

UNIVERSITY OF
BIRMINGHAM



Mitigating the impact of COVID-19
in the United Kingdom through
data to safeguard patients with
cancer and an immunogenomic
analysis of colorectal cancer

Name: Thomas Starkey

Student ID: [REDACTED]

Degree: PhD Cancer + Genomic FT (A300)

PhD Supervisors: Professor Andrew Beggs and
Dr Lennard Lee

2024

Word count: 42,119

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

Contents

ACKNOWLEDGEMENTS	6
STATEMENT OF DISRUPTION AND CHANGES TO RESEARCH.....	7
ABSTRACTS.....	9
An immunogenomic analysis of colorectal cancer.....	9
Mitigating the impact of COVID-19 in the UK through data to safeguard patients with cancer.....	9
CHAPTER 1: OVERVIEW OF METHODS FOR CHARACTERISING ANTI-TUMOUR IMMUNE RESPONSES IN CANCER	10
The anti-tumour immune response and colorectal cancer – a brief overview	11
Immunogenomic methodology – resolving anti-tumour immune responses.....	13
Immunogenomics pre-next generation sequencing and the utility of non-NGS methods.....	13
The power of next generation sequencing on a genome-wide and exome-wide level.....	16
Targeted next-generation sequencing – a directed clinical and research tool.	21
Considerations for application and scalability of high-throughput technologies in clinical settings.....	24
Concluding remarks and the potential of the Cancer Immunity Cycle as a novel way of defining immune function.....	25
CHAPTER 2: IMMUNOGENOMIC ANALYSIS OF COLORECTAL CANCER PRIMARY TUMOURS AND ASSOCIATION WITH TUMOUR FEATURES	27
INTRODUCTION AND AIMS.....	28
METHODS.....	30
Transcriptomic (RNA sequencing) datasets/patient cohorts.....	30
The Cancer Genome Atlas (TCGA).....	30
Neoadjuvant chemoradiotherapy (CRT) rectal cancer cohort	30
Methods for molecular subtyping and immunogenomic classification of CRC	31
Consensus Molecular Subtyping (CMS)	31
Colorectal Cancer Intrinsic Subtyping (CRIS)	31
Coordinate Immune Response Cluster (CIRC).....	32
Tumour Immunogenomic Report (TIGER).....	32
Comparison of classification methods	33
Data analysis and visualisation.....	34
Differential gene expression and geneset enrichment analysis	34
Hierarchical clustering and heatmap generation	34
Somatic mutations and tumour mutational load	34

Statistical methods.....	34
Data visualisation.....	34
RESULTS.....	36
Tumour immunogenomic report profiles in CRC at the primary tumour level.....	36
Baseline characteristics of the TCGA COADREAD cohort.....	36
Tumour immunogenomic report (TIGER) signatures and correlation with ESTIMATE	37
Relationship between TIGER immune profiling and CRC molecular subtypes	37
Immune subtyping using TIGER at the transcriptomic level.	39
Association of TIGER with CRC tumour biology, somatic mutations, and clinical outcomes.....	40
Transcriptomic profiling of locally advanced rectal cancers and association with response to neoadjuvant chemoradiotherapy (CRT)	48
Colorectal cancer subtyping and gene set enrichment by CRT treatment response.....	48
Cancer Immunity Cycle expression profiling	49
Neoadjuvant CRT response and cancer antigen presentation	51
DISCUSSION.....	54
CHAPTER 3: IMMUNOGENOMIC ANALYSIS OF LOCAL AND DISTANT METASTASES OF COLORECTAL CANCER.....	58
INTRODUCTION AND AIMS.....	59
METHODS.....	61
Transcriptomic (RNA sequencing) datasets/patient cohorts.....	61
Lymph Node Metastases (METALLIC).....	61
Distant Metastases	61
RNA extraction, sequencing, and data processing	62
Lymph Node Metastases and Distant Metastases	62
CaArray_notte-00422 patient cohort.....	62
Sequencing batch correction.....	62
Data analysis and visualisation – RNA sequencing	63
Differential gene expression and geneset enrichment analysis	63
Hierarchical clustering and heatmap generation	63
Molecular subtyping and co-ordinate immune response cluster (CIRC).....	63
Cancer Immunity Cycle (CIC) Stage profiling.....	64
xCell analysis pipeline and other methods of immune characterisation.....	64
Statistical methods	65
Data visualisation.....	65

RESULTS	66
Immune and epithelial transcriptomic profiling of CRC lymph node metastases	66
Immune architecture of distant metastases of colorectal cancer	80
DISCUSSION	88
CHAPTER 4: OVERVIEW OF THE CHARACTERISTICS OF SARS-COV-2 AND IMPACT OF COVID-19 AT THE POPULATION LEVEL	92
COVID-19 in people living with cancer in early 2020 – a brief overview.	93
A brief outline of COVID-19 in the United Kingdom in 2020	94
Biology of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)	95
SARS-CoV-2 transmission and structure	95
Physiological impact and implications of SARS-CoV-2 infection in humans	96
SARS-CoV-2 viral evolution	98
Therapeutic potential of understanding SARS-CoV-2	100
Diagnostic testing	100
Vaccines	101
Antiviral agents and anti-SARS-CoV-2 monoclonal antibodies	102
Population-based and cohort-based studies of COVID-19 disease	104
COVID-19 disease at the population level	104
COVID-19 disease in patients with cancer	104
Evolution of COVID-19 disease phenotype	105
Role of COVID-19 cancer patient registries	106
Concluding remarks and remaining challenges in patients with cancer	106
CHAPTER 5: MONITORING THE IMPACT OF CORONAVIRUS DISEASE 2019 (COVID-19) IN PATIENTS WITH CANCER AT THE START OF THE PANDEMIC	108
INTRODUCTION AND AIMS	109
METHODS	110
UK Coronavirus Cancer Monitoring Project (UKCCMP) overview (2020)	110
Data collection and processing	110
Participant inclusion criteria	111
Office for National Statistics (ONS) comparison dataset	111
Statistical analyses and data visualisation	111
RESULTS	113
Outcomes in patients with cancer by demographics and primary tumour subtype in 2020	113
UKCCMP cohort characteristics	113
UKCCMP comparison with ONS cancer census	113
Case-fatality rates by patient demographics	115

Case-fatality rates by primary tumour subtype	117
Clinical features and outcomes of patients with haematological and solid tumour malignancies	119
DISCUSSION.....	121
CHAPTER 6: AN EVALUATION OF HOW THE IMPACT OF CORONAVIRUS DISEASE 2019 (COVID-19) HAS CHANGED IN PATIENTS WITH CANCER ACROSS THE PANDEMIC	125
INTRODUCTION AND AIMS.....	126
METHODS.....	128
UK Coronavirus Cancer Programme (UKCCP) datasets	128
UK Coronavirus Cancer Programme (UKCCP) overview	128
UKCCP dataset curation and linkage	128
COVID-19 infection and clinical outcome definitions.....	130
Data processing	131
COVID-19 evaluation study designs and outcome measures.....	131
Primary 2-dose SARS-CoV-2 vaccination.....	131
Booster 3-dose SARS-CoV-2 vaccination	132
Post-vaccine COVID-19 antibody responses.....	133
Temporal evaluation of COVID-19	133
Statistical analyses and data visualisation.....	134
RESULTS	135
Primary 2-dose SARS-CoV-2 vaccination.....	135
Cohort and data characteristics.....	135
SARS-CoV-2 vaccine effectiveness in patients with cancer and the non-cancer population	135
Vaccine effectiveness in cancer patient subgroups.....	137
Vaccine effectiveness in the context of clinically relevant patient demographics	140
Vaccine effectiveness waning by cancer subgroup	142
Booster 3-dose SARS-CoV-2 vaccination	143
Cohort characteristics	143
SARS-CoV-2 second-dose primary and third-dose booster vaccine effectiveness	143
COVID-19 third-dose booster vaccination protection and age.....	146
Post-vaccine COVID-19 antibody responses.....	149
Cohort characteristics	149
Post-vaccination COVID-19 antibody responses in patients with cancer and non-cancer population	149

Post-vaccination COVID-19 antibody responses by cancer patient subgroup	151
COVID-19 antibody status, breakthrough SARS-CoV-2 infection and hospitalisation	151
Quantitative SARS-CoV-2 antibody responses, breakthrough infections and hospitalisation	153
Temporal evaluation of COVID-19 disease phenotype	156
Cohort characteristics	156
Temporal changes in COVID-19 outcomes in cancer and non-cancer populations	156
COVID-19 disease phenotype in 2022	158
COVID-19 disease phenotype by primary cancer subtype	160
COVID-19 disease phenotype by systemic anticancer therapy (SACT) types and classes	162
DISCUSSION.....	164
ACKNOWLEDGEMENTS AND CONTRIBUTION DISCLOSURES	171
Chapter 1: Overview of methods for characterising anti-tumour immune responses in cancer	171
Chapter 2: Immunogenomic analysis of colorectal cancer primary tumours and association with tumour features	171
Chapter 3: Immunogenomic analysis of local and distant metastases of colorectal cancer.....	171
Chapter 4: Overview of the characteristics of SARS-CoV-2 and impact of COVID-19 at the population level.....	172
Chapter 5: Monitoring the impact of Coronavirus Disease 2019 (COVID-19) in patients with cancer at the start of the pandemic.....	172
Chapter 6: An evaluation of how the impact of coronavirus disease 2019 (covid-19) has changed in patients with cancer across the pandemic.	173
REFERENCES	174
SUPPLEMENTARY MATERIAL	198
Experimental code	198
Supplementary tables	198
Supplementary figures	279

ACKNOWLEDGEMENTS

First, I would like to thank my PhD supervisors, Professor Andrew Beggs, and Dr Lennard Lee, for their invaluable guidance, support, and encouragement throughout my PhD journey. Their support, mentorship and expertise have been instrumental in my development as a researcher and afforded me opportunities which I could not have envisioned when I started my doctorate.

I am incredibly grateful to my colleagues and collaborators, especially within the Institute of Cancer and Genomic Sciences at the University of Birmingham and the UK COVID-19 Cancer Programme (UKCCP) for their feedback and willingness to share their knowledge and expertise from which I learned a great deal. I am also humbled to have had the opportunity to help inform clinical decision-making and impact cancer care during the COVID-19 pandemic through my research and thank the patients, advocates and healthcare staff who worked tirelessly on the frontlines during the pandemic.

I would also like to acknowledge the Institute of Cancer and Genomic Sciences at the University of Birmingham for providing me with financial support during my PhD studies. This funding has supported my research and provided opportunities including enabling me to attend national and international academic conferences to present my research. Among the most memorable highlights of my doctorate was the opportunity to attend and present my work at the American Association for Cancer Research (AACR) annual meeting in the United States which would not have been possible without the financial support of the Institute and Blood Cancer UK.

To my family, especially my Mum, Dad, and brother Ben, and my housemates of the last four years Hannah, Jamie, and Richard, I am deeply indebted for your unwavering love and support throughout my PhD journey. You have always been there for me, offering encouragement and helping me to stay motivated.

STATEMENT OF DISRUPTION AND CHANGES TO RESEARCH

In line with University of Birmingham regulations for postgraduate researchers (PGRs), I am outlining how my research was disrupted during the COVID-19 pandemic and how my research topics have evolved to mitigate the impact of the pandemic. In addition, I will outline how some of the research in my thesis has contributed original research to a new area of research which did not exist when I commenced my doctorate in September 2019.

My original research project sought to better understand the immunogenomics of colorectal cancer (CRC) and identify tumour determinants of this using a variety of *in vitro* and/or *in vivo* methodology. To this end, I have successfully performed immunogenomic profiling of CRC primary tumours on a transcriptomic level through using RNA sequencing from multiple CRC cohorts. I performed RNA extraction prior to sequencing as well as processing of sequencing data from in-house datasets as set out in my thesis as well as using multiple publicly available datasets, all of which I have analysed and disseminated within my thesis in an original fashion. As I had carried out much of the RNA extraction prior to the start of the COVID-19 pandemic, I was able to perform *in silico* analyses without requiring physical access to the University of Birmingham. However, potential *in vitro* and/or *in vivo* comparisons of my findings on the transcriptomic level have not been possible to complete within the timeframe of my doctorate. This is set out in my declaration of disruption to planned research form submitted alongside my thesis as per University of Birmingham regulations.

Considering national priorities during the Coronavirus Disease 2019 (COVID-19) pandemic and aligning to strengths and programmes at the University of Birmingham, I utilised skills that I developed during my bioinformatics and statistical training to address key issues facing individuals with cancer at the start of and throughout the COVID-19 pandemic. To this end, I have undertaken extensive original research evaluating the impact of COVID-19 in people living with cancer using clinician-reported data collected as part of the UK Coronavirus Cancer Monitoring Project (UKCCMP) and from population-level public health and secondary care registries. My body of research includes identifying associations between cancer features with adverse outcomes relating to COVID-19 and evaluating efficacy of SARS-CoV-2 vaccines in patients with cancer to determine cancer patient subgroups who have been and remain extremely clinically vulnerable to COVID-19 infection and severe disease.

I count myself fortunate that due to the timeliness of my work, and the efficiency I was able to deliver it, practice change emerged directly from my research with significant parts disseminated throughout the research community through publication in peer-review journals including *The Lancet Oncology*, *European Journal of Cancer*, *JAMA Oncology* and *Scientific Reports*. More importantly, my research has in part helped to inform clinical decision-making, with my work and that of the wider UK COVID Cancer programme (UKCCP) being directly

attributed to have given confidence in giving systemic anticancer therapy during the pandemic, highlighting the need to better protect cancer patients who have poorer outcomes following COVID-19 infection and reduced protection from COVID-19 vaccine, and establishing a clinical rationale for monoclonal antibody prophylaxis in clinically vulnerable cancer population subgroups.

I am humbled to have had the opportunity to help inform clinical decision-making and impact cancer care during the COVID-19 pandemic through my research. To this end, the role of senior researchers and leadership at the University Birmingham in facilitating and delivering COVID-19 cancer research during the pandemic cannot be understated. This includes Dr Lennard Lee, Professor Jean-Baptiste Cazier and Professor Gary Middleton who co-led the UK Coronavirus Cancer Monitoring Project (UKCCMP), which later evolved into the UKCCP, and Professor Pamela Kearns who coordinated the paediatric UKCCMP. This was made possible by the vision of Professor David Adams who was pro-vice chancellor of the College of Medical and Dental Sciences during the COVID-19 pandemic, enabling the rapid deployment of scientists and resources to facilitate project delivery.

The COVID-19 pandemic disrupted and altered my original research project and plans. However, it was important that I played my role to form and expand the field of COVID-19 and its impact on patients with cancer. Despite these changes, I believe that my research fulfils the requirements of the doctoral programme as set out by University of Birmingham regulations. I have clearly defined contributions of colleagues which have enabled me to carry out my research at the end of my thesis where this is applicable as well as instances where my research findings have been published in peer-reviewed journals.

