis with pathway activation. However, this is likely a more complex process and it is well characterised that macrophages and other myeloid populations can also be potent IL6 producers (75). CAFs and MSCs recruit tumour-promoting myeloid cells to the TME, and stromally dense tumour are associated with increased presence of macrophages of the M2 phenotype. (231, 232). In prostate cancer models CAFs and M2 macrophages have been shown to work together to promote an immunosuppressive cancer-promoting microenvironment (232). There may be a network of feedback events which orchestrate STAT3 hyperactivation in GMS2-like phenotypes, ultimately leading to the poor patient outcomes observed. Further research to understand the mechanisms of stroma-immune-tumour crosstalk leading to aberrant STAT3 tumour signalling is necessary.

Underlying biology of tumour phenotypes

We hypothesised that patients with high levels of STAT3 signalling may have similar underlying biology in terms of their mutational landscape and transcriptomic profiles. Overall, the results from the Glasgow combined discovery cohort indicated that targeting STAT3 activation in patients may be of clinical benefit, particularly in patients with stromal-rich tumours. The protein levels of some pathway members were prognostic, however there is a lack of evidence regarding the factors driving hyperactivated STAT3 phenotypes. Mutation and transcriptomic profiles were investigated with respect to high/low $pSTAT3^{tyr705}$ groups as a step towards understanding the biology potentially driving phenotypes observed. Identifying mutations and transcriptomic patterns associated with aberrant STAT3 signalling was also investigated to gain insight into any other cellular signalling pathways that could be dual targeted with STAT3 to produce a synergistic therapeutic benefit.

Mutations

At the time of writing to our knowledge no previous studies had investigated the underlying biology of activated STAT3 phenotypes in any type of cancer. Data from chapter 7 showed patients with hyperactivated STAT3 phenotypes were more likely to have MAP2K mutations. This was of interest due to emerging evidence that crosstalk between STAT3 and MAPK signalling is implicated in cancer (89). Dual targeting of both pathways has already been investigated in mouse models of pancreatic ductal adenocarcinoma, with promising results in terms of tumour regression in xenograft models using MEK inhibitors (233). Similar results were obtained in a study of melanoma cell lines, where dual inhibition of MAPK and STAT3 yielded better effects on reducing cell growth than inhibiting one pathway alone (182).

In addition to investigating the mutation patterns of STAT3 phenotypes, mutations in pathway member genes included in the panel were also analysed for association with prognosis and clinical features of disease. In the Glasgow combined cohort, a combined score of presence of mutation in ≥1 of JAK1, JAK2 or JAK3 gene/s was not significantly associated with CSS. However, patients with a mutation in at least one JAK gene were more likely to present with unfavourable clinical characteristics including higher T stage, dMMR disease and more peritoneal involvement and marginal involvement. The most characterised somatic mutation in the literature from JAK/STAT signalling pathway is JAK *V1617F*, which is acquired in almost all cases of myeloproliferative cancers (234). In terms of other cancers, its prevalence is much rarer, with one study of ~10500 cancer cases identified mutations in only 0.2% of cases (234). In the Glasgow combined cohort, the rate of mutation at least one JAK member was 11%. Mutations in the STAT3 gene itself are rare in solid cancers, and it is generally accepted that hyperactivated signalling can only be driven by mutation in upstream pathway members (216). Data from chapter 7 suggests that in CRC, aberrant STAT3

signalling is not driven by mutation in upstream kinases, with JAK mutations not associated with pSTAT3^{tyr705} protein expression when assessed via chi-squared analyses.

Transcriptomics

Given that mutation profile did not seem to drive STAT3 activation in the Glasgow combined cohort, the transcriptomic profile of 100 patients was assessed for association with pSTAT3^{tyr705} groups. Data from chapter 8 identified no significant differentially expressed genes between pSTAT3^{tyr705} protein expression groups in the full cohort or when the top 25 expressers, 'extremes', of tyrosine 705 expression were selected for analysis. This may be due to the heterogeneity of cancer and too small a sample size to pull out transcriptomic differences. Bulk RNA-sequencing was performed probing for ~22000 genes, which may have been too large a dataset to observe differential genes. The use of 'extremes' was employed to remove patients near the threshold line of high/low pSTAT3^{tyr705} protein expression, however still no differentially expressed genes were identified. Further work could include increasing the size of the cohort and/or analysing the groups with genes filtered to only include cancer-associated genes or the genes used to define CMS. For the purposes of analyses, the adjusted p value (padj) was set to 0.05 for significance however, some studies use padj <0.1, so subsequent analyses could be performed with a higher padj which may reveal some differentially expressed genes between groups.

Further subsequent analyses could be to analyse the prognostic effect of gene expression counts of pathway members. In previous studies gene set enrichment analysis in pre-cancerous colon polyps has implicated JAK/STAT3 signalling in driving CMS1 disease (235). STAT3-associated gene signatures have also been implicated in specifically driving basal-like breast cancers, but not luminal A or luminal B disease (236). In a 2019 renal cell cancer (RCC) study by Robinson et al, 32 STAT3-associated genes were assessed for association with renal cancer subtypes. This study identified one subtype (clear cell RCC) which was enriched for STAT3 at the transcriptomic level. A similar method could be employed with the data from chapter 8 in this thesis whereby the 32 gene signature could be analysed relative to CMS, GMS and phenotypic subtypes to investigate if akin to the protein data, CMS4/GMS2/Stromal subtype patients would be most likely to benefit from pathway inhibitors (237).

In vitro studies

The IHC-based protein data from chapters 4, 5 and 6 suggested inhibiting STAT3 activation could be of therapeutic benefit for a subset of CRC patients and thus the hypothesis was tested in preclinical models. Several Jakinibs are already approved for clinical use in myeloproliferative and inflammatory disorders and therefore repurposing for colon or rectal cancer is a good option for safety purposes and speed of clinical trials. Jak inhibitors already approved for clinical use include Ruxolitinib (JAK1/2), Tofacitinib (JAK2/3), Baracitinib (JAK1/2) and Upadacitinib (JAK1) (238).

Previous studies in preclinical models have shown inhibition of STAT3 activation causes decreased levels of angiogenesis, proliferation, invasion, metastases, and increased apoptosis in number of tumour types including head and neck squamous cell carcinoma, ovarian, breast and gastric cancer (206, 239, 240). Similarly, colon cancer cell lines HCT116, HT29 and SW480s have shown decreased cell proliferation and increased apoptosis following treatment with JAK inhibitor AZD1480 (78). Comparable results have also been observed in SW620 colon cancer cell lines in another study (241). In chapter 9 of this thesis, seven 2D CRC cell lines were investigated for effects on cell viability following treatment with Ruxolitinib and Tofacitinib. Of the cells assayed, CMS4 categorised line HCT116 showed the most promising results following treatment with JAK2/3 inhibitor Ruxolitinib, with significant reductions observed at doses of 1mM, 0.1mM and 10μM. This result corroborated our hypothesis driven by the protein based IHC data that suggested stromal subtyped patients were the most likely subgroup to benefit from therapeutic targeting of IL6/JAK/STAT3. Data from cell viability of CMS1, CMS2 and CMS3 cell lines were more modest with responses only observed at 1mM drug concentration which is not representative of clinical drug doses. Although the reduction in cell viability is indicative of potential therapeutic benefit further research is required to investigate the effects of Ruxolitinib/Tofacitinib on tumour cell apoptosis, proliferation, and cell cycle arrest. The use of 2D models is cost-effective and easy to perform, however cell lines do not recapitulate human disease effectively. Utilising seven different lines with varying genetic and transcriptomic backgrounds helped account for different tumour types, however development of protocol to grow patient-derived organoids enabled assessment of Jak inhibitors in profoundly better models of CRC disease than the 2D lines.

In the available tumouroids treatment with JAK1/2 inhibitor Ruxolitinib reduced cell viability, increased apoptosis, and decreased proliferation as measured by WST-1 assay and IF staining, respectively. CMS3 and CMS4 subtyped PDOs observed the most promising response to Jakinib treatment, which supports the hypothesis formed from 2D cell line work and tissue-based studies from this thesis. When blinded GMS phenotyping was performed on the original resection block, Sanger31 PDO, which showed a good response to inhibition of STAT3 was GMS2 and stromal phenotypic subtype. Due to difficulties with growth conditions and extensive protocol optimisation, only 12 lines were utilised for the purposes of this thesis. To gain a more thorough understanding of patient populations likely to benefit from targeting STAT3 activation a larger pool of PDOs would need to be used. Further work which could be employed to make the drug screening system higher throughput could be utilisation of microfluidic devices, which enable assessment of individual organoids post-treatment and allow for a gradient of drug concentrations to be applied in one assay. This technology was developed by collaborators (Dr Michele Zagnoni and Dr Theresa Mulholland) at the University of Strathclyde and has already proved successful in the assessment of primary prostate cancer spheroids with standard of care chemotherapeutic agents (242).

Future Perspectives

Future work following on from this thesis should look to validate that STAT3 signal transduction is associated with poor outcomes in stage II-IV CRC and further investigate if inhibiting the pathway would be therapeutically beneficial in the clinic.

Expanding organoid data

One of the most novel findings of this study was that at the protein level, aberrant STAT3 signalling associated with the worst prognosis in patients with stromally-dense tumour classified as GMS2. This finding should be validated in another large retrospective CRC cohort. Preliminary data from inhibiting STAT3 signal transduction in CRC cell lines and PDOs suggested that CMS4/GMS2/stromal phenotype tumours observed a greater reduction in the percentage of viable cells post-treatment than CMS1/2/3 tumours and treatment was associated with increased apoptosis and decreased proliferation where investigated. Future work should expand on these data by testing Jakinibs on other Sanger organoid lines and new lines established from surplus Glasgow patient tissue.