ABSTRACTS

An immunogenomic analysis of colorectal cancer

Anti-tumour immune responses in colorectal cancer (CRC) have been defined in the context of immune cell infiltration and checkpoint regulation. However, functional immunogenomic changes and interplay with other features of CRC biology have not been fully described. I performed transcriptomic immune expression profiling in CRC primary tumours, pre-CRT treatment rectal cancers, local lymph node metastases and adjacent normal tissue, and secondary liver metastases and lung metastases. This included developing a pipeline analysing the Cancer Immunity Cycle and cancer antigen presentation characterised using validated gene sets, with immune expression profiles subsequently compared to other CRC features. I observed enrichment of functional immune gene sets defining immune cell infiltration into tumours and cancer cell killing in microsatellite instability and right-sided CRC tumours. Cancer antigen presentation signatures were highly expressed in rectal cancers with good CRT responses and expression within lymph nodes associated with early-stage CRC tumour proliferation and aggressiveness, corresponding with clinical outcomes. Higher tumour-infiltrating lymphocyte expression, driven by increased MHC Class II antigen presentation and cDC1 dendritic cell expression activity, was observed in CRC lung metastases. Better understanding immune functionality underlying anti-tumour responses is important for identifying CRC patients with increased risk of CRC tumour recurrence/metastasis and may elucidate novel targets for immunotherapy.

Mitigating the impact of COVID-19 in the UK through data to safeguard patients with cancer

Individuals with a cancer diagnosis were reported to have poor clinical outcomes associated with COVID-19 following infection. However, cancer is a heterogeneous group of diseases and studies assessing individual risk, and benefits of COVID-19 vaccines, had not previously been performed using population-level data. I evaluated COVID-19 disease phenotype in patients with cancer compared to the general population and within cancer subgroups in 2020 and over the course of the pandemic, assessing the effects of cancer features and demographics, and determining the efficacy of vaccines from SARS-CoV-2 antigen and antibody testing data. An acute COVID-19 disease phenotype was observed in patients with cancer in 2020 but became less severe over the course of the pandemic. COVID-19 vaccines provided effective protection against infections, severe disease, and COVID-related mortality for many patients with cancer. However, individuals with cancer show lower levels of vaccine protection and more rapid waning, reduced post-vaccine antibody responses, and remain at higher risk of COVID-19 hospitalisation and death compared to the general population. Individuals with haematological malignancies had the lowest level of vaccine protection and poorest post-COVID-19 outcomes. The overall impact of COVID-19 is heterogeneous but individuals with cancer still require protection from being exposed to COVID-19.

CHAPTER 1: OVERVIEW OF METHODS FOR CHARACTERISING ANTI-TUMOUR IMMUNE RESPONSES IN CANCER

The anti-tumour immune response and colorectal cancer – a brief overview

Colorectal cancer (CRC) is the 4th most common cancer in the United Kingdom, with over 42,000 new cases diagnosed every year (CRUK, 2017) ¹. The majority of CRC tumours (>95%) originate in epithelial cells of the gastrointestinal tract ², of which later stage cancers are characterised by metastatic spread to surrounding lymph nodes and more distant organs including the liver and lungs ^{3,4}. This process of tumourigenesis and subsequent metastatic spread is enabled by genetic changes either promoting tumour evolution or growth and/or by facilitating evasion of host immune responses which would otherwise 'seek and destroy' tumour cells ⁵.

Typically, as tumours evolve and diverge from the phenotype of normal cells, the immune system is proposed to recognise these cells through detection of tumour-specific neoantigens resulting in the activation and migration of antigen presenting cells (e.g., dendritic cells) to secondary lymphoid organs (e.g., lymph nodes) before processing tumour antigen and presentation on surface major-histocompatibility complex molecules to immune effector cells ^{5,6}. This process results in the activation of cytotoxic effector cells such as CD8+ T cells, NK cells, NKT cells and gamma-delta T cells which destroy the target cell in a contact-dependent process ^{7–11}. Following tumour recognition and cytotoxic T cell activation, lysosomal or cytoplasmic granules are released which leads to target cell lysis ^{11–13}. However, a tumour may also acquire the capacity to escape or modulate the host immune response to avoid detection and/or cancer cell death. This includes the process of immunoediting whereby tumour cells are specifically selected to enable immune escape, and pro-tumour responses such as the release of immunosuppressive factors such as tumour growth factor beta (TGF- β) and recruitment of immunosuppressive immune cells (regulatory T cells, myeloid-derived suppressor cells) to evade lymphocyte-induced cancer cell death ^{5,14}. Consequently, the balance between the anti-tumour immune response and the ability of a tumour to escape the immune response determines whether tumour growth is limited or promoted.

Across multiple cancer types, intra-tumoural heterogeneity has been identified as an important mechanism shaping anti-tumour immune responses in addition to tumour evolution and plasticity. This mechanism refers to the coexistence of distinct tumour cell subpopulations with different genetic and phenotypic profiles within an individual tumour or between a primary tumour and corresponding metastasis in advanced disease ¹⁵. High intra-tumoural heterogeneity has been shown to correspond with a more immunologically suppressed tumour microenvironment ¹⁶. This shaping of an immunologically quiescent tumour microenvironment is reported to correspond with features of tumour evolution including clonal neoantigen burden, sub-clonal neo-antigenic heterogeneity ¹⁷ and immune editing within the tumour thereby facilitating immune evasion ¹⁸. More pertinently, patients with high intra-tumoural heterogeneity have been reported to have poorer disease prognoses and outcomes; a trend which is notable for individuals with solid tumour malignancies ¹⁹. For example, individuals

presenting with head and neck squamous cell carcinomas (HNSCCs), lung adenocarcinomas or glioblastomas that have high intra-tumoural heterogeneity have worse overall and progression-free survival outcomes compared to those with tumours with lower heterogeneity^{20–22}. This is also important in the context of anticancer treatment response, with intra-tumoural heterogeneity shown to be associated with poorer responses to immunotherapy for multiple cancer subtypes, an observation which is even more pronounced for tumours which have metastasised to the liver^{18,23}. Taken together, this indicates an important role for intra-tumoural heterogeneity for shaping anti-tumour immunity across several cancer subtypes.

More specifically to colorectal cancer (CRC), a number of different features of anti-tumour responses have been directly linked to patient prognosis, identifying association of high Th1 memory cell, NK cell and CD8+ T cell infiltration with better patient prognosis^{24–27} whereas high infiltrates of other cell types such as MDSCs and tumour-associated macrophages are often observed in patients with poorer prognosis^{28–30}. As a result, the interplay of tumour biology and anti-tumour immune responses is frequently associated with or even deterministic of response to anticancer therapies. This is the case with several targeted immunotherapies such as Pembrolizumab (anti-PD1) and Ipilimumab (anti-CTLA4) which inhibit immune-regulatory checkpoints, thus remove the “brakes” from the host immune response and enabling infiltration of immune effector cells such as cytotoxic T-lymphocytes (CTL) and natural killer (NK) cells into the tumour to destroy cancer cells^{31–33}. For example, CRC patients with defective DNA mismatch repair (microsatellite instability, MSI) respond better to immunotherapy than patients with microsatellite stability (MSS)^{32–36} with Le et al confirming positive patient treatment response rates assessed by progression-free survival following treatment with Pembrolizumab (anti-PD1) were 78% and 11% for MSI and MSS tumours, respectively³⁶. This is even more pronounced for late-stage (metastatic) MSI CRC patients treated with pembrolizumab³⁴ or nivolumab/ipilimumab (anti-PD1/anti-CTLA4) combination therapy³⁷.

Other features of CRC primary tumours and the tumour microenvironment beyond microsatellite instability have been reported to correspond with immune biology and treatment responses. This includes the location of primary tumour development, with differences in anti-tumour immune responses reported between right-sided colon tumours compared to left-sided colon and rectal tumours^{38–40}. Right-sided colon tumours have been shown to have higher levels of immune activation and comprise a greater proportion of immunologically ‘hot’ tumours whereas left-sided colon or rectal tumours are more likely to be immunologically ‘cold’ or quiescent^{41,42}. This also corresponds with differences in epithelial tumour biology including differences in driver mutation frequencies between left-sided and right-sided colon tumours (e.g., *BRAF* and *KRAS*) which may be indicators of CRC disease prognosis^{43,44}. Moreover, individuals presenting with right-sided tumours also reported to have differential responses to some anticancer treatment classes including chemotherapy and immunotherapy when compared to those with left-sided tumours^{43,44}.

In addition to tumour location and mismatch repair status, anti-tumour immune response features have also been defined in the context of molecular subtypes, combining immune profiles with epithelial biology and other tumour features to classify CRC tumours ^{45–47}. These classifications have identified molecular phenotypes of tumours corresponding with high or low immune expression and patient outcomes in retrospective analyses of CRC tumours. This includes Consensus Molecular Subtyping (CMS), identifying an immunologically active phenotype (CMS1) corresponding to higher likelihood of microsatellite instability (MSI) and better prognoses when classifying CRC tumours from The Cancer Genome Atlas (TCGA) ⁴⁵. Similarly, *Thorsson et al* defined the majority of CRCs as having a ‘wound healing’ or ‘TGF-beta predominant’ molecular phenotype which correspond with different immune profiles and patient prognoses ⁴⁷. In combination, these studies highlight the relationship between primary CRC tumour biology, anti-tumour immune profiles and their importance in relation to overall disease prognosis. Beyond the primary tumour disease setting, there is also evidence of other factors determining immunotherapy efficacy such as the phenotypic nature and anatomical side of CRC metastases. A small study of 24 metastatic CRC patients treated with regorafenib (VEGFR2-TIE2 tyrosine kinase inhibitor) and nivolumab (anti-PD1) identified that patients with lung metastases had response rates to immunotherapy of 50% irrespective of MSI status ⁴⁸. While it has been previously reported that the molecular profile of lung metastases are relatively unique compared with other metastases (e.g., liver, brain, bone) ⁴⁹, the reason for why these metastatic CRC patients appear to respond well to immunotherapy irrespective of MSI status remains unclear. Understanding the anti-tumour immune response and interaction with tumour biology remains crucial for identifying patients who are likely to respond well to cancer treatment and provide an insight into patient outcomes. Several different methodologies to elucidate mechanisms underlying immune responses and identify phenotypic markers informative for clinical decision-making, from targeted protein expression assays to next-generation sequencing methods on a whole genome level. Here, various methods previously used to better understand the anti-tumour immune response in the context of cancer have been explored.

Immunogenomic methodology – resolving anti-tumour immune responses.

Immunogenomics pre-next generation sequencing and the utility of non-NGS methods

Tumour characterization and phenotyping is a valuable tool to drive better clinical prognostication and new therapeutic opportunities and has become commonplace in cancer research and hospital settings. Screening of tumours to identify this information is particularly informative for clinical decisions for potential treatments, namely biotherapeutics targeting specific cancer and anti-tumour immune response pathways. The earliest phenotyping methods were performed prior to the availability of next generation sequencing or the ability to fully discern gene-

level and genome-level changes and frequently focused on protein expression of known cancer-associated biomarkers. Many of these assays harnessed the capacity of antibodies to identify, quantify and visualize protein changes by western blotting, ELISA, and immunohistochemistry, respectively. Of these, immunohistochemistry remains as a tool for visualizing cancer biomarker expression in the clinical setting and can provide low-level diagnostic information in a timely manner.

Immunohistochemistry (IHC) involves the sectioning of patient samples, antigen retrieval and immunostaining utilizing specific primary antibodies against biomarkers of interest and enzyme-substrate interaction whereby an enzyme-conjugated antibody binds to primary antibodies and luminesces when the primary antibody is bound to target antigens/proteins ⁵⁰. In colorectal cancer (CRC), IHC remains commonly used for determining microsatellite instability (MSI) in the hospital setting based on the expression a panel of DNA mismatch repair (MMR) proteins namely MLH1, MSH2, MSH6 and PMS2, deficiency of which is associated with poor tumour differentiation and high TIL infiltration ^{51–53}. Thus, this approach is useful for identifying MSI patients who are more likely to respond well to immune checkpoint blockade ⁵⁴. Beyond clinical settings, expression of PD-L1 and markers of tumour infiltrating lymphocytes (TILs) such as CD8, TBET (Th1 pathway transcription factor) and GATA 3 (Th2 pathway transcription factor) have been visualised using IHC and has been associated with other features of tumour biology such as MSI status and BRAF/KRAS mutation status (e.g., BRAF V600E) ^{54–56}. As PD-L1 expression is known to provide a means of immune evasion via downregulation of T cell responses, this offers a mechanistic explanation for MSI patient response to PD-1/PD-L1 checkpoint blockade ^{32,57,58}. Carcinoembryonic antigen (CEA), important for metastatic dissemination of cancer cells, can also be assessed using IHC to monitor disease progression as higher CEA expression is observable in metastatic CRC patients ^{59,60}. Furthermore, CEA overexpression has been linked to the inhibition of the cytotoxic activity of NK cells and T lymphocytes ⁶¹, thus decreasing anti-tumour immune activity.

However, potentially one innovative use of IHC in CRC in recent years is the development of Immunoscore, a CRC tumour classification system based upon the density of CD3+ and CD8+ effector T-cells within a tumour and its invasive margin ^{62,63}. This alternative method of tumour classification focuses primarily on the host immune response rather than tumour characteristics and has been directly linked to patient outcomes ^{62,63}. When combining Immunoscore with IHC analysis of MSI status, Pages et al found that patients with a high Immunoscore had significantly higher disease-free survival irrespective of MSI status, identifying the potential clinical utility of assessing CD3+ and CD8+ density as an indicator of CRC patient prognosis ⁶³. Furthermore, Immunoscore has also been combined with immune profiling using next-generation sequencing to characterise anti-tumour immune responses in several cancer types, including CRC ^{42,64–69}. Overall, the IHC-based Immunoscore method in isolation measures immune cell density and perhaps acts as a useful prognostic indicator ^{62,63,70} rather than fully quantifying or describing immune cell function.

While immunohistochemistry is useful for quantifying and visualising cancer biomarkers, other techniques such as flow cytometry and mass cytometry (CyTOF) can measure characteristics of individual cells such as cell surface and cytoplasmic protein expression. These techniques utilise antibodies conjugated with fluorophores or metal isotopes to bind to individual cells which enables single cell counting, quantification and sorting ^{71,72}. This has been harnessed in the context of CRC to better understand immune cell phenotypes and compositions within the tumour microenvironment, including for effector T lymphocytes ^{73–78}. In one such example, Zhang et al identified significantly elevated populations of an abnormal CD4⁺ T cell phenotype (EpCAM⁺PD-L1⁺CCR5⁺CCR6⁺) in CRC using single-cell mass cytometry which was also associated with abnormal activation of pp38 MAPK-pMAPKAPK2 signalling implicated in the regulation of Th1 immune responses ^{74,75}. Both cytometric methods have also been combined to identify novel cell phenotypes including a CD103⁺CD69⁺ innate lymphocyte population which was strongly correlated with infiltration of effector T cells into tumours ⁷⁹. This cell subset was primarily observed in MMR-deficient (MSI) tumours and the finding that activated gamma-delta T cells were almost exclusively found in MSI tumours ⁷⁹, in keeping with previous observations and with implications for immune checkpoint blockade ^{32,34}. Cytometric methodologies have also identified changes in immune cell populations following acute or chronic chemotherapy, with treatment characterised by a depletion of mature CD56^{dim} CD16⁺ NK cell populations and corresponding increase in immature CD56^{dim} CD16⁻ and CD56^{bright} NK cell populations ⁷⁵. These findings could be of clinical relevance, due to the importance of mature NK cells (CD16⁺) for mediating NK cell responses including cytotoxic activity involved in cancer cell killing ^{8,75,76,80}. Taken together, if we address the issues of diagnostics standardisation and assurance, there could be potential utility of cytometric methods to observe changes in immune cell populations and identify novel abnormal cell populations associated with cancer that has been contextualised with respect to anticancer treatment.

In addition to definition of anti-tumour immune responses by infiltrating immune cell populations and biomarkers, the characterisation of cytokines released to coordinate, stimulate, and regulate immune responses is also a key component. One method used to assess cytokine profiles was real-time or quantitative polymerase chain reaction (RT-PCR/qPCR), which simultaneously combines the conventional PCR process of target DNA amplification with sequence-specific fluorescently-labelled DNA probes that hybridise to the complementary DNA sequences during PCR ⁸¹. This enables real-time quantification of gene products and has been successfully used to identify differences in cytokine expression profiles in stage I lung cancer patients ⁸². Most notably, levels of five cytokines (CCL3, IL-8, IL-1 β , CXCL10, IL-2R α) were elevated in the peripheral blood mononuclear cells of patients before and after surgical tumour resection compared with non-cancer patients undergoing surgery ⁸². Commercially available multiplex immunoassays, combining sandwich enzyme-linked immunosorbent assay (ELISA) technology with fluorescent beads have also been utilised to identify prognostic and early detection cytokine markers in breast and ovarian cancers, respectively ^{83,84}. In the ovarian cancer study, serum concentrations of IL-6, IL-8, and VEGF,

EGF, and CA-125 serum cytokines were elevated in cancer patients in comparison to healthy controls ⁸⁴. Furthermore, cytokine profiling of metastatic melanoma patients following anti-PD1 therapy using multiplex immunoassays identified increased co-secretion of effector and stimulatory cytokines (Granzyme B, IFN- γ , MIP-1 α , IL-8) in patients who responded well to treatment ^{83,85}. Thus, upon further experimental validation, immunoassays may facilitate the identification of candidate biomarkers for early cancer detection and/or prognosis. In summary, qPCR and multiplex assay methods can be successfully implemented to quantify expression levels of immune-related proteins, using cytokines as an example, and potential for early-stage cancer detection.

Overall, there are significant opportunities and caveats in terms of using non-next-generation-sequencing (NGS) based methods. One major advantage of some of the methods described is the utility to rapidly and cheaply identify and quantify known cancer and immune biomarkers to inform clinical decision-making. Methods such as IHC can also provide a limited characterisation of immune cell populations and heterogeneity in the tumour microenvironment. However, these methods alone provide little information about immune function and the relationship between anti-tumour responses and tumour biology. As a result, NGS-based methods may potentially provide more opportunities for better understanding the anti-tumour immune response and interactions with other pathways.

The power of next generation sequencing on a genome-wide and exome-wide level

The continued development and increasing availability of next generation sequencing (NGS) methods have enabled high throughput characterization of mechanisms underlying tumour formation and cancer progression at multiple levels, from base substitutions at the gene level changes to genome-wide and chromosomal alterations. In synchrony with technological advances, the fidelity and versatility of NGS methods have significantly improved while the cost and time of sequencing runs have decreased concomitantly ^{86,87}. This in turn has facilitated population-wide DNA and RNA sequencing of cancer patients including the Cancer Genome Atlas (TCGA) ⁸⁸, HapMap ⁸⁹ and the 5,000,000 genomes project (formerly 100,000 genomes project) ^{90–92} which provide inexhaustible repositories of NGS data for the scientific research community. This combined advancement of sequencing methods and data availability has and continues to enable the identification of novel genetic variants in colorectal cancer, including suppression of anti-tumour immune responses facilitating immune escape.

Next-generation DNA sequencing determines the order of nucleotides of DNA molecules, comprising adenine, guanine, cytosine, and thymine bases and can resolve the sequence of the entire human genome. The most common method of NGS DNA sequencing uses sequencing-by-synthesis technology involving the creation of DNA libraries, application to a flow cell, bridge (DNA) amplification and incorporation of fluorescently-labelled deoxynucleoside triphosphate (dNTP) molecules with reversible terminators during sequencing cycles to enables

highly accurate base-by-base sequencing in a timely manner⁹³. The breadth and depth of NGS sequencing can be modified based on experimental requirements with whole genome sequencing (WGS) providing information about the whole genome whereas exome sequencing (WES) is targeted to protein-coding genomic regions and subsequently can be performed at a greater sequencing depth⁹⁴. Importantly, both levels of coverage can identify a broad range of genetic variation including single nucleotide polymorphisms (SNPs), insertions and deletions (INDELS) and structural variation⁹⁴. As a result, both methods have been applied to elucidate genomic changes associated with cancer and anti-tumour immune responses.

On a whole-genome level, genetic variants of both germline (inherited) and somatic (tumour-specific) nature have been observed in cancer patients associated with the anti-tumour immune responses. This includes DNA mismatch repair mutations (e.g., MLH1), BRAF and KRAS associated with higher immune cell infiltration into tumours⁹⁵, tumour proliferation⁹⁶ and inhibition of anti-tumour immune responses⁹⁷, respectively. To discover further genetic variants and determine association to anti-tumour immunity, non-gene candidate-driven approaches have been utilised to call and annotate variants from DNA sequencing data^{98–100}. Genome-wide association studies (GWAS) have been utilised to identify germline (heritable) genetic changes and associate these variants with immune infiltration in solid tumours. Using TCGA data, Shahamatdar et al discovered a novel SNP in the SIK1 gene (rs3366) associated with tumour follicular helper T cell population and germline variation in 23 genes, including four known to be involved in cytokine signalling⁹⁹. This has been taken further in the context of polygenic risk scores (PRS) which enables quantification of the level of disease risk based on genetic variants with potential utility already demonstrated by multiple studies for predicting a CRC diagnosis^{101,102}. Moreover, this will be assessed as part of the UK's 'Our Future Health' programme incorporating genetic and non-genetic data from five million individuals^{103,104}. Another non-hypothesis-driven approach, expression quantitative trait locus (eQTL) analysis, has also been utilised in combination with gene expression data to assess variation in specific DNA regions with expression of anti-tumour immune-related genes¹⁰⁰. In one such study, Lim et al discovered that germline variants determined the expression of the *ICOSLG* and *ERAP2* immune-related genes hypothesised to act as an inducible T cell co-stimulator and in the process of class I MHC antigen processing, respectively^{105,106}. Importantly, these studies demonstrate the utility of non-candidate driven NGS methods to discover novel genomic variation and provide contextualisation in relation to other cancer features.

Specifically, to tumour biology, DNA sequencing has been harnessed to identify and contextualise somatic features in terms of cancer biology. This includes the identification of driver mutations, known to confer a selective advantage during cancer evolution and in some instances, by facilitating immune escape^{107–109}. Driver mutations such as *PIK3CA* in colorectal cancer (CRC), including E542K and R38H, are known to alter PI3K signalling, of which pathway-activation mutations provide a mechanism for tumours to evade anti-tumour immune responses through promotion of an inhibitory myeloid microenvironment^{107,110–112}. Somatic variant annotation has also

facilitated the calculation of tumour mutational burden (TMB), defined as the total number of cancer-specific somatic mutations, in individual tumours ^{113,114}. This has been reported to correspond with anti-tumour immune responses and patient outcome following PD-1/PD-L1 checkpoint blockade therapy ¹¹⁵. Copy number variation (CNV) is also observable in several different cancer types and is characterised by gains and/or losses of copies of specific genes ^{116–119}. For some cancers, this has been directly associated with anti-tumour immune responses ^{120,121} including CRC where Lu et al showed that a low CNV burden (i.e. fewer copy number changes) was correlated with higher immune signature expression in multiple patient cohorts ¹¹⁸. Moreover, Lu et al proposed that CNV could be utilised to predict CRC patient response to immune checkpoint blockade as an alternative to TMB and microsatellite instability ¹¹⁸.

Combinatory analyses involving somatic variant annotation, TMB and CNV have subsequently been used to describe the mutational landscape of several cancer types ^{108,114,122–124}. For example, Kandoth et al discern differences in mutation frequency across cancer types, including those relating to individual genes implicated in PI3K, TGF-beta and Wnt signalling pathways whereby the mutation frequency of genes within these pathways was notably higher in colorectal cancers ¹⁰⁸. Moreover, these pathways are reported to interact with anti-tumour immunity with the mutation status of some of genes comprising these pathways genes such as *PIK3CA* have been directly linked to cytotoxic T cell infiltration, PD-L1 immune checkpoint expression and clinical outcomes in microsatellite stable CRC patients following immunotherapy ^{108,111}.

DNA sequencing data and somatic mutation information has also been used to identify and predict cancer neoantigens, tumour-specific antigens not present in normal tissues, which arise from altered protein structure as a consequence of DNA damage (e.g., mutations) and have been shown to initiate anti-tumour immune responses ^{125–129}. This has been demonstrated by Rooney et al, who identified correlations between the total number of predicted neoepitopes (neoantigens) and immune cytolytic activity (CYT) in multiple cancer types including breast, cervical and stomach cancers as well as association with HLA binding peptide-specific mutations ¹²⁸. Thus, this is in keeping with previous studies identifying relationships between TMB and patient outcomes as high CYT is associated with better patient outcomes ^{130,131}. Furthermore, Snyder et al identified candidate neoantigens predicted to have a strong binding affinity to MHC class I molecules in melanoma patients ¹²⁹. These neoantigens were comprised of tetrapeptides (4 amino acids in length) resulting from somatic mutations in several genes including *MUC4* for which overexpression has been implicated in multiple cancer types ^{129,132,133}. However, most importantly, Snyder et al identified that these neoantigens were shared between patients with good response to CTLA-4 checkpoint blockade therapy but absent in poor responders ¹²⁹. As a consequence of the importance of neoantigens in cancer and immune biology, dedicated neoantigen prediction pipelines and databases have been created (e.g., NeoPredPipe and TSNAdb) and neoantigens have become an attractive candidate for cancer vaccine and adoptive T-cell therapy development ^{134,135}.

In contrast to DNA sequencing, RNA sequencing captures genetic information at the transcript level, following transcription but prior to translation of messenger RNA (mRNA) into functional proteins. Perhaps of most interest, mRNA comprises a protein coding sequence which is flanked by untranslated regions (5' and 3' UTRs) in addition to a 5' cap and 3' poly A tail ¹³⁶. Different levels of sequencing coverage can be utilised for RNA sequencing ranging from total RNA-seq for full transcript information to 3' RNA-seq of coding mRNAs for gene expression information ^{137,138}. The basic RNA-seq process entails reverse transcription of extracted RNA to cDNA, which is subsequently fragmented, followed by sequencing of fragment ends and computational genome mapping of sequencing reads ¹³⁷. Mammalian cells comprise several types of RNA aside from mRNA such as micro RNAs (miRNAs) and small nucleolar RNAs (snoRNAs) involved in RNA silencing and post-transcriptional modification of RNA, respectively ¹³⁷. Indeed, both of these RNA types have been associated with cancer with implications for anti-tumour immune responses ^{117,139–141}. For example, the presence of a tumour-suppressing miRNA (miRNA-204-5p) in breast cancer has been implicated in the regulation of tumour growth, metastasis, and remodelling of the immune microenvironment ¹⁴⁰. More specifically, miRNA-204-5p expression corresponded with changes in myeloid and lymphocyte cell populations and upregulation of immunological pathways ¹⁴⁰. As miRNA-204-5p expression is often downregulated in several cancer types ^{140,142,143}, this implies a role for miRNAs in immune regulation and therefore are potential candidates for development of targeted anticancer treatment.

Though total RNA sequencing has provided useful insight into cancer biology, RNA sequencing-based tumour characterisation has primarily focused on capturing gene-encoding mRNA and generation of gene expression profiles. One way in which this has been utilised is to subtype tumours based on molecular biology. Specifically to colorectal cancer (CRC), two major classifications have been defined from gene expression profiles, namely Consensus Molecular Subtyping (CMS) and Colorectal Cancer Intrinsic Subtyping (CRIS) ^{45,46}. The first of these subtyping methods, CMS, combined previously published tumour classifications and RNA sequencing data from over 3,000 tumours with a computational network analysis which determined four distinct CRC tumour subtypes ⁴⁵. Subsequently, utilising both DNA and RNA sequencing data, each CMS subtype was found to have a unique molecular profile, including differences in expression of immune cell activation and infiltration signatures whereby one subtype, CMS1, was characterised by high expression of immune signatures ⁴⁵. Similarly, the CRIS subtypes were also associated with other features of tumour biology inferred from NGS data ⁴⁶ and both tumour classification methods also identified differences in patient outcomes based on assigned molecular subtype ^{45,46}. Importantly, both classification systems describe observed features of (primarily epithelial) tumour biology based on NGS data but provide little information about processes underlying the anti-tumour responses. Thus, several research groups have subsequently sought to use NGS data to characterise and subtype cancers based on anticancer immune responses, although classifications such as CMS and CRIS have not yet entered clinical practice ^{47,144–146}.

In one such example of immune characterisation of cancer, Thorsson et al analysed the expression of 160 cancer-specific and general immune gene signatures in 10,000 tumours from the TCGA ⁴⁷. Each geneset comprised a list of genes which have been identified based on a particular phenotype (e.g., CD8+ tumour-infiltrating T lymphocyte). This analysis resulted in tumour subtyping into 6 groups based on immune biology ranging from interferon-gamma dominant tumours characterised by high CD8+ cell expression and intratumoural heterogeneity to immunologically quiet tumours characterised by high macrophage and low lymphocyte expression ⁴⁷. These 6 immune subtypes were subsequently associated with neoantigen counts, tumour-specific somatic mutations and patient outcomes whereby patients with “immunologically quiet” tumours had poorer survival outcomes ⁴⁷. Furthermore, the proportion of tumours classified as each of the 6 immune subtypes varied markedly between different cancer subtypes, with colorectal tumours primarily classified as having a “wound healing” phenotype defined by high proliferative capacity and intratumoural heterogeneity ⁴⁷. Another such pan-cancer analysis also utilised TCGA data to develop a new immune analysis tool referred to as The Cancer Immunome Atlas (<https://tcia.at/>) whereby an “immunophenoscore” was allocated to each tumour based on expression of immune-specific genes and genesets ¹⁴⁷. In keeping with previous studies, Charoentong et al identified distinct immunophenotypes based on cancer subtype, with colorectal cancers characterised by high tumour mutational load and association of higher CD8+ T cell populations with patient survival ^{115,147,148}. Furthermore, notable cancer genotypes including BRAF and RAS mutation status were directly linked to immunophenoscore and expression of immune-stimulators and –inhibitors including immune checkpoints ¹⁴⁷. Further analysis of two cohorts of melanoma patients treated with anti-CTLA-4 and anti-PD1 antibodies implicated immunophenoscores and association with cancer genotypes as a predictor of patient response to immune checkpoint blockade agents ¹⁴⁷.