IL6/JAK/STAT3 in the different TME components

Limitations of the current study include a lack of proper recapitulation of the TME during drug screening. Blockade of STAT3 activation in tumour cells alone is no sufficient evidence that pathway inhibition would be therapeutically useful due to the lack of immune and stromal populations in the model. This could be combated in the first instance using co-culture methods, and the effect of STAT3 inhibitors on fibroblasts and immune populations alongside the tumour cells should be investigated. Aside from coculture moving towards mouse models would provide a more comprehensive argument for moving Jakinibs into the clinic. The Sansom laboratory have developed a number of mouse models which recapitulate CMS4/GMS2 diseases such as AKPT and KPN models.

The effects of STAT3 blockade could be investigated in these models in relation to each specific compartment of the tumour microenvironment. Inhibiting the pathway shows promise as high tumoural expression of the pathway is associated with worse outcome, however if the same observation is not seen when immune cells/stromal cells are scored there may be paradoxical evidence for targeting the pathway. Investigating protein expression of the pathway amongst the inflammatory infiltrate and tumour-stroma represents another key piece of future work, which could be easily put into practise with all IHC staining already completed. These scores could then be assessed for association with CSS relative to subtype, MMR status and tumour subsite as well as assessed for association with tumour expression of each marker.

Multiplex IF staining could be used to assess vehicle and STAT3 inhibited tumours via staining for Ki67, Caspase 8, pSTAT3, CD68, CD3, aSMA and panCK to investigate the spatial relationship between key members of the TME.

Another important future study of interest could be to develop 3D tumouroid cocultures with cancer-associated stroma and/or macrophages. These models could be utilised to understand the main source of IL6 and investigate other interactions between primary tumour cells and stroma/macrophages. This represents challenges due to difficulty in deriving sufficient primary macrophages/stromal cells to grow in culture. If developed, these coculture systems could also be utilised to assess Jakinib affects in an even more representative model of disease than PDOs alone and could be investigated in the context of cRT combination therapy also.

Staining validation

Future work should include staining the Glasgow combined cohort and a validation cohort for the activation sites of the JAK family members; pJAK1, pJAK2, pJAK3 and pTYK2. This is necessary as we observed no association between total STAT3 and prognosis, however a prognostic effect was observed for expression of STAT3 activation sites in GMS2 patients. Additionally, there is evidence that the ratio of STAT1 to STAT3 could be an important prognostic measure in CRC, so staining the Glasgow combined array for STAT1 could represent another experiment of interest (243). It would also be of interest to perform STAT1 and STAT3 staining via IF to investigate co-expression levels on a cell-cell basis and relate data to survival outcome and clinicopathological characteristics.

Mechanistic work

In order to investigate the mechanisms of signal transduction, IL6R could be silenced from cell lines, and the drug screening performed with and without IL6 stimulation. Results could be evaluated using cell viability, apoptosis TUNEL assays, cell cycle assays, and western blots probing for pSTAT3 performed on cell lysates from each treatment group. Mouse models could be utilised to perform silencing of IL6/JAK/STAT3 in each compartment of the TME with downstream effects on tumour volume and number investigated. Much of the evidence in the literature suggests that STAT3 activation in the tumour, immune cells and in the surrounding stroma is associated with poor outcome and a promotion of the tumour.

JAK inhibitors and chemo-radiotherapy

Data from the literature and preliminary data from Mr Campbell Roxburgh's proof of concept rectal cancer study shows an increase of IL6/JAK/STAT3 signalling following treatment with radiotherapy (244). This suggests combination therapy with JAK inhibitors and radiotherapy should be investigated

in future work using PDOs and PDX animal models. In head and neck cancer lines, inhibiting STAT3 has been shown to enhance sensitivity to radiotherapy, with a greater effect observed under hypoxic conditions (245). The radiosensitizing effect of STAT3 inhibition has also been observed in oesophageal cancer models, whereby blocking STAT3 activation caused downregulation of hypoxia-inducing factor-1 and VEGF, and enhanced response to radiation (246). This could provide a link with the data from this thesis, with high tumour-stroma potentially associated with driving hypoxia through formation of a physical barrier around the tumour (247). JAK inhibition should also be investigated in combination with standard of care chemotherapy, and with other kinase inhibitors to obtain maximal pathway inhibition.

Conclusion

Glasgow microenvironment score and phenotypic subtype histological scoring methods were prognostic in the stage I-IV Glasgow combined cohort, high risk stage II-III TransSCOT clinical trial cohort and stage IV synchronous resection cohort. However, both histological subtyping methods did not stratify patient survival in the Australian TMA cohort, suggesting full sections are needed to accurately perform GMS and phenotypic subtyping. Additionally, in stage I-II screen detected CRC cohort, DM-CRC-TMA, GMS, and phenotypic subtype did not associate with clinicopathological characteristics indicating histological subtyping may not be useful for early-stage disease. We hypothesised that histological subtypes may be driven by different underlying aberrant cellular signalling pathways. As a first step towards investigating this the IL6/JAK/STAT3 pathway was investigated in the Glasgow combined cohort and an association with worse prognosis was observed for key pathway members. This effect was potentiated in GMS2 stromal tumours in the case of IL6, cytoplasmic IL6R, cytoplasmic JAK2, membranous JAK1/2 and pSTAT3^{tyr705}/pSTAT3^{ser727}. There were no distinct patterns of genomic or transcriptomic expression associated with protein pSTAT3^{tyr705} groups, however patients with high pSTAT3^{tyr705} protein expression were more likely to have mutations in the MAP2K gene. Inhibiting STAT3 activation by targeting JAK1/2 and JAK2/3 reduced cell viability in CMS4 classified cell lines (HCT116) and patient derived organoids, particularly those classified as CMS4/GMS2/stromal subtype. PDOs treated with Jakinibs also displayed reduced proliferation and increased apoptosis. Further research is required to validate these findings and to confirm that inhibiting IL6/JAK/STAT3 represents a promising therapeutic target for CRC patients with stromal-rich GMS2 tumours. If validated, this method represents a highly translatable method to transform CRC treatment, due to the simplicity of GMS histological subtyping and repurpose-able nature of JAK inhibitors.

References

1. Quirke P, Williams GT, Ectors N, Ensari A, Piard F, Nagtegaal I. The future of the TNM staging system in colorectal cancer: time for a debate? Lancet Oncol. 2007;8(7):651-7.
2. Guinney J, Dienstmann R, Wang X, de Reyniès A, Schlicker A, Soneson C, et al. The consensus molecular subtypes of colorectal cancer. Nature Medicine. 2015;21:1350.
3. Isella C, Brundu F, Bellomo SE, Galimi F, Zanella E, Porporato R, et al. Selective analysis of cancer-cell intrinsic transcriptional traits defines novel clinically relevant subtypes of colorectal cancer. Nat Commun. 2017;8:15107.
4. Dunne PD, Alderdice M, O'Reilly PG, Roddy AC, McCorry AMB, Richman S, et al. Cancer-cell intrinsic gene expression signatures overcome intratumoural heterogeneity bias in colorectal cancer patient classification. Nat Commun. 2017;8:15657.
5. Mesker WE, Junggeburt JM, Szuhai K, de Heer P, Morreau H, Tanke HJ, et al. The carcinoma-stromal ratio of colon carcinoma is an independent factor for survival compared to lymph node status and tumor stage. Cell Oncol. 2007;29(5):387-98.
6. Park JH, McMillan DC, Powell AG, Richards CH, Horgan PG, Edwards J, et al. Evaluation of a tumor microenvironment-based prognostic score in primary operable colorectal cancer. Clin Cancer Res. 2015;21(4):882-8.
7. Klintrup K, Makinen JM, Kauppila S, Vare PO, Melkko J, Tuominen H, et al. Inflammation and prognosis in colorectal cancer. Eur J Cancer. 2005;41(17):2645-54.
8. Antonia K. Roseweir DCM, Paul G. Horgan, Joanne Edwards. Colorectal cancer subtypes: Translation to routine clinical pathology. Cancer Treatment Reviews 2017;57:1-7.
9. Huynh J, Chand A, Gough D, Ernst M. Therapeutically exploiting STAT3 activity in cancer - using tissue repair as a road map. Nat Rev Cancer. 2019;19(2):82-96.
10. Fleischmann R, Kremer J, Cush J, Schulze-Koops H, Connell CA, Bradley JD, et al. Placebo-controlled trial of tofacitinib monotherapy in rheumatoid arthritis. N Engl J Med. 2012;367(6):495-507.
11. Berg KCG, Eide PW, Eilertsen IA, Johannessen B, Bruun J, Danielsen SA, et al. Multi-omics of 34 colorectal cancer cell lines - a resource for biomedical studies. Mol Cancer. 2017;16(1):116.
12. Arnold M, Sierra MS, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global patterns and trends in colorectal cancer incidence and mortality. Gut. 2017;66(4):683-91.
13. Soriano LC, Soriano-Gabarro M, Garcia Rodriguez LA. Trends in the contemporary incidence of colorectal cancer and patient characteristics in the United Kingdom: a population-based cohort study using The Health Improvement Network. BMC Cancer. 2018;18(1):402.
14. Clark GR, Anderson AS, Godfrey TG, Strachan JA, Fraser CG, Steele RJ. Variation in changes in the incidence of colorectal cancer by age and association with screening uptake: an observational study. BMJ Open. 2020;10(9):e037925.
15. Quyn AJ, Fraser CG, Stanners G, Carey FA, Carden C, Shaukat A, et al. Uptake trends in the Scottish Bowel Screening Programme and the influences of age, sex, and deprivation. J Med Screen. 2018;25(1):24-31.
16. Steele RJ, McClements PL, Libby G, Black R, Morton C, Birrell J, et al. Results from the first three rounds of the Scottish demonstration pilot of FOBT screening for colorectal cancer. Gut. 2009;58(4):530-5.
17. Haggar FA, Boushey RP. Colorectal cancer epidemiology: incidence, mortality, survival, and risk factors. Clin Colon Rectal Surg. 2009;22(4):191-7.
18. Johnson CM, Wei C, Ensor JE, Smolenski DJ, Amos CI, Levin B, et al. Meta-analyses of colorectal cancer risk factors. Cancer Causes Control. 2013;24(6):1207-22.
19. Galiatsatos P, Foulkes WD. Familial adenomatous polyposis. Am J Gastroenterol. 2006;101(2):385-98.