A sizeable proportion of immune characterisation to date has utilised NGS data obtained from “bulk” tumour samples comprising an admixture of immune and stromal cell types which comprise the tumour microenvironment. To overcome this caveat, methods for separating and sequencing individual cells have been developed to obtain gene expression profiles on a unicellular level. Several methods of single cell RNA sequencing (scRNA-seq) are available, including Drop-seq and the commercially available 10X genomics platform, and are based on the principle of microfluidic cell separation and capture prior to sequencing ^{149,150}. In colorectal cancer (CRC), scRNA-seq has been applied to obtain transcriptomic profiles of individual T cells divided into 20 cellular subsets and identify genetic features within each subset ¹⁵¹. From these profiles, Zhang et al observed significant differences in CD4+ and CD8+ T cell clonal expansion, migration and developmental transition based on transcriptomic expression and tracking of T cell receptors (TCRs) ¹⁵¹. Unique immune-related expression profiles were also obtained for each cell subset, including those associated with immune regulation and immune checkpoints ¹⁵¹. Perhaps most uniquely, utilisation of scRNA-seq can provide unique immune cell profiles for individual patients due to the ability to determine precise cell compositions which can also be extended beyond

immune cells ¹⁵² and potentially used to investigate cancer evolution. Despite the potential of scRNA-seq, this method remains uncommon because of cost compared with “bulk” tumour RNA sequencing.

Taken together, NGS-based methods have been successfully implemented to enable tumour classification based on features of epithelial and immune biology. Many of these have been based on the use of transcriptomic (RNA sequencing) patient cohorts and the relationship with other tumour features delineated from corresponding whole-genome or –exome sequencing data. To date, this experimental approach has focused on changes in immune cell infiltration or alterations in tumour biology enabling evasion of host immune responses and the development of single cell sequencing has enabled the capture of genomic information on an individual cell level. However, to further complement these methods, targeted NGS methods have been applied to gain a deeper understanding of immune cell variation derived from specific genomic regions which may have implications for the anti-tumour immune response.

Targeted next-generation sequencing – a directed clinical and research tool.

Conscious of the fact that host immune responses can be determined by variation within specific genomic regions, targeted NGS sequencing methods have been applied to increase sequencing depth and elucidate roles of specific genomic regions. This includes, though not restricted to, targeted assays to call RAS and BRAF mutations as standard clinical practice for several solid tumour malignancies ¹⁵³. One such multi-targeted NGS panel for colorectal carcinoma tests for KRAS, NRAS and BRAF mutations whereas a similar panel for assessing non-small cell lung cancer-associated variation within EGFR, ALK, BRAF, KRAS and MET (based on version 7.3 as of 20th September 2023) ¹⁵³. Further to this, targeted NGS can also be applied for characterising immune effector cell diversity and discerning the role of immunogenomics in the presentation of tumour antigens and neoantigens to immune effectors by highly specialized antigen presenting cells.

One of these important methods, T cell receptor (TCR) sequencing, enables the identification of unique T cell repertoires, generation of which is known to be important for TCR interactions and binding to peptide antigens presented by major-histocompatibility complex (MHC) molecules ^{154,155}. This interaction is enabled by the creation of unique TCR repertoires specific to MHC molecules and presented peptide antigens through V(D)J recombination, a specialised somatic recombination mechanism confined solely to the primary lymphoid organs (e.g., thymus for T cell maturation) ^{156,157}. Structurally for the majority of T cells, TCRs consist of paired alpha and beta chains, comprised of a unique combination of variable (V), diversity (D) and joining (J) segments whereby V/J and D gene segments encode alpha/gamma and beta/kappa TCR chains, respectively ^{156,157}. This in turn generates complimentary-determining regions (CDRs) on T cell receptors which enables T cell binding to presented tumour neoantigens ^{158,159}. Overall, three CDRs are generated for each T cell receptor, with CDRs 1 and 2 required as part

of TCR-MHC complex formation and CDR3 allowing for direct contact of TCRs to antigens or neoantigens presented by MHC ^{158,159}. Importantly, CDRs 1 and 2 have limited variability whereas CDR3 are highly variable and determine unique TCRs ^{158,159}. As a result, TCR sequencing have largely focused on elucidating CDR3 regions for profiling the basis of T cell function in anti-tumour immune responses ^{160–163}.

TCR repertoire diversity was first investigated using flow cytometry or TCR spectratyping of CDR3 regions ^{158,160}. Flow cytometry-based methods used monoclonal antibodies to determine the proportion of T cells with TCR-beta receptors whereas spectratyping involved PCR amplification of TCRs and subsequent separation by size using polyacrylamide gel electrophoresis (PAGE) ^{158,160}. However, NGS-based methods are now commonly used and often utilise RNA as starting material, due to the availability of gene expression (mRNA sequencing) and exclusion of intronic regions which prevents sequencing of non-important residual V(D)J/VDH rearrangements ^{162,164}. Further developments in TCR sequencing such as of unique molecular identifiers (UMIs), which ligate to target molecules during cDNA synthesis, enables the calculation of absolute count of RNA transcripts processed in a single sample and can be used to distinguish low frequency clonotypes and sequencing errors ^{165,166}. Thus, NGS methods for investigating TCR diversity have become the 'gold standard.'

In one such example on a pan-cancer level, Li et al identified 600,000 CDR3 sequences from paired-end RNA sequencing data from 9142 samples in The Cancer Genome Atlas (TCGA) comprising 29 cancer types ¹⁶³. The study found that V(D)J recombination formed the basis for TCR diversity, which was strongly positively correlated to tumour mutation load, inferring that either T cell diversity is driven by tumour-acquired mutations (immunoediting) or that somatic events occur in the tumour in response to T-cell-driven anti-tumour responses ¹⁶³. Between different cancer types, T cell diversity was also found to vary markedly with CRC and non-small-cell lung cancer both showing high diversity ¹⁶³. Furthermore, specific cancer types were characterised by alpha/beta and gamma/delta T cell profiles, identifying unique TCR profiles for low grade gliomas (LGG) and kidney renal clear cell carcinoma (KIRC) which were characterised by markedly different CDR3 sequence length distribution and amino acid conservation compared with CDR3 profiles of peripheral blood cells in healthy individuals ¹⁶³. However, perhaps the most important finding was the association of two cancer/testis (CT) antigens expressed by tumours, SPAG5 and TSSK6, with tumour infiltrating lymphocyte (TIL) diversity ¹⁶³ and thus could be potential anticancer vaccine targets.

Variation of other immune-specific genomic regions including human leukocyte antigen (HLA) proteins are also reported to be deterministic of host immune responses. HLA proteins comprise major histocompatibility complex (MHC) molecules whereby HLA genes are highly polymorphic and form the genetic basis of three MHC classes, two of which are directly involved in antigen presentation and therefore are integral to the presentation of cancer neoantigens to immune effector cells ^{167–170}. The antigen-presentation pathway for the first of these classes, MHC class I, involves the processing of intracellular virus or tumour antigens into peptides, transportation to the

endoplasmic reticulum (ER) and loaded into MHC class I protein complexes^{168,171}. These complexes are comprised of a heavy-chain encoded by the HLA-A, HLA-B and HLA-C genes as well as β 2-microglobulin (β 2m) which subsequently present antigens on the cell surface to cytotoxic CD8+ T cells¹⁷². Conversely, the MHC class II pathway involves presentation of extracellular antigens by MHC class II complexes (encoded by HLA-D genes) to CD4+ T cells via displacement of class II-associated invariant chain peptides (CLIP)¹⁷². Different class MHC molecules are expressed by different cell types whereby class I MHC molecules can be expressed by most nucleated cell types whereas constitutive class II MHC molecule expression is usually limited to professional antigen-presenting cells (e.g., dendritic cells)^{171,172}. Furthermore, the polymorphic nature of HLA genes enables the generation of a highly diverse MHC molecules which bind to a select repertoire of peptide antigens and in turn facilitate the presentation of these antigens to immune effector cells. Therefore, several NGS-based HLA genotyping assays now exist for analysis of HLA polymorphisms, including targeted MSK-IMPACT next generation sequencing and a tool enabling complete HLA typing from Oxford Nanopore sequencing data developed at the University of Birmingham^{169,173–176}.

To date, analysis of HLA genotypes has identified associations between different HLA genotypes and anticancer immune responses including HLA class I (MHC class I) molecules involved in neoantigen presentation to CD8+ T cells¹⁷⁵. In a clinical context, HLA class I homozygosity (loss of variation) in at least one of the HLA genes (HLA-A/B/C) was associated with reduced overall survival of melanoma and non-small cell lung cancer patients treated with ICB agents targeting CTLA-4 and/or PD-1/PD-L1¹⁷⁵. This phenomenon that was observed both in normal tissue and in tumour cells from patients with maximum heterozygosity of the HLA class I gene loci in 'normal' cells¹⁷⁵. Owing to previous observations of loss of HLA heterozygosity in cancer patients^{177,178}, this is indicative that HLA diversity is integral for binding a wide repertoire of tumour neoantigens as part of MHC antigen presentation pathways in turn enabling neoantigen recognition by CD8+ T cells. Moreover, specific HLA genotypes have been directly linked to cancer patient response to ICB agents^{179–181}, further indicating the utility of HLA typing.

Another key component of the anti-tumour immune response is the function of cytotoxic natural killer (NK) cells, capable of eliminating virus-infected cells in a rapid and non-specific manner¹⁸². NK cell activity, including killing function, is regulated by killer-cell immunoglobulin-like receptors (KIRs) which interact with MHC class I complexes ('ligands') determined in part by HLA genotype^{171,183,184}. KIR receptors are highly variable which allows for interaction with a broad range of HLA 'ligands' to elicit and regulate NK cell responses^{171,183,184}. An assessment of the germline genotype and copy number of 100 KIR-related genes in cancer patients within the TCGA identified higher copy numbers of inhibitory KIR genes in cervical squamous cell carcinoma and uterine carcinosarcoma¹⁸⁵. This was subsequently associated with poorer patient prognosis, suggesting that cytotoxic NK cell function is suppressed in tumours with high copy numbers of inhibitory KIR genes¹⁸⁵. Secondly, in a study specific to colorectal cancer (CRC), KIR and KIR ligand genotypes associated with disease recurrence and disease-free

survival in stage I/III patients following surgical tumour resection were identified using a single specific primer PCR (SSP-PCR) based method ¹⁸³. Most notably, a NK cell-activation KIR gene combination (2DS2-C1) was more frequently observed in patients with recurrence-free survival whereas inhibitory KIR gene combinations (e.g., KIR2DL1) were more common in patients with disease recurrence ¹⁸³. As a result, combinations of NK cell-activating and inhibitory KIRs are indicated to be important for regulating NK cell activities which is reflected in patient outcomes ¹⁸⁴.

To summarise, targeted sequencing approaches designed to obtain a greater depth of information related to immune-specific regions of the human genome have aided the discovery of immune diversity relevant to cancer and anticancer therapies. However, these approaches focus on specific regions and therefore do not fully account for other genomic changes influencing the anti-tumour immune response and interplay with tumour biology, which are captured by DNA and RNA sequencing on a genome- or exome-wide level. Overall, a potential technology direction that may yield more information is a combination of targeted and broad-scale NGS methodologies, or deeper investigation of targeted regions (e.g., TCR, HLA) from whole genome or exome sequencing data is an effective experimental strategy to elucidate mechanisms underpinning anti-tumour immune responses and implications for cancer treatment.

Considerations for application and scalability of high-throughput technologies in clinical settings

High-throughput next-generation sequencing (NGS) methods have been successfully implemented to elucidate features of tumour biology, anti-tumour immune responses and link this to disease progression and prognosis. The fidelity and coverage provided by these methods have enabled tumour categorisation or classification based on features of epithelial and immune biology using broad-range and targeted assays, including testing for KRAS and BRAF mutations in colorectal cancers now used in clinical practice ¹⁵³. However, challenges remain for the applicability and clinical scalability of implementing NGS technologies in routine clinical practice including granular molecular phenotyping of individual cancers.

Firstly, potential candidates for cancer biomarkers derived using high-throughput NGS methods require rigorous analytic and clinical validation prior to use in clinical settings ^{186,187}. Considerations for the first of these, analytical validity, include sample collection, processing and storage as these can affect outcomes and reproducibility of biomarker studies. This is especially pertinent for preparing RNA for NGS which can degrade quickly during sample collection, extraction and library preparation stages, in turn reducing sample quality for sequencing ¹⁸⁸. This is noted in the Cancer Research UK (CRUK) early detection and diagnosis of cancer roadmap with a lack of quality sample availability, especially repeated samples taken over time and pre-diagnosis, as a barrier to biomarker development ^{186,189}. In addition to sample preparation and quality, any high-throughput NGS

assay needs to accurately detect the biomarker(s) of interest (e.g., high specificity and sensitivity) and that this can be consistently replicated across different clinical laboratories and testing centres ^{186,190,191}. It is therefore important to measure and account for these issues for when considering clinical scalability. Subsequently, clinical validity needs to be ascertained whereby the degree to which a biomarker can be used to reliably identify individuals with the disease of interest and/or indicate risk of developing disease.

Beyond analytical and clinical validity, a strong candidate biomarker also needs to have demonstrable utility for use in clinical practice and informing clinical guidelines. In the context of cancer, clinical utility could include application for predicting risk of developing cancer, detecting or diagnosing disease (e.g., cancer screening), determining likely disease prognosis and/or probability of recurrence, or predicting response to anticancer therapies ^{187,192}. Following assay validity and clinical utility, financial constraints are also important to consider in the context of cancer biomarker development and testing. The CRUK early detection and diagnosis of cancer roadmap notes that there is often 'insufficient funding for validation and translation' of candidate biomarkers and that industry (research and development) investment can be limited due to a complex regulatory environment and unclear protectability of biomarkers (i.e., a patentable product) coupled with comparatively poor investment returns historically ¹⁸⁹. As a result, cancer biomarker testing using high-throughput NGS methods remains limited with this typically restricted to assessing changes in a relatively small number of genes.

Subsequent to biomarker development, scaling up high-throughput NGS methods into clinical practice is also resource-intensive ^{192,193}. The initial setup costs of NGS infrastructure are expensive and the process can generate large amounts of data which requires processing, analysis and storage. Moreover, this process is dependent on expertise in genomics and NGS methods, such as clinical bioinformaticians, for workflow development and interpreting outputs and training of clinical personnel to understand the results derived from NGS testing in order to bring about changes in clinical practice. The workflow subsequently needs to be standardized and quality control measures implemented to ensure accurate NGS testing and result interpretation which can be replicated across different clinical settings ¹⁹⁴. However, all of these steps require significant resources at a time where these are constrained in clinical centres, thereby potentially limiting the accessibility of high-throughput NGS testing in some settings. Taken together, the complex nature of cancer biomarker development, validation and test implementation in clinical settings remains a significant consideration for cancer biomarker development and testing using high-throughput next-generation sequencing methods,

Concluding remarks and the potential of the Cancer Immunity Cycle as a novel way of defining immune function.

To date, several methodologies have been successfully implemented to uncover features of the anti-tumour immune response in cancer. Anti-tumour immune responses in cancer have been well-defined in the context of immune cell infiltration and checkpoint regulation, and association with features of tumour biology described in the context of targeted immunotherapies. Indeed on a transcriptomic level using RNA sequencing data, features of the anti-tumour immune response have been quantified and characterised on a single gene or geneset level ^{47,144,147,195}. However, the functionality of the anti-tumour immune response goes beyond single genes and genesets and functional characterisation on the transcriptomic level remaining limited in scope. Moreover, the anti-tumour immune response is a multistep process involving presentation of cancer antigens, T cell priming and activation, and T cell trafficking of which tumour biology can interact with individual or multiple stages ¹⁹⁶.

The process of immune surveillance and control of tumour growth defining the anti-tumour immune response is accepted to be best modelled by the *Cancer Immunity Cycle*, a 7-stage process initiated by release of cancer cell antigens ^{196,197}. This is then followed by cancer antigen presentation, priming and activation of T-cells within the lymph nodes, trafficking of T cells to tumours, infiltration of t-cells into tumours, immune recognition of cancer cells, killing of cancer cells, which in turn leads to more release of cancer cell antigens ^{196,197}. To date, the *Cancer Immunity Cycle* has been described in the context of immune cell infiltration and immune markers as well as targeted anticancer therapies ^{197–199} but has not yet been quantified or functionally characterised at the transcriptomic level in colorectal cancer. One major strength of this approach would be to utilise previously validated and published genes and genesets known to be important in cancer biology and host immune responses from the outset to characterise the *Cancer Immunity Cycle* stages. Therefore, as a novel approach to understanding the anti-tumour immune response, quantification and functional characterisation of the *Cancer Immunity Cycle* may enable a much more accurate and comprehensive understanding of tumour immunogenomic biology for the scientific research community.

CHAPTER 2: IMMUNOGENOMIC ANALYSIS OF COLORECTAL CANCER PRIMARY TUMOURS AND ASSOCIATION WITH TUMOUR FEATURES

INTRODUCTION AND AIMS

Understanding the immunological processes underlying anti-tumour responses and corresponding changes in tumour biology is important for development of targeted anticancer therapies and determining potential efficacy. To date, anti-tumour immune responses in colorectal cancer (CRC) have been defined in the context of immune cell infiltration and checkpoint regulation, and transcriptomically quantified on a single gene or geneset level to subtype tumours on an immunogenomic basis and provide a snapshot of immune function. However, the functionality of the host immune response goes beyond single genesets and is a multistep process involving presentation of cancer antigens, T cell priming and activation, and T cell trafficking defined by the *Cancer Immunity Cycle* ^{196,197}. Thus, my first and primary aim is to define anti-tumour response in colorectal cancer (CRC) based on immunogenomic analysis of the cancer immunity cycle to quantify and characterise immune function using RNA sequencing data as the data input. Following the curation of a new analysis pipeline to characterise the anti-tumour immune response in the context of the *Cancer Immunity Cycle*, I will explore the association of this new method of immunogenomic profiling with primary CRC tumour biology. This includes molecular subtyping based on epithelial biology using RNA sequencing data and other genomic changes identifiable using exome sequencing data, namely analysis of mutational changes. Furthermore, I will investigate patient features associated with clinical outcomes and response to anticancer treatments such as microsatellite instability within data from The Cancer Genome Atlas (TCGA). In combination, this will enable a better understanding and characterisation of anti-tumour immune responses at the primary tumour level and any interplay with other features of cancer biology.

Adding mechanistic detail of the immunological processes underlying anti-tumour responses and corresponding changes in tumour biology is also important for determining potential treatment efficacy, including the association of immunophenotype with response to anticancer treatments. This includes individuals with locally-advanced rectal cancers following treatment with standard-of-care neoadjuvant chemoradiotherapy (CRT) ^{200,201}. Clinical responses following neoadjuvant CRT are associated with a number of molecular markers, including mutations in APC, KRAS, TP53 or high intertumoural heterogeneity which confer poorer clinical outcomes ²⁰². There is also evidence that higher blood lymphocyte, tumour lymphocyte, tumour CD8 and tumour CD4 counts are associated to better CRT responses ^{203–210}. However, higher counts of more immunosuppressive cells, such as FOXP3 T-cells, MDSCs and neutrophils are associated with poorer outcomes ^{206,207,211}. Given that these observations suggest that CRT might induce an immunogenic environment or inhibit tumour immunosuppressive mechanisms ^{212,213}, but the importance of the immune microenvironment pre-neoadjuvant CRT as a determinant for treatment response remains unclear. I therefore aim to identify tumour cellular phenotypes that determine the

anti-tumoural immune response and their association with clinical outcomes using transcriptomic data from a local-advanced rectal cancer cohort comprising individuals preceding neoadjuvant chemoradiotherapy (CRT) treatment.

METHODS

Transcriptomic (RNA sequencing) datasets/patient cohorts

The Cancer Genome Atlas (TCGA)

The Cancer Genome Atlas (TCGA) dataset is publicly available for the analysis of primary tumour and corresponding normal tissue gene expression. RNA sequencing data, somatic mutation data and corresponding clinical information from the TCGA colorectal patient cohort (COADREAD) was utilised and comprised 474 colon primary tumours, 166 rectal primary tumours, 41 normal colon tissue and 10 normal rectal tissue. In addition, the transcriptomic profiles of 50 normal liver and 110 normal lung specimens were utilised. RNA sequencing gene expression data was obtained in the form of normalized RSEM count data at the mean. Exome sequencing data was available and subsequently obtained for 489 colorectal primary tumours for somatic mutation analysis. Corresponding clinical information included patient characteristics (e.g., age of cancer diagnosis, gender), cancer staging (TNM and AJCC) and patient outcomes. Ethical approval for the data was available from the National Institute of Health (NIH). Expression data from the TCGA was obtained directly from the Broad Institute website (<https://gdac.broadinstitute.org>).

Neoadjuvant chemoradiotherapy (CRT) rectal cancer cohort

I utilised publicly-available bulk- and single-cell- RNA sequencing datasets from rectal cancer patients prior to neoadjuvant CRT, which comprised the transcriptomic profiles of 90 rectal cancer biopsies²⁰⁹ and single tumour epithelial cells (n=180) from five rectal cancer primary tumours, respectively¹⁵². These datasets are available under GEO accessions GSE109057 and GSE81861. In each case, clinical response to neoadjuvant CRT was quantified by Dworak tumour regression grade (TRG) and Neoadjuvant rectal score (NAR)²⁰⁹. Good clinical responses determined by TRG and NAR are indicated by a TRG of 3-4^{214,215} or NAR <8.0^{216,217}. Where applicable, Affymetrix probe IDs were normalised and converted to gene symbols using the 'gcrma' (v2.56.0) and biomaRt (v2.42.0) R packages, respectively.

Methods for molecular subtyping and immunogenomic classification of CRC

Consensus Molecular Subtyping (CMS)

Consensus molecular subtyping (CMS) of CRC was designed to resolve inconsistencies among previously reported CRC classifications based on gene expression data and enable clinical translation. It is based on the evaluation on six previously published CRC subtyping algorithms across 18 CRC sample sets based on a network-based approach to assess the associations between each of the six algorithms⁴⁵. This comprised Jaccard similarity coefficients to assess similarities and differences between algorithms and a Markov cluster algorithm to identify patterns, an analysis pipeline that in turn corresponded with four CRC subtypes. These subtypes corresponded with specific features in CRC biology including microsatellite instability and immune activation (CMS 1), canonical signalling pathway activation (e.g., WNT, MYC), metabolic signature expression and dysregulation (CMS 3) and a subtype characterised by epithelial-mesenchymal transition (EMT), stromal invasion and angiogenesis with poorer prognostic outcomes (CMS 4)⁴⁵. To perform CMS classification within the TCGA data for comparative analyses, I utilised the CMSclassifier R package (version 1.0.0, random forest machine-learning algorithm) using methods as previously described⁴⁵.

Colorectal Cancer Intrinsic Subtyping (CRIS)

Colorectal cancer intrinsic subtyping (CRIS) sought to categorise CRC tumours based on intrinsic epithelial tumour cell biology and overcome the challenge of stromal tumour content disproportionately influencing previous CRC transcriptional classifications. To counter this, CRIS classification was based on the use of patient-derived xenografts (PDXs) where the stromal components of the original CRC tumour had been substituted by murine tissue and gene expression data ascertained using human-specific expression assays⁴⁶. A non-negative matrix factorization (NMF) algorithm was subsequently implemented, identifying five CRIS subtypes with distinctive molecular and phenotypic profiles⁴⁶. This included subtypes characterised by enrichment of microsatellite instability and/or KRAS mutations (CRIS-A), epithelial-mesenchymal transition and poor prognosis (CRIS-B), epidermal growth factor receptor (EGFR) signalling and sensitivity to Cetuximab (CRIS-C), IGF2 overexpression and WNT signalling enrichment (CRIS-D) and a subtype characterised by TP53 mutations (CRIS-E)⁴⁶. For performing CRIS classification within the TCGA data for comparative analyses, I utilised the CRISclassifier R package (version 1.0.0) using methods as previously described⁴⁶.

Coordinate Immune Response Cluster (CIRC)

In contrast to CMS and CRIS, the coordinate immune response cluster (CIRC) was developed to identify differences in tumour immunophenotype and anti-tumour immune responses in CRC with this hypothesised to be important for improving the efficacy of stratified immunotherapy approaches ²¹⁸. The cluster is based on a 28-gene signature comprising markers of type 1 T helper (Th1) cell subsets, immune-related chemokines, adhesion molecules, major histocompatibility complex (MHC) class II molecules, and immune checkpoint genes including PD-1 and PD-L1 ²¹⁸. Once developed, CIRC scoring was performed using gene expression data from CRC tumours within The Cancer Genome Atlas (TCGA) and aggregating Z-score normalised transcript counts of the 28 genes in each tumour. Two-dimensional hierarchical clustering methods were subsequently applied to CIRC score data, identifying four patient groups with distinct molecular profiles, with groups displaying higher CIRC expression associated with microsatellite instability, DNA polymerase gene (POL) mutations and presence of inhibitory molecules targetable by immune checkpoint blockade such as CTLA-4 and PD-L1 ²¹⁸. On the other hand, groups with lower CIRC expression being characterised by RAS mutations and predict for poorer immune infiltration ²¹⁸. For evaluating CIRC in my analyses, I calculated co-ordinate Immune Response Cluster (CIRC) scores for each sample using methods previously described ²¹⁸. In brief, z scores were calculated for the 28 CIRC genes ²¹⁸ using the `scale()` function of R version 3.6.1 (`center = TRUE`, `scale = FALSE`) and averaged to generate CIRC scores.

Tumour Immunogenomic Report (TIGER)

The Tumour Immunogenomic Report (TIGER) is based on quantification each of the 7 stages of the Cancer Immunity Cycle (CIC) ¹⁹⁶ using publicly-available genesets from peer-review publications. In brief, the CIC is initiated by release of cancer cell antigens and subsequently followed by cancer antigen presentation, priming and activation of T-cells within the lymph nodes, trafficking of T cells to tumours, infiltration of t-cells into tumours, immune recognition of cancer cells, and killing of cancer cells ^{196,197}. Thus, TIGER comprises genesets representative of the 7 stages of the CIC, as well as signatures defining immune regulation and activation which were used to calculate geneset expression scores for each cycle stage. I determined gene signature expression values by calculating the geometric mean of transcript counts of each gene within each signature, firstly performing this using bulk transcriptomic data in the form of normalized RSEM count data at the mean obtained from The Cancer Genome Atlas CRC domain. The TIGER analysis pipeline generated geneset expression scores for 47 genesets used to define each CIC stage, stage constituents and immune regulation or activation signatures. Full information on the CIC genesets comprising the TIGER pipeline is available in Supplementary Tables 1-3. Subsequently, subtypes based on CIC stage expression were determined through performing a K-means analysis with gap statistic to determine the optimal number of clusters and unsupervised hierarchical clustering for CRC tumour subtyping. The 'complete' linkage method and 'euclidean' distance algorithm were used as parameters for clustering.

Comparison of classification methods

A summary of the key features of each molecular or immunogenomic classification of colorectal cancer (CRC) are outlined in Table 1, including key characteristics and previous investigation using TCGA data.

	CMS	CRIS	CIRC	TIGER
<i>Subtyping</i>	Yes – 4 subtypes	Yes – 5 subtypes	Not for development but later used for subtyping	Not for development but later used for subtyping
<i>Key characteristics</i>	CMS 1: High immune signature expression, microsatellite instability, high mutation count. CMS 2: High canonical signalling pathway expression, low immune signature expression. CMS 3: High metabolic pathway expression, KRAS mutations. CMS 4: Epithelial-mesenchymal transition, high TGF-beta signalling expression, poor disease prognosis. Derived using network analysis and a Markov cluster algorithm.	CRIS A: Microsatellite instability, mucinous, KRAS/TP53 mutations in MSS tumours. CRIS B: Epithelial-mesenchymal transition, high TGF-B signalling expression. CRIS C: High EGF-R expression, sensitive to Cetuximab treatment. CRIS D: High canonical signalling expression (MYC, WNT), IGF2 overexpression. CRIS E: Low immune expression, KRAS/TP53 mutations (MSS). Derived using a non-negative matrix factorization method.	Signature comprising 28 genes, including Th1 and immune checkpoint markers. Calculated using z-score normalised RNA counts data from individual genes. Signature defines Th1 anti-tumour immune responses in CRC. Higher CIRC expression associated with microsatellite instability, right-sided tumours, increased methylation and BRAF/PIK3CA mutations. Lower CIRC expression associated with left-sided tumours and KRAS mutations.	Characterisation of the 7 stages of the Cancer Immunity Cycle. 7 stages characterised using gene signatures published in peer-reviewed studies. Geometric mean of genes within each signature calculated. Quantification of immune function using RNA sequencing data from CRC primary tumours in the TCGA. Subtypes later derived using unsupervised hierarchical clustering with euclidean clustering distance algorithm.
<i>Analysis using TCGA data</i>	Yes	Yes	Yes	Yes
<i>Reference</i>	45	46	218	N/A

Table 1: Features of the molecular and immunogenomic classifications of CRC evaluated and discussed in this chapter. CMS: Consensus Molecular Subtyping. CRIS: Colorectal Intrinsic Subtyping. CIRC: Coordinate Immune Response Cluster. TIGER: Tumour Immunogenomic Report. TCGA: The Cancer Genome Atlas.

Data analysis and visualisation

Differential gene expression and geneset enrichment analysis

I performed Differential gene expression analyses using the DESeq2 R package (version 1.26.0) ²¹⁹. Analyses were performed on gene counts using the Wald test method and applying p value correction using the Benjamini-Hochberg method. Significant differences in gene expression were defined as a Benjamini-Hochberg P value of below 0.05. I generated volcano plots for differential expression analyses using the ggplot2 (version 3.3.0) and calibrate (version 1.7.5) R packages. I performed gene set enrichment analyses for 50 Hallmark genesets from the Molecular Signatures Database (MSigDB) using the GSEA software package version 4.0.1 (Broad Institute). Gene list inputs were pre-ranked by log2 fold change from DESeq2 analyses.

Hierarchical clustering and heatmap generation

Unsupervised hierarchical clustering and heatmap figure generation required the R Studio (R version 3.6.1), pheatmap (version 1.0.12) and RColorBrewer (version 1.1-2) R packages. Heatmaps were also generated using pheatmap where hierarchical clustering was not performed.

Somatic mutations and tumour mutational load

I processed whole exome sequencing data using the 'MAFTools' R package (version 2.2.10) to identify non-synonymous somatic mutations and calculate tumour mutational load (TML) for each patient, defined as the total number of non-synonymous somatic mutations.

Statistical methods

I performed statistical analyses in R versions 3.6.1 and 3.6.3, and the following R packages were utilized: broom, car, digest, dplyr, foreign, ggplot2, MASS, peperr, survival, survminer, sva, tidyverse. T tests and Mann-Whitney U tests were performed using the t.test and wilcox.test R functions, respectively. I fitted patient survival data using the 'survfit' function and corresponding Kaplan-Meier plots generated using the 'ggsurvplot' function.