20. Moller P, Seppala T, Bernstein I, Holinski-Feder E, Sala P, Evans DG, et al. Cancer incidence and survival in Lynch syndrome patients receiving colonoscopic and gynaecological surveillance: first report from the prospective Lynch syndrome database. Gut. 2017;66(3):464-72.
21. Belli S, Aytac HO, Karagulle E, Yabanoglu H, Kayaselcuk F, Yildirim S. Outcomes of surgical treatment of primary signet ring cell carcinoma of the colon and rectum: 22 cases reviewed with literature. Int Surg. 2014;99(6):691-8.
22. Pino MS, Chung DC. The chromosomal instability pathway in colon cancer. Gastroenterology. 2010;138(6):2059-72.
23. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell. 1990;61(5):759-67.
24. Murphy KM, Zhang S, Geiger T, Hafez MJ, Bacher J, Berg KD, et al. Comparison of the microsatellite instability analysis system and the Bethesda panel for the determination of microsatellite instability in colorectal cancers. J Mol Diagn. 2006;8(3):305-11.
25. Richman S. Deficient mismatch repair: Read all about it (Review). Int J Oncol. 2015;47(4):1189-202.
26. Lin A, Zhang J, Luo P. Crosstalk Between the MSI Status and Tumor Microenvironment in Colorectal Cancer. Front Immunol. 2020;11:2039.
27. Nazemalhosseini Mojarad E, Kuppen PJ, Aghdaei HA, Zali MR. The CpG island methylator phenotype (CIMP) in colorectal cancer. Gastroenterol Hepatol Bed Bench. 2013;6(3):120-8.
28. Curtin K, Slattery ML, Samowitz WS. CpG island methylation in colorectal cancer: past, present and future. Patholog Res Int. 2011;2011:902674.
29. Barault L, Charon-Barra C, Jooste V, de la Vega MF, Martin L, Roignot P, et al. Hypermethylator phenotype in sporadic colon cancer: study on a population-based series of 582 cases. Cancer research. 2008;68(20):8541-6.
30. Carmichael JC, Mills S. Anatomy and Embryology of the Colon, Rectum, and Anus. In: Steele SR, Hull TL, Read TE, Saclarides TJ, Senagore AJ, Whitlow CB, editors. The ASCRS Textbook of Colon and Rectal Surgery. Cham: Springer International Publishing; 2016. p. 3-26.
31. Baran B, Mert Ozupek N, Yerli Tetik N, Acar E, Bekcioglu O, Baskin Y. Difference Between Left-Sided and Right-Sided Colorectal Cancer: A Focused Review of Literature. Gastroenterology Res. 2018;11(4):264-73.
32. Meguid RA, Slidell MB, Wolfgang CL, Chang DC, Ahuja N. Is there a difference in survival between right- versus left-sided colon cancers? Ann Surg Oncol. 2008;15(9):2388-94.
33. Patel M, McSorley ST, Park JH, Roxburgh CSD, Edwards J, Horgan PG, et al. The relationship between right-sided tumour location, tumour microenvironment, systemic inflammation, adjuvant therapy and survival in patients undergoing surgery for colon and rectal cancer. Br J Cancer. 2018;118(5):705-12.
34. Treanor D, Quirke P. Pathology of colorectal cancer. Clin Oncol (R Coll Radiol). 2007;19(10):769-76.
35. Dukes CE. The classification of cancer of the rectum. J Pathol Bacteriol. 1932;35(3):323-32.
36. Haq AI, Schneeweiss J, Kalsi V, Arya M. The Dukes staging system: a cornerstone in the clinical management of colorectal cancer. Lancet Oncol. 2009;10(11):1128-.
37. Delibegovic S. Introduction to Total Mesorectal Excision. Med Arch. 2017;71(6):434-8.
38. Iveson TJ, Kerr RS, Saunders MP, Cassidy J, Hollander NH, Tabernero J, et al. 3 versus 6 months of adjuvant oxaliplatin-fluoropyrimidine combination therapy for colorectal cancer (SCOT): an international, randomised, phase 3, non-inferiority trial. Lancet Oncol. 2018;19(4):562-78.
39. Conroy T, Bosset JF, Etienne PL, Rio E, Francois E, Mesgouez-Nebout N, et al. Neoadjuvant chemotherapy with FOLFIRINOX and preoperative chemoradiotherapy for patients with locally advanced rectal cancer (UNICANCER-PRODIGE 23): a multicentre, randomised, open-label, phase 3 trial. Lancet Oncol. 2021;22(5):702-15.

40. Roxburgh CSD, Strombom P, Lynn P, Cercek A, Gonen M, Smith JJ, et al. Changes in the multidisciplinary management of rectal cancer from 2009 to 2015 and associated improvements in short-term outcomes. Colorectal Dis. 2019;21(10):1140-50.
41. Reynolds NA, Wagstaff AJ. Cetuximab: in the treatment of metastatic colorectal cancer. Drugs. 2004;64(1):109-18; discussion 19-21.
42. Cao D, Zheng Y, Xu H, Ge W, Xu X. Bevacizumab improves survival in metastatic colorectal cancer patients with primary tumor resection: A meta-analysis. Sci Rep. 2019;9(1):20326.
43. Overman MJ, Ernstoff MS, Morse MA. Where We Stand With Immunotherapy in Colorectal Cancer: Deficient Mismatch Repair, Proficient Mismatch Repair, and Toxicity Management. Am Soc Clin Oncol Educ Book. 2018;38:239-47.
44. Goldman JW, Dvorkin M, Chen Y, Reinmuth N, Hotta K, Trukhin D, et al. Durvalumab, with or without tremelimumab, plus platinum-etoposide versus platinum-etoposide alone in first-line treatment of extensive-stage small-cell lung cancer (CASPIAN): updated results from a randomised, controlled, open-label, phase 3 trial. Lancet Oncol. 2021;22(1):51-65.
45. Tamberi S, Grassi E, Barbera MA, Corbelli J, Papiani G, Montroni I, et al. A phase II study of capecitabine plus concomitant radiation therapy followed by durvalumab (MEDI4736) as preoperative treatment in rectal cancer (PANDORA). Journal of Clinical Oncology. 2020;38(4).
46. Buikhuisen JY, Torang A, Medema JP. Exploring and modelling colon cancer inter-tumour heterogeneity: opportunities and challenges. Oncogenesis. 2020;9(7):66.
47. Alexander PG, Roseweir AK, Pennel KAF, van Wyk HC, Powell A, McMillan DC, et al. The Glasgow Microenvironment Score associates with prognosis and adjuvant chemotherapy response in colorectal cancer. Br J Cancer. 2021;124(4):786-96.
48. Roseweir AK, McMillan DC, Horgan PG, Edwards J. Colorectal cancer subtypes: Translation to routine clinical pathology. Cancer Treat Rev. 2017;57:1-7.
49. Roseweir AK, Park JH, Hoorn ST, Powell AG, Aherne S, Roxburgh CS, et al. Histological phenotypic subtypes predict recurrence risk and response to adjuvant chemotherapy in patients with stage III colorectal cancer. J Pathol Clin Res. 2020;6(4):283-96.
50. Galon J, Mlecnik B, Bindea G, Angell HK, Berger A, Lagorce C, et al. Towards the introduction of the 'Immunoscore' in the classification of malignant tumours. J Pathol. 2014;232(2):199-209.
51. Galon J, Lanzi A. Immunoscore and its introduction in clinical practice. Q J Nucl Med Mol Imaging. 2020;64(2):152-61.
52. Inthagard J, Edwards J, Roseweir AK. Immunotherapy: enhancing the efficacy of this promising therapeutic in multiple cancers. Clin Sci (Lond). 2019;133(2):181-93.
53. Dahlin AM, Henriksson ML, Van Guelpen B, Stenling R, Oberg A, Rutegard J, et al. Colorectal cancer prognosis depends on T-cell infiltration and molecular characteristics of the tumor. Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc. 2011;24(5):671-82.
54. Ye L, Zhang T, Kang Z, Guo G, Sun Y, Lin K, et al. Tumor-Infiltrating Immune Cells Act as a Marker for Prognosis in Colorectal Cancer. Front Immunol. 2019;10:2368.
55. Betts G, Jones E, Junaid S, El-Shanawany T, Scurr M, Mizen P, et al. Suppression of tumour-specific CD4(+) T cells by regulatory T cells is associated with progression of human colorectal cancer. Gut. 2012;61(8):1163-71.
56. Suzuki T, Hayman L, Kilbey A, Edwards J, Coffelt SB. Gut gammadelta T cells as guardians, disruptors, and instigators of cancer. Immunol Rev. 2020;298(1):198-217.
57. Yang C, Wei C, Wang S, Shi D, Zhang C, Lin X, et al. Elevated CD163(+)/CD68(+) Ratio at Tumor Invasive Front is Closely Associated with Aggressive Phenotype and Poor Prognosis in Colorectal Cancer. Int J Biol Sci. 2019;15(5):984-98.
58. Liou GY, Storz P. Reactive oxygen species in cancer. Free Radic Res. 2010;44(5):479-96.
59. Noy R, Pollard JW. Tumor-associated macrophages: from mechanisms to therapy. Immunity. 2014;41(1):49-61.

60. Zhu Q, Zhang X, Zhang L, Li W, Wu H, Yuan X, et al. The IL-6-STAT3 axis mediates a reciprocal crosstalk between cancer-derived mesenchymal stem cells and neutrophils to synergistically prompt gastric cancer progression. Cell Death Dis. 2014;5:e1295.
61. Sconocchia G, Eppenberger S, Spagnoli GC, Tornillo L, Droeser R, Caratelli S, et al. NK cells and T cells cooperate during the clinical course of colorectal cancer. Oncoimmunology. 2014;3(8):e952197.
62. Schurch CM, Bhate SS, Barlow GL, Phillips DJ, Noti L, Zlobec I, et al. Coordinated Cellular Neighborhoods Orchestrate Antitumoral Immunity at the Colorectal Cancer Invasive Front. Cell. 2020;183(3):838.
63. Zwing N, Failmezger H, Ooi CH, Hibar DP, Canamero M, Gomes B, et al. Analysis of Spatial Organization of Suppressive Myeloid Cells and Effector T Cells in Colorectal Cancer-A Potential Tool for Discovering Prognostic Biomarkers in Clinical Research. Front Immunol. 2020;11:550250.
64. O'Malley G, Heijltjes M, Houston AM, Rani S, Ritter T, Egan LJ, et al. Mesenchymal stromal cells (MSCs) and colorectal cancer - a troublesome twosome for the anti-tumour immune response? Oncotarget. 2016;7(37):60752-74.
65. Gieniec KA, Butler LM, Worthley DL, Woods SL. Cancer-associated fibroblasts-heroes or villains? Brit J Cancer. 2019;121(4):293-302.
66. O'Malley G, Treacy O, Lynch K, Naicker SD, Leonard NA, Lohan P, et al. Stromal Cell PD-L1 Inhibits CD8(+) T-cell Antitumor Immune Responses and Promotes Colon Cancer. Cancer Immunology Research. 2018;6(11):1426-41.
67. Pennel KAF, Park JH, McMillan DC, Roseweir AK, Edwards J. Signal interaction between the tumour and inflammatory cells in patients with gastrointestinal cancer: Implications for treatment. Cell Signal. 2019;54:81-90.
68. Yu H, Jove R. The stats of cancer - New molecular targets come of age. Nat Rev Cancer. 2004;4(2):97-105.
69. Johnson DE, O'Keefe RA, Grandis JR. Targeting the IL-6/JAK/STAT3 signalling axis in cancer. Nat Rev Clin Oncol. 2018;15(4):234-48.
70. Ghoreschi K, Laurence A, O'Shea JJ. Janus kinases in immune cell signaling. Immunol Rev. 2009;228(1):273-87.
71. Wen Z, Zhong Z, Darnell JE, Jr. Maximal activation of transcription by Stat1 and Stat3 requires both tyrosine and serine phosphorylation. Cell. 1995;82(2):241-50.
72. Guthrie GJ, Roxburgh CS, Richards CH, Horgan PG, McMillan DC. Circulating IL-6 concentrations link tumour necrosis and systemic and local inflammatory responses in patients undergoing resection for colorectal cancer. Br J Cancer. 2013;109(1):131-7.
73. Shiga K, Hara M, Nagasaki T, Sato T, Takahashi H, Sato M, et al. Preoperative Serum Interleukin-6 Is a Potential Prognostic Factor for Colorectal Cancer, including Stage II Patients. Gastroenterol Res Pract. 2016;2016:9701574.
74. Morikawa T, Baba Y, Yamauchi M, Kuchiba A, Nosho K, Shima K, et al. STAT3 expression, molecular features, inflammation patterns, and prognosis in a database of 724 colorectal cancers. Clin Cancer Res. 2011;17(6):1452-62.
75. Toyoshima Y, Kitamura H, Xiang H, Ohno Y, Homma S, Kawamura H, et al. IL6 Modulates the Immune Status of the Tumor Microenvironment to Facilitate Metastatic Colonization of Colorectal Cancer Cells. Cancer Immunol Res. 2019;7(12):1944-57.
76. Gujam FJ, McMillan DC, Edwards J. The relationship between total and phosphorylated STAT1 and STAT3 tumour cell expression, components of tumour microenvironment and survival in patients with invasive ductal breast cancer. Oncotarget. 2016;7(47):77607-21.
77. Tam L, McGlynn LM, Traynor P, Mukherjee R, Bartlett JM, Edwards J. Expression levels of the JAK/STAT pathway in the transition from hormone-sensitive to hormone-refractory prostate cancer. Br J Cancer. 2007;97(3):378-83.

78. Wang SW, Hu J, Guo QH, Zhao Y, Cheng JJ, Zhang DS, et al. AZD1480, a JAK inhibitor, inhibits cell growth and survival of colorectal cancer via modulating the JAK2/STAT3 signaling pathway. Oncol Rep. 2014;32(5):1991-8.
79. Chang Q, Bournazou E, Sansone P, Berishaj M, Gao SP, Daly L, et al. The IL-6/JAK/Stat3 feed-forward loop drives tumorigenesis and metastasis. Neoplasia. 2013;15(7):848-62.
80. Sun Q, Shang YK, Sun FK, Dong XW, Niu J, Li FN. Interleukin-6 Promotes Epithelial-Mesenchymal Transition and Cell Invasion through Integrin beta 6 Upregulation in Colorectal Cancer. Oxidative Medicine and Cellular Longevity. 2020;2020.
81. Nagasaki T, Hara M, Nakanishi H, Takahashi H, Sato M, Takeyama H. Interleukin-6 released by colon cancer-associated fibroblasts is critical for tumour angiogenesis: anti-interleukin-6 receptor antibody suppressed angiogenesis and inhibited tumour-stroma interaction. Br J Cancer. 2014;110(2):469-78.
82. Chen L, Wang S, Wang Y, Zhang W, Ma K, Hu C, et al. IL-6 influences the polarization of macrophages and the formation and growth of colorectal tumor. Oncotarget. 2018;9(25):17443-54.
83. Park SJ, Nakagawa T, Kitamura H, Atsumi T, Kamon H, Sawa S, et al. IL-6 regulates in vivo dendritic cell differentiation through STAT3 activation. J Immunol. 2004;173(6):3844-54.
84. Hagenstein J, Melderis S, Nosko A, Warkotsch MT, Richter JV, Ramcke T, et al. A Novel Role for IL-6 Receptor Classic Signaling: Induction of RORgammat(+)Foxp3(+) Tregs with Enhanced Suppressive Capacity. J Am Soc Nephrol. 2019;30(8):1439-53.
85. Toyoshima Y, Kitamura H, Xiang HH, Ohno Y, Homma S, Kawamura H, et al. IL6 Modulates the Immune Status of the Tumor Microenvironment to Facilitate Metastatic Colonization of Colorectal Cancer Cells. Cancer Immunology Research. 2019;7(12):1944-57.
86. Nagasaki T, Hara M, Nakanishi H, Takahashi H, Sato M, Takeyama H. Interleukin-6 released by colon cancer-associated fibroblasts is critical for tumour angiogenesis: anti-interleukin-6 receptor antibody suppressed angiogenesis and inhibited tumour-stroma interaction. Brit J Cancer. 2014;110(2):469-78.
87. Patel M, Horgan PG, McMillan DC, Edwards J. NF-kappaB pathways in the development and progression of colorectal cancer. Transl Res. 2018;197:43-56.
88. Blando JM, Carbajal S, Abel E, Beltran L, Conti C, Fischer S, et al. Cooperation between Stat3 and Akt signaling leads to prostate tumor development in transgenic mice. Neoplasia. 2011;13(3):254-65.
89. Rybak AP, Bristow RG, Kapoor A. Prostate cancer stem cells: deciphering the origins and pathways involved in prostate tumorigenesis and aggression. Oncotarget. 2015;6(4):1900-19.
90. Yazici Y, Curtis JR, Ince A, Baraf H, Malamet RL, Teng LL, et al. Efficacy of tocilizumab in patients with moderate to severe active rheumatoid arthritis and a previous inadequate response to disease-modifying antirheumatic drugs: the ROSE study. Ann Rheum Dis. 2012;71(2):198-205.
91. Choy EH, De Benedetti F, Takeuchi T, Hashizume M, John MR, Kishimoto T. Translating IL-6 biology into effective treatments. Nat Rev Rheumatol. 2020;16(6):335-45.
92. Lamb YN, Deeks ED. Sarilumab: A Review in Moderate to Severe Rheumatoid Arthritis. Drugs. 2018;78(9):929-40.
93. Angevin E, Tabernero J, Elez E, Cohen SJ, Bahleda R, van Laethem JL, et al. A phase I/II, multiple-dose, dose-escalation study of siltuximab, an anti-interleukin-6 monoclonal antibody, in patients with advanced solid tumors. Clin Cancer Res. 2014;20(8):2192-204.
94. Keystone EC, Taylor PC, Drescher E, Schlichting DE, Beattie SD, Berclaz PY, et al. Safety and efficacy of baricitinib at 24 weeks in patients with rheumatoid arthritis who have had an inadequate response to methotrexate. Ann Rheum Dis. 2015;74(2):333-40.
95. Thilakasiri PS, Dmello RS, Nero TL, Parker MW, Ernst M, Chand AL. Repurposing of drugs as STAT3 inhibitors for cancer therapy. Semin Cancer Biol. 2021;68:31-46.
96. Sen M, Thomas SM, Kim S, Yeh JI, Ferris RL, Johnson JT, et al. First-in-human trial of a STAT3 decoy oligonucleotide in head and neck tumors: implications for cancer therapy. Cancer Discov. 2012;2(8):694-705.