Data visualisation

Data visualization and generation of all figures utilized several packages within R versions 3.6.1 and 3.6.3. I used functions from the ggplot2 R package to generate bar plots (geom_bar) with or without standard error bars

(geom_errorbar), boxplots (geom_boxplot, stat_boxplot), dotplots (geom_dotplot, geom_line), and spaghetti plots (geom_point, geom_line). I generated marginal plots using the ggmargin function of the ggextra R package. I also used the Stars (grDevices) and SankeyNetworks (NetworkD3) functions to generate star plots and sankey plots, respectively. I generated correlograms by applying a Pearson correlation coefficient between continuous variables and using the corrplot function of the 'corrplot' R package.

RESULTS

Tumour immunogenomic report profiles in CRC at the primary tumour level

Baseline characteristics of the TCGA COADREAD cohort

To evaluate the immunogenomics of CRC at the primary tumour level, I utilised The Cancer Genome Atlas (TCGA) COADREAD dataset comprising CRC tumours from 640 individuals. The baseline characteristics of this cohort are outlined in Table 2.

Characteristic	Total (n=640)
<i>Age</i>	
Age median (IQR)	68.0 (58.0-76.0)
<i>Gender</i>	
Female	300 (46.9%)
Male	340 (53.1%)
<i>Tumour location/site</i>	
Right Colon (COAD)	272 (42.5%)
Left Colon (COAD)	182 (28.5%)
Unspecified part of Colon (COAD)	20 (3.1%)
Rectum (READ)	166 (25.9%)
<i>Stage at diagnosis</i>	
Stage 1	110 (17.2%)
Stage 2	238 (37.2%)
Stage 3	185 (28.9%)
Stage 4	93 (14.5%)
Unknown	14 (2.2%)
<i>Microsatellite status</i>	
MSI high (msi-h)	92 (14.4%)
MSI low (msi-l)	106 (16.5%)
MSS (mss)	439 (68.6%)
Unknown	3 (0.5%)
<i>Death during follow-up period</i>	
Yes (OS Event)	133 (20.8%)
No (No OS Event)	505 (78.9%)
Unknown	2 (0.3%)
<i>Median follow-up period (days)</i>	
All patients (IQR)	669.5 (379.0-1096.0)
Patients with death event (IQR)	442.0 (188.0-1096.0)
Patients with no death event (IQR)	730.0 (419.0-1096.0)
<i>Median follow-up period (years)</i>	
All patients (IQR)	1.83 (1.04-3.00)
Patients with death event (IQR)	1.21 (0.51-3.00)
Patients with no death event (IQR)	2.00 (1.15-3.00)

Table 2: Baseline characteristics of The Cancer Genome Atlas (TCGA) colorectal (COADREAD) cohort. The COADREAD cohort comprised primary tumours from 640 individuals with a diagnosis of colorectal cancer (CRC).

Tumour immunogenomic report (TIGER) signatures and correlation with ESTIMATE

To quantify and functionally characterise the anti-tumour immune response based on the *Cancer Immunity Cycle* (CIC) in the context of colorectal cancer (CRC), I developed a new analysis pipeline, the Tumour ImmunoGENomic Report (TIGER). I first applied the TIGER analysis pipeline to publicly available mRNA sequencing data from colon and rectal cancer primary tumours within The Cancer Genome Atlas (TCGA) dataset (TCGA COADREAD) to determine efficacy. TIGER comprises 29 genesets representative of the 7 CIC stages and 18 genesets defining immune regulation and activation to enable the functional characterization and quantification of anti-tumour immune responses using RNA sequencing data (Supplementary Table 1). I determined gene signature expression by calculating the geometric mean of transcript counts of each gene within each signature which had been obtained in the form of normalized RSEM count data at the mean prior to analysis. Application of the TIGER analysis pipeline generated geneset expression scores for each Cancer Immunity Cycle stage, stage constituents and immune regulation or activation signatures and therefore demonstrating successful quantification.

Having determined the efficacy of the TIGER pipeline to quantify the CIC stages, I sought to determine associations between expression of CIC stages and other features of CRC tumour biology. Firstly, I compared TIGER/CIC stage expression with CRC tumour purity and immune content determined using the ESTIMATE R package ²²⁰, identifying that expression of all TIGER stages except for stage 3 (T cell priming and activation) were significantly correlated with a higher tumour immune content and lower tumour purity (Figure 1), and further supporting the efficacy of the TIGER analysis pipeline.

Relationship between TIGER immune profiling and CRC molecular subtypes

Subsequently, I performed molecular subtyping of CRC primary tumours using two major molecular subtyping methodologies for CRC, consensus molecular subtyping (CMS) and colorectal intrinsic signatures (CRIS) and compared with expression of genesets comprising the CIC stages and immune regulation or activation. I utilised the CMSclassifier (random forest) and CRISclassifier R packages to determine the molecular subtypes of the TCGA CRC primary tumours for CMS and CRIS, respectively ^{45,46}. From this analysis, I identified that distinct TIGER profiles corresponded to CMS subtypes determined using CMSclassifier, with increased expression of TIGER genesets in tumours classified as “CMS1” or “CMS4”, CMS classifications which are characterised by mismatch repair deficiency (microsatellite instability) and mesenchymal phenotypes, respectively (Figure 2) ⁴⁵. However, principal component analysis of CIC stages by CMS subtype shows large overlap between CMS subtypes, perhaps indicating subtle rather than major differences in immune responses between different epithelial CMS subtypes. Moreover, the TIGER profiles of tumours classified within different CRIS subtypes identified using CRISclassifier were not distinct (data not shown).

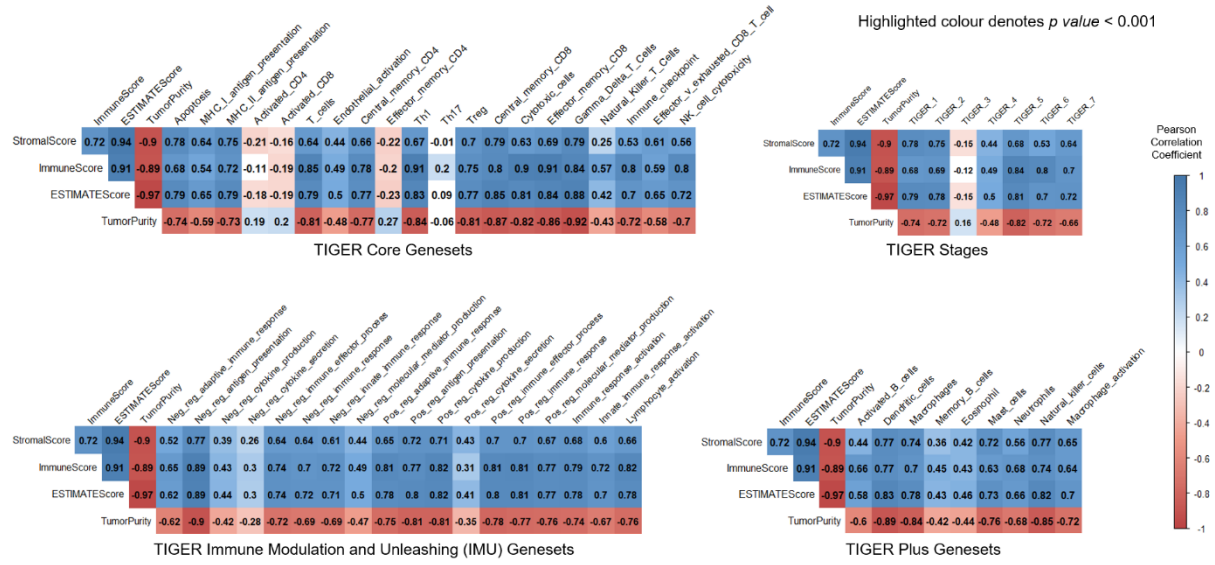


Figure 1: Correlogram profiles of TIGER stages and signatures with ESTIMATE scores for immune and stromal sample compositions and tumour purity. Pearson correlation coefficients highlighted in colour denotes p value < 0.001.

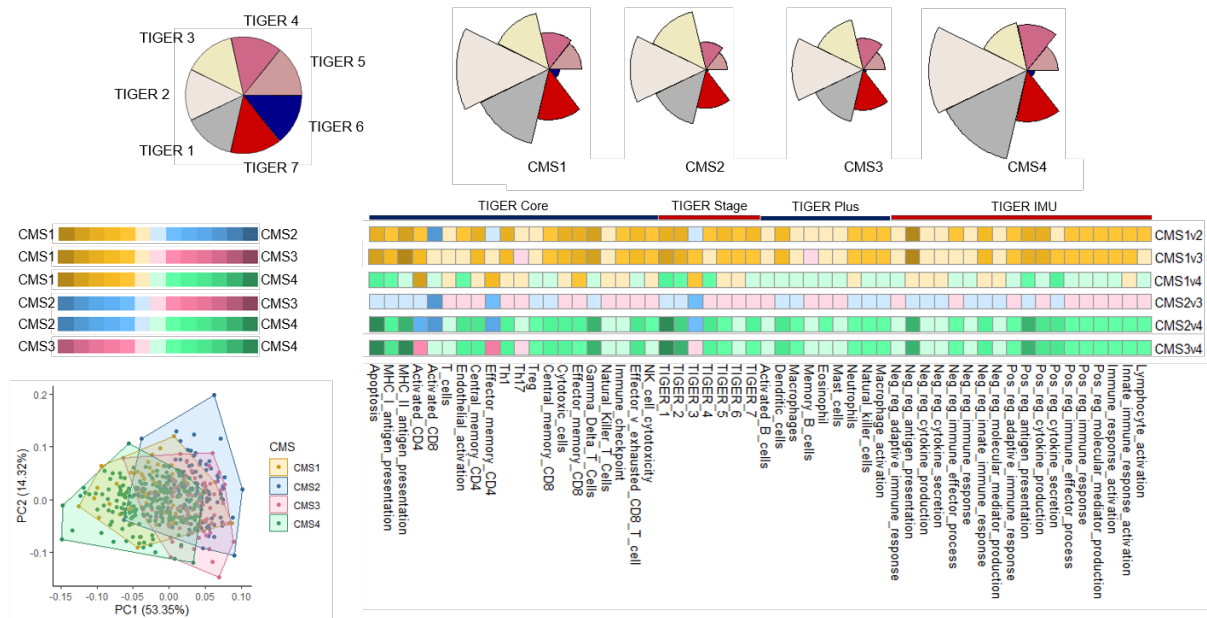


Figure 2: TIGER profiles by consensus molecular subtype (CMS) tumour classification. Top) Star plots of CMS subtype by TIGER (Cancer Immunity Cycle) Stage. Left) T-SNE plot of CMS subtype variation by TIGER features. Right) Heatmap of relative TIGER feature expression changes between CMS subtypes.

Immune subtyping using TIGER at the transcriptomic level.

To date, the major molecular subtyping methods for colorectal cancer including CMS and CRIS have been based on epithelial tumour biology ^{45,46} as opposed to immune phenotype, and thus global patterns of activation from immunogenomic signatures are poorly defined. Therefore, to better understand these patterns and identify subtypes based on immune biology, I performed hierarchical clustering using the degree of activation of different cancer immunity cycle (CIC) stages ¹⁹⁶ as defined by TIGER. K-means analysis with gap statistic based on CIC stage expression determined that the optimal number of clusters were four. Subsequently, I performed hierarchical clustering and identified four clusters with distinct TIGER profiles (Figure 3a). One cluster, CIC cluster 1, was characterised by strong activation of all CIC stages (CIC cluster 4, hyper-immune), a second cluster was characterised by poor/limited activation of CIC stages (hypo-immune), and a third and fourth cluster which appeared to be characterised by an early block in the cancer immunity cycle, termed as “uninitiated immunogenomic phenotype” and a late block in the cancer immunity cycle, “unpropagated immunogenomic phenotype” (Figure 3b).

It is widely recognised that there is significant interplay between epithelial tumour biology and the anti-tumour immune response ^{47,144}, for which CRC primary tumours can be classified based on epithelial biology (CMS/CRIS) or immune features (CIC immune clusters). As a result, I subsequently investigated the association of the four CIC immune cluster subtypes with epithelial-based tumour subtypes (CMS and CRIS). From this analysis, I identified that a high proportion of tumours in the hyper-immune cluster (CIC cluster 1) were also classified as CMS4 tumours (mesenchymal, 79.8% - 67/84) based on epithelial biology whereas I observed low proportions of CMS2 and CMS3 tumours in the hyper-immune cluster (3.6% and 1.2%, respectively) (Figure 3c). This is in keeping with observations of increased expression of cancer immunity cycle stages in CMS4-classified tumours and previous identification of increased immune responses in CMS1 and CMS4 cancers ⁴⁵. Conversely, a high proportion of primary tumours in the hypo-immune cluster were classified as CMS2 (45.5% - 60/132) or CMS3 (41.7% - 55/132) tumours which are associated with lower immune response expression ^{221,222}. However, the CIC immune clusters were not significantly linked to CRIS subtyping (data not shown).

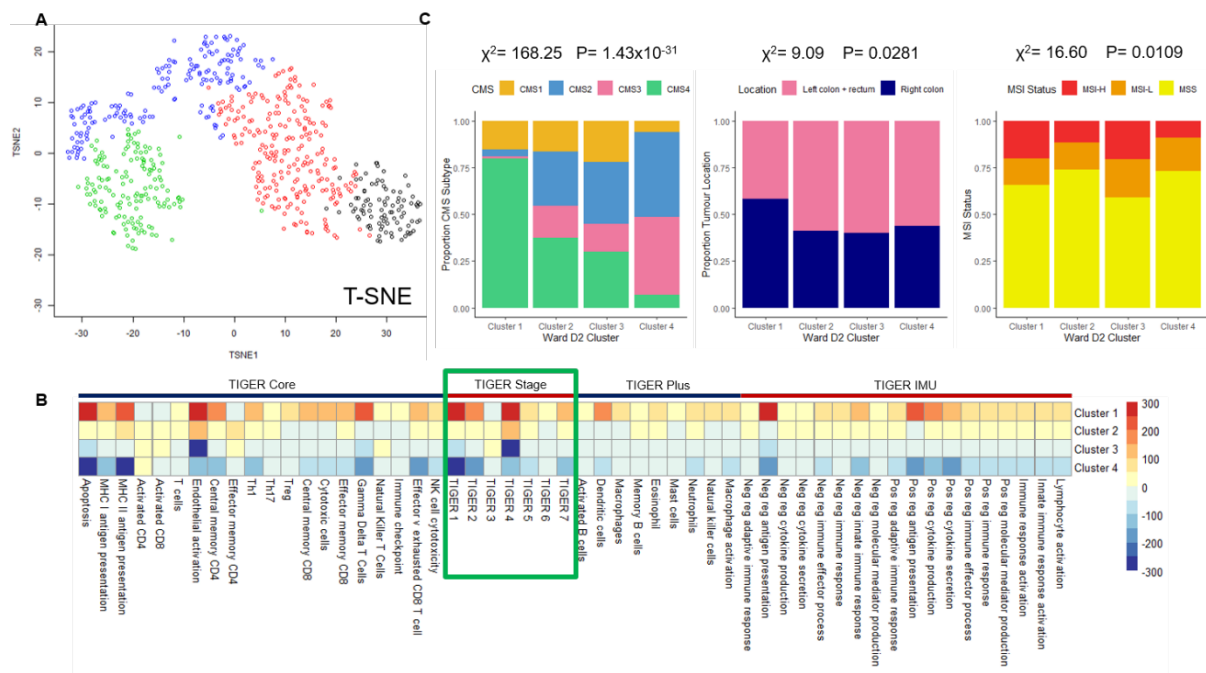


Figure 3: Cancer Immunity Cycle (CIC) cluster determination and molecular features. A) T-SNE highlighting distribution of CRC primary tumours within the four CIC clusters. B) Heatmap showing expression of TIGER features in each CIC cluster with CIC stages highlighted in a green box. C) Bar charts of sample proportions with chi-squared analysis by CMS subtype, tumour location and microsatellite instability for each CIC cluster.

Association of TIGER with CRC tumour biology, somatic mutations, and clinical outcomes

In addition to interplay with epithelial tumour biology, other patient and tumour factors may determine anti-tumour immune responses. Thus, I sought to determine the effects of patient and tumour factors on immunogenomic responses characterised using TIGER. I observed no significant differences in immune responses for notable patient factors including patient age and sex. In terms of tumour factors, I identified that the microsatellite instability (MSI) status of a tumour was the greatest driver of differences in immunogenomic profiles, corresponding with increased expression of T cell infiltration (CIC stage 5) signatures including NKT cells, TH1 cells, cytotoxic T lymphocytes and effector memory CD8 cells, immune checkpoint signature (CIC stage 6), and effector CD8 and NK cell cytotoxicity signatures (CIC stage 7) (Figure 4). Moreover, further analyses comparing the TIGER profiles of MSI-high (MSI-H) tumours with MSI-low (MSI-L) and MSS tumours identified significant differences in expression of Class II MHC antigen presentation (CIC stage 2), activated CD4 cells (CIC stage 3) and central memory CD4 cells (CIC Stage 5) in addition to Th1 cells and CD8+ T cells as previously described (Figure 4)⁹⁵. These findings in the context of T cell infiltration would be in keeping with known literature whereby MSI tumours are characterised by higher levels of immune cell infiltration^{32,33,95} and indicates the importance of microsatellite instability in

determining anti-tumour immune responses. However, perhaps most interesting was the association of other effector cells with MSI status, including NK and NKT cells, which has not been highlighted previously.

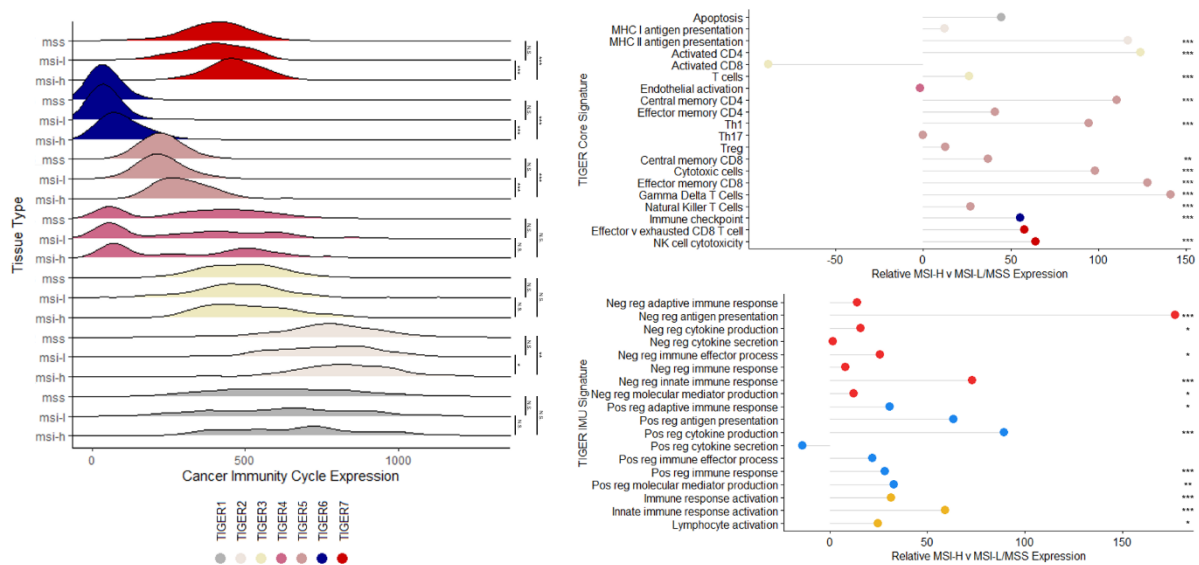


Figure 4: Cancer Immunity Cycle (TIGER) stages and constituent geneset expression changes based on microsatellite instability status. Left) Ridgeline plot showing differences in TIGER stage by MSI status. Right) Horizontal dot plots showing relative expression changes in MSI-high tumours compared with MSI-low and MSS tumours. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

As the tumour progresses, invading, involving more nodes and becoming metastatic, a shift in the immune architecture was also observed. This is frequently defined by lymphovascular invasion (LVI), for which I observed association with upregulation of negative regulation of cytokine production ($p < 0.001$) and secretion ($p < 0.001$) and production of molecular mediators ($p = 0.018$), indicating immunosuppression which has previously been observed in cancers with LVI ^{223,224}. Moreover, I identified decreased expression of activated CD4+ T cells (CIC Stage 3), NKT cells (CIC Stage 5) and NK cell cytotoxicity (CIC Stage 7) in stage 4 tumours compared with earlier tumour stages (stages 1 and 2 but not stage 3) (Figure 5) indicative of metastatic cancers acquiring the ability to evade host immune responses ²²⁵. Conversely, I found that other patient and tumour features such as age, sex/gender and presence of colorectal polyps were not associated with significant changes in TIGER immune profiles.

It has also been hypothesised that the primary tumour location within the colorectum is associated with differences in anti-tumour immune responses ^{38–40}, whereby right-sided colon tumours have been shown to have higher levels of immune activation ⁴¹. I therefore compared the TIGER immune profiles of right-sided colon, left-sided colon, and rectal tumours. I observed no significant differences in the immune profiles of left-colon and rectal tumours, and therefore compared right-sided colon tumours against left-sided colon and rectal tumours as a

combined group. This comparison identified enrichment of CIC stages 4-7 in right-sided colon tumours, indicating increased immune effector cell infiltration and cancer cell killing (Figure 6). In keeping with this observation, I found that several T cell signatures were more highly expressed in right-sided tumours including central memory CD4+ T cells, Th1 cells, cytotoxic T-lymphocytes, effector memory CD8+ T cells and gamma-delta T cells (all $p < 0.05$). Furthermore, I observed that signatures defining positive regulation of immune responses and immune activation were significantly upregulated in right-sided tumours in keeping with previous studies (Figure 6) ⁴¹.

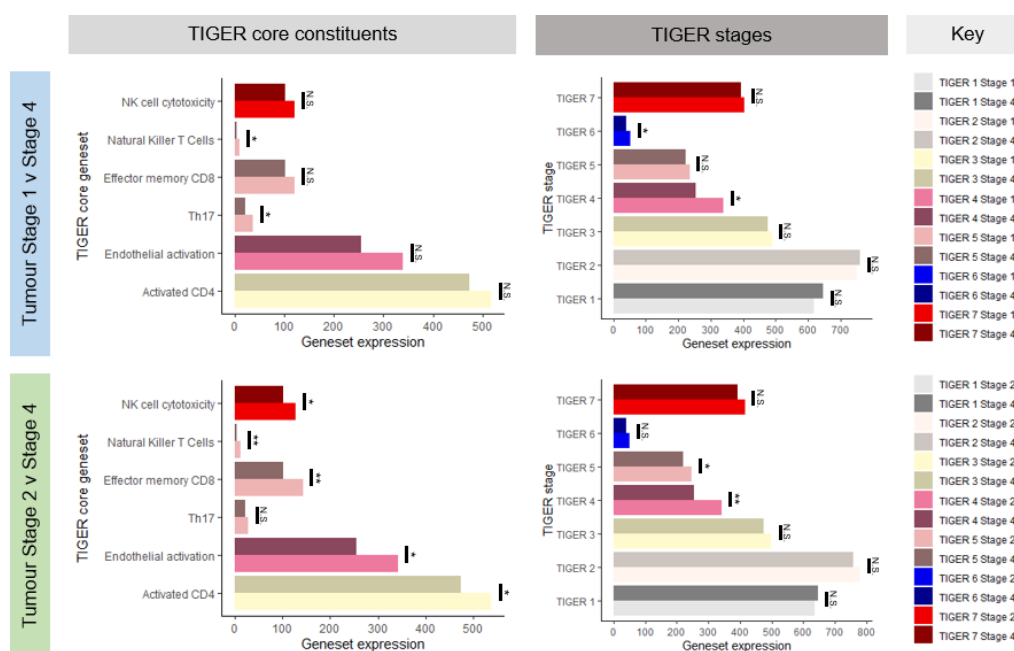


Figure 5: Bar plots showing expression differences between TIGER/CIC constituents (left) and stages (right) by tumour stage. Significant differences were observed between early stage (1/2) and stage 4 cancers. No significant differences were observed between stage 3 and 4 cancers (not shown). * $p < 0.05$, ** $p < 0.01$, * $p < 0.001$. 110, 238, 185 and 93 CRC tumours were classed as stage 1, stage 2, stage 3 and stage 4, respectively.**

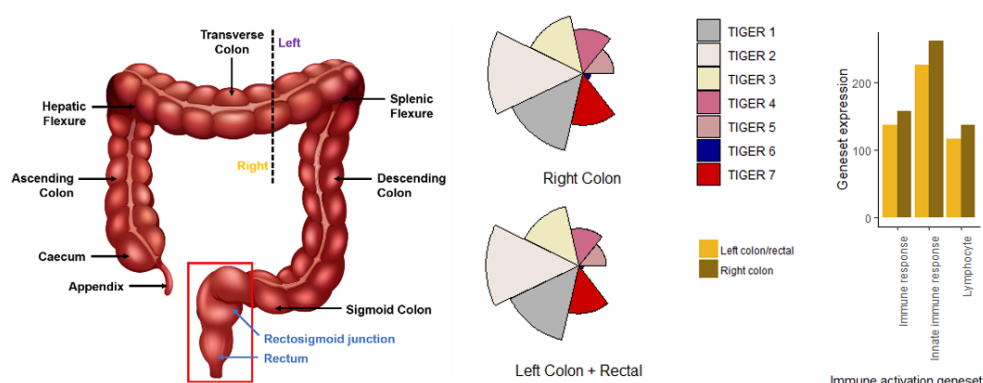
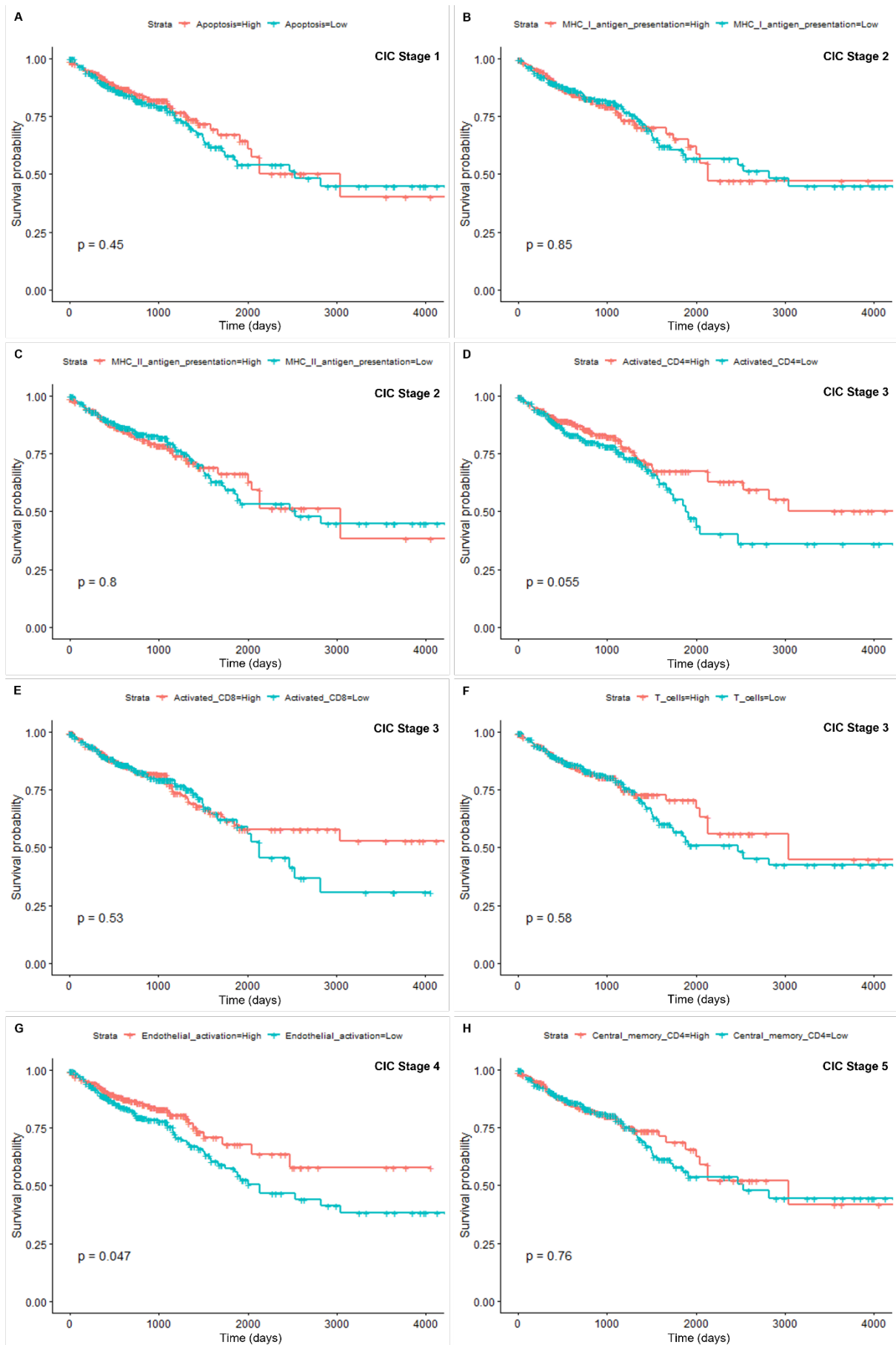