97. Jonker DJ, Nott L, Yoshino T, Gill S, Shapiro J, Ohtsu A, et al. Napabucasin versus placebo in refractory advanced colorectal cancer: a randomised phase 3 trial. Lancet Gastroenterol Hepatol. 2018;3(4):263-70.
98. Park JH, Richards CH, McMillan DC, Horgan PG, Roxburgh CSD. The relationship between tumour stroma percentage, the tumour microenvironment and survival in patients with primary operable colorectal cancer. Ann Oncol. 2014;25(3):644-51.
99. Bankhead P, Loughrey MB, Fernandez JA, Dombrowski Y, Mcart DG, Dunne PD, et al. QuPath: Open source software for digital pathology image analysis. Sci Rep-Uk. 2017;7.
100. Broutier L, Andersson-Rolf A, Hindley CJ, Boj SF, Clevers H, Koo BK, et al. Culture and establishment of self-renewing human and mouse adult liver and pancreas 3D organoids and their genetic manipulation. Nat Protoc. 2016;11(9):1724-43.
101. Nagtegaal ID, Quirke P, Schmoll HJ. Has the new TNM classification for colorectal cancer improved care? Nat Rev Clin Oncol. 2011;9(2):119-23.
102. Quirke P, Williams GT, Ectors N, Ensari A, Piard F, Nagtegaal I. The future of the TNM staging system in colorectal cancer: time for a debate? Lancet Oncol. 2007;8(7):651-7.
103. Park JH, McMillan DC, Powell AG, Richards C, Horgan PG, Edwards J, et al. Evaluation of a Tumor Microenvironment-Based Prognostic Score in Primary Operable Colorectal Cancer. Clin Cancer Res. 2015;21(4):882-8.
104. Roseweir AK, McMillan DC, Horgan PG, Edwards J. Colorectal cancer subtypes: Translation to routine clinical pathology. Cancer Treatment Reviews. 2017;57:1-7.
105. Alexander PG, Roseweir AK, Pennel KAF, van Wyk HC, Powell A, McMillan DC, et al. The Glasgow Microenvironment Score associates with prognosis and adjuvant chemotherapy response in colorectal cancer. Br J Cancer. 2020.
106. Park JH, van Wyk H, Roxburgh CSD, Horgan PG, Edwards J, McMillan DC. Tumour invasiveness, the local and systemic environment and the basis of staging systems in colorectal cancer. Br J Cancer. 2017;116(11):1444-50.
107. Iveson TJ, Kerr RS, Saunders MP, Cassidy J, Hollander NH, Tabernero J, et al. 3 versus 6 months of adjuvant oxaliplatin-fluoropyrimidine combination therapy for colorectal cancer (SCOT): an international, randomised, phase 3, non-inferiority trial. Lancet Oncol. 2018;19(4):562-78.
108. Park JH, van Wyk H, McMillan DC, Edwards J, Orange C, Horgan PG, et al. Preoperative, biopsy-based assessment of the tumour microenvironment in patients with primary operable colorectal cancer. Journal of Pathology Clinical Research. 2020;6(1):30-9.
109. Trinh A, Trumpi K, De Sousa EMF, Wang X, de Jong JH, Fessler E, et al. Practical and Robust Identification of Molecular Subtypes in Colorectal Cancer by Immunohistochemistry. Clin Cancer Res. 2017;23(2):387-98.
110. Hansen TF, Kjaer-Frifeldt S, Lindebjerg J, Rafaelsen SR, Jensen LH, Jakobsen A, et al. Tumor-stroma ratio predicts recurrence in patients with colon cancer treated with neoadjuvant chemotherapy. Acta Oncol. 2018;57(4):528-33.
111. Becker C, Fantini MC, Wirtz S, Nikolaev A, Lehr HA, Galle PR, et al. IL-6 signaling promotes tumor growth in colorectal cancer. Cell Cycle. 2005;4(2):217-20.
112. Waldner MJ, Foersch S, Neurath MF. Interleukin-6--a key regulator of colorectal cancer development. Int J Biol Sci. 2012;8(9):1248-53.
113. Grivennikov S, Karin E, Terzic J, Mucida D, Yu GY, Vallabhapurapu S, et al. IL-6 and Stat3 Are Required for Survival of Intestinal Epithelial Cells and Development of Colitis-Associated Cancer (vol 15, pg 103, 2009). Cancer Cell. 2009;15(3):241-.
114. Waugh DJ, Wilson C. The interleukin-8 pathway in cancer. Clin Cancer Res. 2008;14(21):6735-41.
115. Squarize CH, Castilho RM, Sriuranpong V, Pinto DS, Jr., Gutkind JS. Molecular cross-talk between the NFkappaB and STAT3 signaling pathways in head and neck squamous cell carcinoma. Neoplasia. 2006;8(9):733-46.

116. Bankhead P, Loughrey MB, Fernandez JA, Dombrowski Y, McArt DG, Dunne PD, et al. QuPath: Open source software for digital pathology image analysis. Sci Rep. 2017;7(1):16878.
117. Zeng J, Tang ZH, Liu S, Guo SS. Clinicopathological significance of overexpression of interleukin-6 in colorectal cancer. World J Gastroenterol. 2017;23(10):1780-6.
118. Holmer R, Watzig GH, Tiwari S, Rose-John S, Kalthoff H. Interleukin-6 trans-signaling increases the expression of carcinoembryonic antigen-related cell adhesion molecules 5 and 6 in colorectal cancer cells. BMC Cancer. 2015;15:975.
119. Zhang R, Qi F, Zhao F, Li G, Shao S, Zhang X, et al. Cancer-associated fibroblasts enhance tumor-associated macrophages enrichment and suppress NK cells function in colorectal cancer. Cell Death Dis. 2019;10(4):273.
120. Ebbing EA, van der Zalm AP, Steins A, Creemers A, Hermsen S, Rentenaar R, et al. Stromal-derived interleukin 6 drives epithelial-to-mesenchymal transition and therapy resistance in esophageal adenocarcinoma. Proc Natl Acad Sci U S A. 2019;116(6):2237-42.
121. Karpinski P, Rossowska J, Sasiadek MM. Immunological landscape of consensus clusters in colorectal cancer. Oncotarget. 2017;8(62):105299-311.
122. Pollard JW. Tumour-educated macrophages promote tumour progression and metastasis. Nat Rev Cancer. 2004;4(1):71-8.
123. Guo Y, Zang Y, Lv L, Cai F, Qian T, Zhang G, et al. IL8 promotes proliferation and inhibition of apoptosis via STAT3/AKT/NFkappaB pathway in prostate cancer. Mol Med Rep. 2017;16(6):9035-42.
124. Rubie C, Frick VO, Pfeil S, Wagner M, Kollmar O, Kopp B, et al. Correlation of IL-8 with induction, progression and metastatic potential of colorectal cancer. World J Gastroenterol. 2007;13(37):4996-5002.
125. Qureshy Z, Johnson DE, Grandis JR. Targeting the JAK/STAT pathway in solid tumors. J Cancer Metastasis Treat. 2020;6.
126. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell. 2000;100(1):57-70.
127. Wang SW, Sun YM. The IL-6/JAK/STAT3 pathway: potential therapeutic strategies in treating colorectal cancer (Review). Int J Oncol. 2014;44(4):1032-40.
128. Bousoik E, Montazeri Aliabadi H. "Do We Know Jack" About JAK? A Closer Look at JAK/STAT Signaling Pathway. Front Oncol. 2018;8:287.
129. Slattery ML, Lundgreen A, Kadlubar SA, Bondurant KL, Wolff RK. JAK/STAT/SOCS-signaling pathway and colon and rectal cancer. Mol Carcinog. 2013;52(2):155-66.
130. Park SY, Lee CJ, Choi JH, Kim JH, Kim JW, Kim JY, et al. The JAK2/STAT3/CCND2 Axis promotes colorectal Cancer stem cell persistence and radioresistance. J Exp Clin Cancer Res. 2019;38(1):399.
131. Ide H, Nakagawa T, Terado Y, Kamiyama Y, Muto S, Horie S. Tyk2 expression and its signaling enhances the invasiveness of prostate cancer cells. Biochem Biophys Res Commun. 2008;369(2):292-6.
132. Woss K, Simonovic N, Strobl B, Macho-Maschler S, Muller M. TYK2: An Upstream Kinase of STATs in Cancer. Cancers (Basel). 2019;11(11).
133. Walters DK, Mercher T, Gu TL, O'Hare T, Tyner JW, Loriaux M, et al. Activating alleles of JAK3 in acute megakaryoblastic leukemia. Cancer Cell. 2006;10(1):65-75.
134. Lin Q, Lai R, Chirieac LR, Li C, Thomazy VA, Grammatikakis I, et al. Constitutive activation of JAK3/STAT3 in colon carcinoma tumors and cell lines: inhibition of JAK3/STAT3 signaling induces apoptosis and cell cycle arrest of colon carcinoma cells. Am J Pathol. 2005;167(4):969-80.
135. Chen B, Lai J, Dai D, Chen R, Li X, Liao N. JAK1 as a prognostic marker and its correlation with immune infiltrates in breast cancer. Aging (Albany NY). 2019;11(23):11124-35.
136. Liu D, Huang Y, Zhang L, Liang DN, Li L. Activation of Janus kinase 1 confers poor prognosis in patients with non-small cell lung cancer. Oncol Lett. 2017;14(4):3959-66.
137. Yu DL, Zhang T, Wu K, Li Y, Wang J, Chen J, et al. MicroRNA-448 suppresses metastasis of pancreatic ductal adenocarcinoma through targeting JAK1/STAT3 pathway. Oncol Rep. 2017;38(2):1075-82.

138. Owen KL, Brockwell NK, Parker BS. JAK-STAT Signaling: A Double-Edged Sword of Immune Regulation and Cancer Progression. Cancers (Basel). 2019;11(12).
139. Corcoran RB, Contino G, Deshpande V, Tzatsos A, Conrad C, Benes CH, et al. STAT3 plays a critical role in KRAS-induced pancreatic tumorigenesis. Cancer research. 2011;71(14):5020-9.
140. Mohanty SK, Yagiz K, Pradhan D, Luthringer DJ, Amin MB, Alkan S, et al. STAT3 and STAT5A are potential therapeutic targets in castration-resistant prostate cancer. Oncotarget. 2017;8(49):85997-6010.
141. Kim DY, Cha ST, Ahn DH, Kang HY, Kwon CI, Ko KH, et al. STAT3 expression in gastric cancer indicates a poor prognosis. J Gastroenterol Hepatol. 2009;24(4):646-51.
142. Kamran MZ, Patil P, Gude RP. Role of STAT3 in cancer metastasis and translational advances. Biomed Res Int. 2013;2013:421821.
143. Kusaba T, Nakayama T, Yamazumi K, Yakata Y, Yoshizaki A, Inoue K, et al. Activation of STAT3 is a marker of poor prognosis in human colorectal cancer. Oncol Rep. 2006;15(6):1445-51.
144. Sawayama H, Miyamoto Y, Ogawa K, Yoshida N, Baba H. Investigation of colorectal cancer in accordance with consensus molecular subtype classification. Ann Gastroenterol Surg. 2020;4(5):528-39.
145. Corvinus FM, Orth C, Moriggl R, Tsareva SA, Wagner S, Pfitzner EB, et al. Persistent STAT3 activation in colon cancer is associated with enhanced cell proliferation and tumor growth. Neoplasia. 2005;7(6):545-55.
146. Lomada D, Jain M, Bolner M, Reeh KA, Kang R, Reddy MC, et al. Stat3 Signaling Promotes Survival And Maintenance Of Medullary Thymic Epithelial Cells. PLoS Genet. 2016;12(1):e1005777.
147. Yang H, Yamazaki T, Pietrocola F, Zhou H, Zitvogel L, Ma Y, et al. STAT3 Inhibition Enhances the Therapeutic Efficacy of Immunogenic Chemotherapy by Stimulating Type 1 Interferon Production by Cancer Cells. Cancer research. 2015;75(18):3812-22.
148. Park JH, van Wyk H, McMillan DC, Quinn J, Clark J, Roxburgh CS, et al. Signal Transduction and Activator of Transcription-3 (STAT3) in Patients with Colorectal Cancer: Associations with the Phenotypic Features of the Tumor and Host. Clin Cancer Res. 2017;23(7):1698-709.
149. Calon A, Espinet E, Palomo-Ponce S, Tauriello DV, Iglesias M, Cespedes MV, et al. Dependency of colorectal cancer on a TGF-beta-driven program in stromal cells for metastasis initiation. Cancer Cell. 2012;22(5):571-84.
150. You L, Wang Z, Li H, Shou J, Jing Z, Xie J, et al. The role of STAT3 in autophagy. Autophagy. 2015;11(5):729-39.
151. Dougados M, van der Heijde D, Chen YC, Greenwald M, Drescher E, Liu J, et al. Baricitinib in patients with inadequate response or intolerance to conventional synthetic DMARDs: results from the RA-BUILD study. Ann Rheum Dis. 2017;76(1):88-95.
152. Talpaz M, Kiladjian JJ. Fedratinib, a newly approved treatment for patients with myeloproliferative neoplasm-associated myelofibrosis. Leukemia. 2021;35(1):1-17.
153. Bieber T, Simpson EL, Silverberg JI, Thaci D, Paul C, Pink AE, et al. Abrocitinib versus Placebo or Dupilumab for Atopic Dermatitis. N Engl J Med. 2021;384(12):1101-12.
154. Verstovsek S, Odenike O, Singer JW, Granston T, Al-Fayoumi S, Deeg HJ. Phase 1/2 study of pacritinib, a next generation JAK2/FLT3 inhibitor, in myelofibrosis or other myeloid malignancies. J Hematol Oncol. 2016;9(1):137.
155. Jasperson KW, Tuohy TM, Neklason DW, Burt RW. Hereditary and familial colon cancer. Gastroenterology. 2010;138(6):2044-58.
156. Steinke V, Engel C, Buttner R, Schackert HK, Schmiegel WH, Propping P. Hereditary Nonpolyposis Colorectal Cancer (HNPCC)/Lynch Syndrome. Dtsch Arztebl Int. 2013;110(3):32-U28.
157. Armaghany T, Wilson JD, Chu Q, Mills G. Genetic alterations in colorectal cancer. Gastrointest Cancer Res. 2012;5(1):19-27.
158. Fodde R, Smits R, Clevers H. APC, signal transduction and genetic instability in colorectal cancer. Nat Rev Cancer. 2001;1(1):55-67.

159. De Filippo C, Luceri C, Caderni G, Pacini M, Messerini L, Biggeri A, et al. Mutations of the APC gene in human sporadic colorectal cancers. Scand J Gastroenterol. 2002;37(9):1048-53.
160. Conlin A, Smith G, Carey FA, Wolf CR, Steele RJ. The prognostic significance of K-ras, p53, and APC mutations in colorectal carcinoma. Gut. 2005;54(9):1283-6.
161. Nguyen HT, Duong HQ. The molecular characteristics of colorectal cancer: Implications for diagnosis and therapy. Oncol Lett. 2018;16(1):9-18.
162. Park JH, Mcmillan DC, Horgan PG, Roxburgh CSD. The clinical utility of a tumour microenvironment-based histopathological score in patients with primary operable colorectal cancer. Journal of Clinical Oncology. 2015;33(3).
163. Molinari F, Signoroni S, Lampis A, Bertan C, Perrone F, Sala P, et al. BRAF mutation analysis is a valid tool to implement in Lynch syndrome diagnosis in patients classified according to the Bethesda guidelines. Tumori. 2014;100(3):315-20.
164. Lieu CH, Golemis EA, Serebriiskii IG, Newberg J, Hemmerich A, Connelly C, et al. Comprehensive Genomic Landscapes in Early and Later Onset Colorectal Cancer. Clinical Cancer Research. 2019;25(19):5852-8.
165. Prendergast GC, Mondal A, Dey S, Laury-Kleintop LD, Muller AJ. Inflammatory Reprogramming with IDO1 Inhibitors: Turning Immunologically Unresponsive 'Cold' Tumors 'Hot'. Trends Cancer. 2018;4(1):38-58.
166. Khurana N, Sikka SC. Interplay Between SOX9, Wnt/beta-Catenin and Androgen Receptor Signaling in Castration-Resistant Prostate Cancer. Int J Mol Sci. 2019;20(9).
167. Pan J, Silva TC, Gull N, Yang Q, Plummer JT, Chen S, et al. Lineage-Specific Epigenomic and Genomic Activation of Oncogene HNF4A Promotes Gastrointestinal Adenocarcinomas. Cancer research. 2020;80(13):2722-36.
168. Grasso CS, Giannakis M, Wells DK, Hamada T, Mu XJ, Quist M, et al. Genetic Mechanisms of Immune Evasion in Colorectal Cancer. Cancer Discov. 2018;8(6):730-49.
169. Stein Y, Aloni-Grinstein R, Rotter V. Mutant p53-a potential player in shaping the tumor-stroma crosstalk. J Mol Cell Biol. 2019;11(7):600-4.
170. Nash GM, Gimbel M, Shia J, Nathanson DR, Ndubuisi MI, Zeng ZS, et al. KRAS Mutation Correlates With Accelerated Metastatic Progression in Patients With Colorectal Liver Metastases. Ann Surg Oncol. 2010;17(2):572-8.
171. Dienstmann R, Vermeulen L, Guinney J, Kopetz S, Tejpar S, Tabernero J. Consensus molecular subtypes and the evolution of precision medicine in colorectal cancer. Nat Rev Cancer. 2017;17(4):268.
172. Chen YC, Lee TH, Tzeng SL. Reduced DAXX Expression Is Associated with Reduced CD24 Expression in Colorectal Cancer. Cells. 2019;8(10).
173. Stein Y, Aloni-Grinstein R, Rotter V. Mutant p53-a potential player in shaping the tumor-stroma crosstalk. J Mol Cell Biol. 2019;11(7):600-4.
174. Albacker LA, Wu J, Smith P, Warmuth M, Stephens PJ, Zhu P, et al. Loss of function JAK1 mutations occur at high frequency in cancers with microsatellite instability and are suggestive of immune evasion. Plos One. 2017;12(11).
175. Li SYD, Ma M, Li H, Waluszko A, Sidorenko T, Schadt EE, et al. Cancer gene profiling in non-small cell lung cancers reveals activating mutations in JAK2 and JAK3 with therapeutic implications. Genome Med. 2017;9.
176. Herreros-Villanueva M, Garcia-Giron C, Er TK. No evidence for JAK2 V617F mutation in colorectal cancer. Brit J Biomed Sci. 2010;67(4):220-2.
177. Peyser ND, Pendleton K, Gooding WE, Lui VWY, Johnson DE, Grandis JR. Genomic and Transcriptomic Alterations Associated with STAT3 Activation in Head and Neck Cancer. Plos One. 2016;11(11).
178. Muzny DM, Bainbridge MN, Chang K, Dinh HH, Drummond JA, Fowler G, et al. Comprehensive molecular characterization of human colon and rectal cancer. Nature. 2012;487(7407):330-7.

179. Nagathihalli NS, Castellanos JA, Lamichhane P, Messaggio F, Shi C, Dai X, et al. Inverse Correlation of STAT3 and MEK Signaling Mediates Resistance to RAS Pathway Inhibition in Pancreatic Cancer. Cancer research. 2018;78(21):6235-46.
180. Gkouveris I, Nikitakis N, Karanikou M, Rassidakis G, Sklavounou A. Erk1/2 activation and modulation of STAT3 signaling in oral cancer. Oncol Rep. 2014;32(5):2175-82.
181. Meng A, Zhang X, Shi Y. Role of p38 MAPK and STAT3 in lipopolysaccharide-stimulated mouse alveolar macrophages. Exp Ther Med. 2014;8(6):1772-6.
182. Zhao K, Lu Y, Chen Y, Cheng J, Zhang W. Dual Inhibition of MAPK and JAK2/STAT3 Pathways Is Critical for the Treatment of BRAF Mutant Melanoma. Mol Ther Oncolytics. 2020;18:100-8.
183. Fragoso MA, Patel AK, Nakamura RE, Yi H, Surapaneni K, Hackam AS. The Wnt/beta-catenin pathway cross-talks with STAT3 signaling to regulate survival of retinal pigment epithelium cells. PLoS One. 2012;7(10):e46892.
184. Eto T, Miyake K, Nosho K, Ohmuraya M, Imamura Y, Arima K, et al. Impact of loss-of-function mutations at the RNF43 locus on colorectal cancer development and progression. Journal of Pathology. 2018;245(4):445-55.
185. Koo BK, van Es JH, van den Born M, Clevers H. Porcupine inhibitor suppresses paracrine Wnt-driven growth of Rnf43;Znrf3-mutant neoplasia. P Natl Acad Sci USA. 2015;112(24):7548-50.
186. Xu L, Wang R, Ziegelbauer J, Wu WW, Shen RF, Juhl H, et al. Transcriptome analysis of human colorectal cancer biopsies reveals extensive expression correlations among genes related to cell proliferation, lipid metabolism, immune response and collagen catabolism. Oncotarget. 2017;8(43):74703-19.
187. Trejo CL, Babic M, Imler E, Gonzalez M, Bibikov SI, Shepard PJ, et al. Extraction-free whole transcriptome gene expression analysis of FFPE sections and histology-directed subareas of tissue. PLoS One. 2019;14(2):e0212031.
188. Pan P, Oshima K, Huang YW, Agle KA, Drobyski WR, Chen X, et al. Loss of FFAR2 promotes colon cancer by epigenetic dysregulation of inflammation suppressors. International Journal of Cancer. 2018;143(4):886-96.
189. Rebe C, Ghiringhelli F. Interleukin-1 beta and Cancer. Cancers. 2020;12(7).
190. Jakobsen MK, Traynor S, Staehr M, Duijf PG, Nielsen AY, Terp MG, et al. The Cancer/Testis Antigen Gene VCX2 Is Rarely Expressed in Malignancies but Can Be Epigenetically Activated Using DNA Methyltransferase and Histone Deacetylase Inhibitors. Frontiers in Oncology. 2021;10.
191. Tilley SK, Kim WY, Fry RC. Analysis of bladder cancer tumor CpG methylation and gene expression within The Cancer Genome Atlas identifies GRIA1 as a prognostic biomarker for basal-like bladder cancer. American Journal of Cancer Research. 2017;7(9):1850-62.
192. Endo M, Kanda M, Sawaki K, Shimizu D, Tanaka C, Kobayashi D, et al. Tissue Expression of Melanoma-associated Antigen A6 and Clinical Characteristics of Gastric Cancer. Anticancer Res. 2019;39(11):5903-10.
193. Wei J, Ma L, Lai YH, Zhang R, Li H, Li C, et al. Bazedoxifene as a novel GP130 inhibitor for Colon Cancer therapy. J Exp Clin Cancer Res. 2019;38(1):63.
194. Trinh A, Ladrach C, Dawson HE, ten Hoorn S, Kuppen PJK, Reimers MS, et al. Tumour budding is associated with the mesenchymal colon cancer subtype and RAS/RAF mutations: a study of 1320 colorectal cancers with Consensus Molecular Subgroup (CMS) data. Brit J Cancer. 2018;119(10):1244-51.
195. Kevans D, Wang LM, Sheahan K, Hyland J, O'Donoghue D, Mulcahy H, et al. Epithelial-Mesenchymal Transition (EMT) Protein Expression in a Cohort of Stage II Colorectal Cancer Patients With Characterized Tumor Budding and Mismatch Repair Protein Status. International Journal of Surgical Pathology. 2011;19(6):751-60.
196. Wang Y, Kartasalo K, Valkonen M, Larsson C, Ruusuvuori P, Hartman J, et al. Predicting molecular phenotypes from histopathology images: a transcriptome-wide expression-morphology analysis in breast cancer. ArXiv. 2020;abs/2009.08917.

197. Schmauch B, Romagnoni A, Pronier E, Saillard C, Maille P, Calderaro J, et al. A deep learning model to predict RNA-Seq expression of tumours from whole slide images. Nat Commun. 2020;11(1):3877.
198. Yu KH, Berry GJ, Rubin DL, Re C, Altman RB, Snyder M. Association of Omics Features with Histopathology Patterns in Lung Adenocarcinoma. Cell Syst. 2017;5(6):620-+.
199. Kather JN, Heij LR, Grabsch HI, Loeffler C, Echle A, Muti HS, et al. Pan-cancer image-based detection of clinically actionable genetic alterations (vol 1, pg 789, 2020). Nat Cancer. 2020;1(11):1129-.
200. Tang W, Zhou M, Dorsey TH, Prieto DA, Wang XW, Ruppin E, et al. Integrated proteotranscriptomics of breast cancer reveals globally increased protein-mRNA concordance associated with subtypes and survival. Genome Med. 2018;10(1):94.
201. Lassmann S, Schuster I, Walch A, Gobel H, Jutting U, Makowiec F, et al. STAT3 mRNA and protein expression in colorectal cancer: effects on STAT3-inducible targets linked to cell survival and proliferation. J Clin Pathol. 2007;60(2):173-9.
202. Sato T, Stange DE, Ferrante M, Vries RG, Van Es JH, Van den Brink S, et al. Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. Gastroenterology. 2011;141(5):1762-72.
203. Lam C, Ferguson ID, Mariano MC, Lin YT, Murnane M, Liu H, et al. Repurposing tofacitinib as an anti-myeloma therapeutic to reverse growth-promoting effects of the bone marrow microenvironment. Haematologica. 2018;103(7):1218-28.
204. Verstovsek S, Mesa RA, Gotlib J, Levy RS, Gupta V, DiPersio JF, et al. A double-blind, placebo-controlled trial of ruxolitinib for myelofibrosis. N Engl J Med. 2012;366(9):799-807.
205. Linnekamp JF, Hooff SRV, Prasetyanti PR, Kandimalla R, Buikhuisen JY, Fessler E, et al. Consensus molecular subtypes of colorectal cancer are recapitulated in in vitro and in vivo models. Cell Death Differ. 2018;25(3):616-33.
206. Klein JD, Sano D, Sen M, Myers JN, Grandis JR, Kim S. STAT3 oligonucleotide inhibits tumor angiogenesis in preclinical models of squamous cell carcinoma. PLoS One. 2014;9(1):e81819.
207. Han ES, Wen W, Dellinger TH, Wu J, Lu SA, Jove R, et al. Ruxolitinib synergistically enhances the anti-tumor activity of paclitaxel in human ovarian cancer. Oncotarget. 2018;9(36):24304-19.
208. Xiong H, Zhang ZG, Tian XQ, Sun DF, Liang QC, Zhang YJ, et al. Inhibition of JAK1, 2/STAT3 signaling induces apoptosis, cell cycle arrest, and reduces tumor cell invasion in colorectal cancer cells. Neoplasia. 2008;10(3):287-97.
209. Ganesh K, Wu C, O'Rourke KP, Szeglin BC, Zheng Y, Sauve CG, et al. A rectal cancer organoid platform to study individual responses to chemoradiation. Nat Med. 2019;25(10):1607-14.
210. Becker C, Fantini MC, Schramm C, Lehr HA, Wirtz S, Nikolaev A, et al. TGF-beta suppresses tumor progression in colon cancer by inhibition of IL-6 trans-signaling. Immunity. 2004;21(4):491-501.
211. Takeda K, Noguchi K, Shi W, Tanaka T, Matsumoto M, Yoshida N, et al. Targeted disruption of the mouse Stat3 gene leads to early embryonic lethality. Proc Natl Acad Sci U S A. 1997;94(8):3801-4.
212. Lin L, Liu A, Peng Z, Lin HJ, Li PK, Li C, et al. STAT3 is necessary for proliferation and survival in colon cancer-initiating cells. Cancer research. 2011;71(23):7226-37.
213. Okita A, Takahashi S, Ouchi K, Inoue M, Watanabe M, Endo M, et al. Consensus molecular subtypes classification of colorectal cancer as a predictive factor for chemotherapeutic efficacy against metastatic colorectal cancer. Oncotarget. 2018;9(27):18698-711.
214. Pogue-Geile KL, Andre T, Song N, Lipchik C, Wang Y, Kim RS, et al. Association of colon cancer (CC) molecular signatures with prognosis and oxaliplatin prediction benefit in the MOSAIC Trial (Multicenter International Study of Oxaliplatin/5FU-LV in the Adjuvant Treatment of Colon Cancer). Journal of Clinical Oncology. 2019;37(15).

215. Park JH, van Wyk H, McMillan DC, Edwards J, Orange C, Horgan PG, et al. Preoperative, biopsy-based assessment of the tumour microenvironment in patients with primary operable colorectal cancer. J Pathol Clin Res. 2020;6(1):30-9.
216. de Jong PR, Mo JH, Harris AR, Lee J, Raz E. STAT3: An Anti-Invasive Factor in Colorectal Cancer? Cancers (Basel). 2014;6(3):1394-407.
217. Knupfer H, Preiss R. Serum interleukin-6 levels in colorectal cancer patients--a summary of published results. Int J Colorectal Dis. 2010;25(2):135-40.
218. Turano M, Cammarota F, Duraturo F, Izzo P, De Rosa M. A Potential Role of IL-6/IL-6R in the Development and Management of Colon Cancer. Membranes (Basel). 2021;11(5).
219. Dowdall JF, Winter DC, Andrews E, Laug WE, Wang JH, Redmond HP. Soluble interleukin 6 receptor (sIL-6R) mediates colonic tumor cell adherence to the vascular endothelium: a mechanism for metastatic initiation? J Surg Res. 2002;107(1):1-6.
220. Johnson DE, O'Keefe RA, Grandis JR. Targeting the IL-6/JAK/STAT3 signalling axis in cancer. Nat Rev Clin Oncol. 2018;15(4):234-48.
221. Song Y, Tang MY, Chen W, Wang Z, Wang SL. High JAK2 Protein Expression Predicts a Poor Prognosis in Patients with Resectable Pancreatic Ductal Adenocarcinoma. Dis Markers. 2020;2020:7656031.
222. Yoshikawa T, Miyamoto M, Aoyama T, Soyama H, Goto T, Hirata J, et al. JAK2/STAT3 pathway as a therapeutic target in ovarian cancers. Oncol Lett. 2018;15(4):5772-80.
223. Miller CP, Thorpe JD, Kortum AN, Coy CM, Cheng WY, Ou Yang TH, et al. JAK2 expression is associated with tumor-infiltrating lymphocytes and improved breast cancer outcomes: implications for evaluating JAK2 inhibitors. Cancer Immunol Res. 2014;2(4):301-6.
224. Liu Q, Ai B, Kong X, Wang X, Qi Y, Wang Z, et al. JAK2 expression is correlated with the molecular and clinical features of breast cancer as a favorable prognostic factor. Int Immunopharmacol. 2021;90:107186.
225. Tang S, Yuan X, Song J, Chen Y, Tan X, Li Q. Association analyses of the JAK/STAT signaling pathway with the progression and prognosis of colon cancer. Oncol Lett. 2019;17(1):159-64.
226. Monnien F, Zaki H, Borg C, Mougin C, Bosset JF, Mercier M, et al. Prognostic value of phosphorylated STAT3 in advanced rectal cancer: a study from 104 French patients included in the EORTC 22921 trial. J Clin Pathol. 2010;63(10):873-8.
227. Lin GS, Chen YP, Lin ZX, Wang XF, Zheng ZQ, Chen L. STAT3 serine 727 phosphorylation influences clinical outcome in glioblastoma. Int J Clin Exp Pathol. 2014;7(6):3141-9.
228. Qin HR, Kim HJ, Kim JY, Hurt EM, Klarmann GJ, Kawasaki BT, et al. Activation of signal transducer and activator of transcription 3 through a phosphomimetic serine 727 promotes prostate tumorigenesis independent of tyrosine 705 phosphorylation. Cancer research. 2008;68(19):7736-41.
229. Yue C, Shen S, Deng J, Priceman SJ, Li W, Huang A, et al. STAT3 in CD8+ T Cells Inhibits Their Tumor Accumulation by Downregulating CXCR3/CXCL10 Axis. Cancer Immunol Res. 2015;3(8):864-70.
230. Irey EA, Lassiter CM, Brady NJ, Chuntova P, Wang Y, Knutson TP, et al. JAK/STAT inhibition in macrophages promotes therapeutic resistance by inducing expression of protumorigenic factors. Proc Natl Acad Sci U S A. 2019;116(25):12442-51.
231. Gok Yavuz B, Gunaydin G, Gedik ME, Kosemehmetoglu K, Karakoc D, Ozgur F, et al. Cancer associated fibroblasts sculpt tumour microenvironment by recruiting monocytes and inducing immunosuppressive PD-1(+) TAMs. Sci Rep. 2019;9(1):3172.
232. Comito G, Giannoni E, Segura CP, Barcellos-de-Souza P, Raspollini MR, Baroni G, et al. Cancer-associated fibroblasts and M2-polarized macrophages synergize during prostate carcinoma progression. Oncogene. 2014;33(19):2423-31.
233. Sahu N, Chan E, Chu F, Pham T, Koeppen H, Forrest W, et al. Cotargeting of MEK and PDGFR/STAT3 Pathways to Treat Pancreatic Ductal Adenocarcinoma. Mol Cancer Ther. 2017;16(9):1729-38.

234. Nielsen C, Birgens HS, Nordestgaard BG, Kjaer L, Bojesen SE. The JAK2 V617F somatic mutation, mortality and cancer risk in the general population. Haematologica. 2011;96(3):450-3.
235. Chang K, Willis JA, Reumers J, Taggart MW, San Lucas FA, Thirumurthi S, et al. Colorectal premalignancy is associated with consensus molecular subtypes 1 and 2. Ann Oncol. 2018;29(10):2061-7.
236. Tell RW, Horvath CM. Bioinformatic analysis reveals a pattern of STAT3-associated gene expression specific to basal-like breast cancers in human tumors. Proc Natl Acad Sci U S A. 2014;111(35):12787-92.
237. Robinson RL, Sharma A, Bai S, Heneidi S, Lee TJ, Kodeboyina SK, et al. Comparative STAT3-Regulated Gene Expression Profile in Renal Cell Carcinoma Subtypes. Front Oncol. 2019;9:72.
238. Reddy V, Cohen S. JAK Inhibitors: What Is New? Curr Rheumatol Rep. 2020;22(9):50.
239. Burke WM, Jin X, Lin HJ, Huang M, Liu R, Reynolds RK, et al. Inhibition of constitutively active Stat3 suppresses growth of human ovarian and breast cancer cells. Oncogene. 2001;20(55):7925-34.
240. Jiang X, Wu M, Xu Z, Wang H, Wang H, Yu X, et al. HJC0152, a novel STAT3 inhibitor with promising anti-tumor effect in gastric cancer. Cancer Manag Res. 2018;10:6857-67.
241. Zhu Z, Li E, Liu Y, Gao Y, Sun H, Ma G, et al. Inhibition of Jak-STAT3 pathway enhances bufalin-induced apoptosis in colon cancer SW620 cells. World J Surg Oncol. 2012;10:228.
242. Mulholland T, McAllister M, Patek S, Flint D, Underwood M, Sim A, et al. Drug screening of biopsy-derived spheroids using a self-generated microfluidic concentration gradient. Sci Rep. 2018;8(1):14672.
243. Nivarthi H, Gordziel C, Themanns M, Kramer N, Eberl M, Rabe B, et al. The ratio of STAT1 to STAT3 expression is a determinant of colorectal cancer growth. Oncotarget. 2016;7(32):51096-106.
244. Park SY, Lee CJ, Choi JH, Kim JH, Kim JW, Kim JY, et al. The JAK2/STAT3/CCND2 Axis promotes colorectal Cancer stem cell persistence and radioresistance. J Exp Clin Canc Res. 2019;38(1).
245. Adachi M, Cui CX, Dodge CT, Bhayani MK, Lai SY. Targeting STAT3 inhibits growth and enhances radiosensitivity in head and neck squamous cell carcinoma. Oral Oncol. 2012;48(12):1220-6.
246. Zhang C, Yang X, Zhang Q, Guo Q, He J, Qin Q, et al. STAT3 inhibitor NSC74859 radiosensitizes esophageal cancer via the downregulation of HIF-1alpha. Tumour Biol. 2014;35(10):9793-9.
247. van Pelt GW, Sandberg TP, Morreau H, Gelderblom H, van Krieken J, Tollenaar R, et al. The tumour-stroma ratio in colon cancer: the biological role and its prognostic impact. Histopathology. 2018;73(2):197-206.