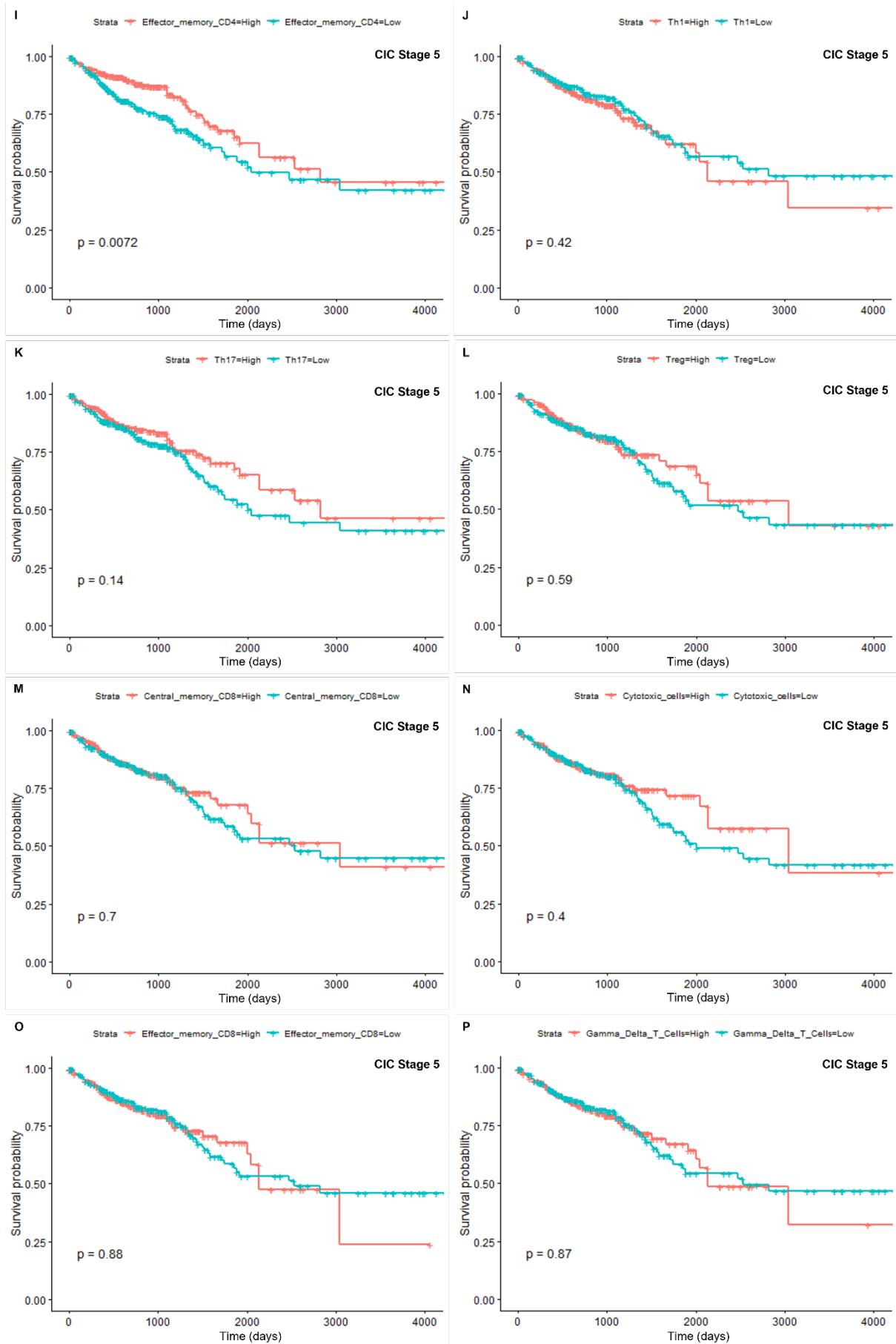
Figure 6: TIGER (CIC) expression changes between right colonic tumours and left colonic or rectal tumours. Left: Anatomy of the colorectum with colon labels in black and rectum labels in blue. Middle: TIGER stages by tumour location. Right: Immune activation geneset expression by tumour location.

Having identified differences in TIGER immune profiles based on patient and tumour features, I subsequently compared these features with the four CIC immune clusters. In the context of the clusters, the hyper-immune cluster (CIC immune cluster 1) was characterised by having more patients with right sided colon primary tumours whereas the hypo-immune cluster (CIC immune cluster 4) was characterised by lower levels of MSI (Figure 3c). However, the CIC immune clusters were not associated with differences in other patient features.

Further to patient and tumour features, several studies have suggested that CRC immune architecture is associated with clinical outcomes^{63,66,148}. Therefore, I compared the TIGER immune profiles of CRC patients known to have died with those of patients still alive at the point of data collection from the TCGA. I bifurcated TIGER immune gene signatures at the median and classified patients as having “high” or “low” expression of each signature. Subsequently, I performed a Kaplan-Meier analysis to assess clinical outcomes. Across all CRC patients, I observed that expression of T cell trafficking, defined by an endothelial activation signature (CIC Stage 4), and effector memory CD4+ (CIC stage 5) were significantly lower in patients who died (Figure 7). This was also reflected in the number of deaths (OS Events) and time to death (OS Time) (Supplementary Table 4). Furthermore, when only comparing clinical outcomes in patients with colon tumours, I observed significant correlation of higher activated CD4+ T cell signature expression (CIC stage 3) with patient survival ($p=0.019$), thus implicating the importance of an activated CD4+ T cell population and T cell trafficking prior to infiltration of effector CD4+ T cells into the tumour^{63,66,148}.

In keeping with other features of tumour biology, the development of somatic mutations might also alter the degree of activation of anti-tumour immune responses^{69,131,222}. Therefore, I interrogated changes in CIC stage and constituent expression in CRC tumours associated with acquired somatic mutations. Whole exome sequencing data for somatic mutation analysis was available for 489 CRC patients, and I applied a threshold of 5% of patients with a mutation of a specific gene (25/489 patients) as a cut-off for mutation analysis. Of all the non-synonymous somatic mutations, mutations of several genes were linked to the differential activation of different stages of the cancer immunity cycle (Figure 8). One notable gene with mutations identified were PIK3CA, linked with increased expression of CIC stage 7 as well as infiltrating T cell subsets (CIC stage 5) including central memory CD4, central memory CD8, NKT and effector CD8 T cells. Furthermore, I observed association of BRAF mutations with increased expression of CIC stages 5 and 7 and mutations in mismatch repair genes characteristic of MSI tumours such as MLH1 and MSH6 with higher expression of CIC stages 4 (T cell trafficking) and 7 (cancer cell killing), respectively (Figure 8). Conversely, mutations in CRC-associated KRAS and APC genes were linked with lower immune activity of CIC stages 2 (cancer antigen presentation) and 5 (T cell infiltration), respectively.





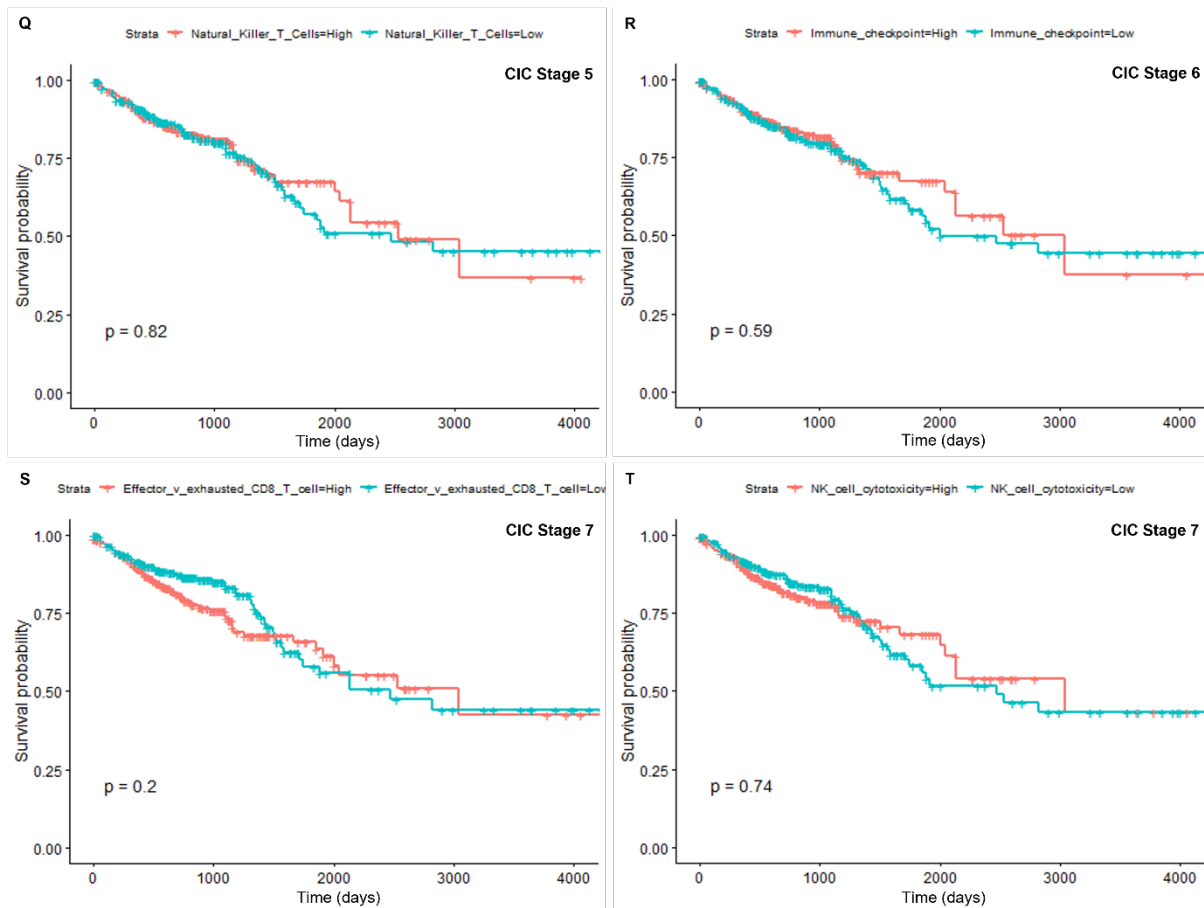


Figure 7: Cancer Immunity Cycle (CIC) constituent survival plots based on high or low gene signature expression. Kaplan-Meier overall patient survival plots based on bifurcated CIC expression categorisation. A) CIC Stage 1, B-C) CIC Stage 2, D-F) CIC Stage 3, G) CIC Stage 4, H-Q) CIC Stage 5, R) CIC Stage 6, S-T) CIC Stage 7. Patients with higher expression of effector memory CD4 (I, CIC stage 5) and endothelial activation (G, CIC stage 4) show significantly higher survival probabilities. The corresponding CIC stage for each constituent gene signature is given in each figure panel.

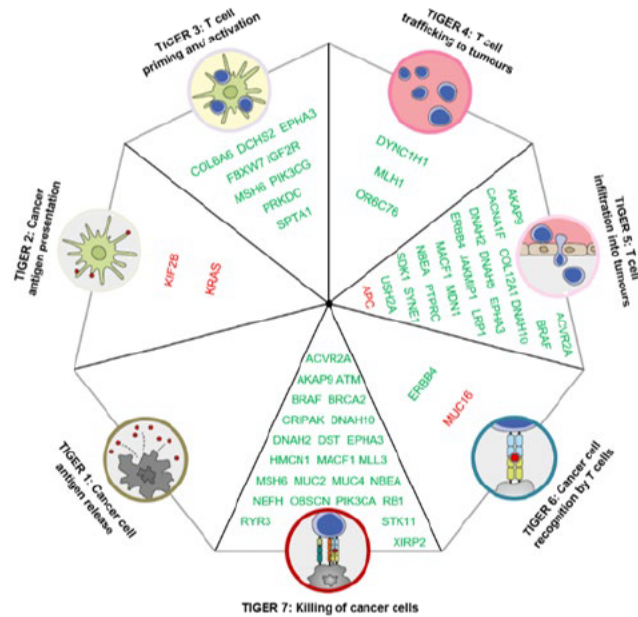


Figure 8: Cancer Immunity Cycle stage plot with association of gene-level somatic mutations to each stage. Mutations in genes shown in green were associated with increased CIC stage expression ($p < 0.05$) whereas mutations in genes shown in red were associated with lower CIC stage expression ($p < 0.05$). Somatic mutations were observed in $>5\%$ (25/489) of CRC patients with exome sequencing data. The 7 CIC stages represent those originally defined by Chen and Mellman, 2013¹⁹⁶.

In addition to individual mutations, it has been proposed that the total number of mutations in a tumour is also associated with immune cell infiltration and activity²²⁶. To assess this relationship in the context of the cancer immunity cycle, I calculated the tumour mutational load (of somatic mutations) for each sample using the MAFtools R package²²⁷. I utilised bootstrapping to calculate a threshold for high tumour mutational load (TML) across all tumours, defined as any sample above the 95% upper confidence interval across 493 patient tumours where mutational load data was available, irrespective of microsatellite stability status. This derived a cut-off of 425.7 non-synonymous mutations to categorise samples as having a high or low TML, with 77 and 416 tumours labelled as TML high and low, respectively. From the subsequent comparison of high and low TML tumours, I identified distinct TIGER profiles characterised by increased expression of innate responses (dendritic cells, natural killer T cells, innate immune response activation) and adaptive T cell responses (activated CD4 cells, effector memory CD8 cells, gamma delta T cells) as well as expression of CIC stage 5-7 in patients with higher tumour mutational load (Supplementary Table 5). Consequently, these finds are in keeping with previous immune profile association with TML or tumour mutational burden^{131,226}.

Transcriptomic profiling of locally advanced rectal cancers and association with response to neoadjuvant chemoradiotherapy (CRT)

Colorectal cancer subtyping and gene set enrichment by CRT treatment response.

Following immunogenomic characterisation at the CRC primary tumour level and identification of changes associated with tumour biology and patient outcomes, I sought to evaluate the relationship between transcriptomic immune profiling and anticancer treatment response. The opportunity to assess anti-tumour immune responses in this treatment context came about through a research collaboration with Dr Takashi Akiyoshi (Department of Gastroenterological Surgery, Cancer Institute Hospital, Tokyo, Japan) and Dr Seichi Mori (Cancer Precision Medicine Centre, Tokyo, Japan) and comprises a dataset which has been previously published ²⁰⁹. More pertinently, neoadjuvant chemoradiotherapy (CRT) is a standard-of-care treatment ^{200,201} whereby patient responses to treatment have previously been reported to be associated with differences in immune cell counts (T-lymphocytes, MDSCs, neutrophils) ^{203–211} and hypothesised to shape the immune microenvironment of the tumour ^{212,213}. The baseline characteristics of the CRT treatment response cohort are outlined in Table 3.

Characteristic	Total (n=90)
<i>Age</i>	
Age median (IQR)	61.0 (50.0-66.0)
<i>Gender</i>	
Female	27 (30.0%)
Male	63 (70.0%)
<i>Stage at diagnosis</i>	
Stage 2	38 (42.2%)
Stage 3	52 (57.8%)
<i>Tumour regression grade (TRG)</i>	
TRG = 1	54 (60.0%)
TRG = 3	14 (15.6%)
TRG = 4	22 (24.4%)
<i>Neoadjuvant rectal score (NAR)</i>	
NAR < 8.0	24 (26.7%)
NAR > 8.0	66 (73.3%)
<i>Death during follow-up period</i>	
Yes (OS Event)	7 (7.8%)
No (No OS Event)	83 (92.2%)
<i>Median follow-up period (days)</i>	
All patients (IQR)	2050.5 (1842.0-2410.5)
Patients with death event (IQR)	1522.0 (653.5-1767.0)
Patients with no death event (IQR)	2105.0 (1892.5-2467.0)
<i>Median follow-up period (years)</i>	
All patients (IQR)	5.62 (5.05-6.60)
Patients with death event (IQR)	4.17 (1.79-4.84)
Patients with no death event (IQR)	5.77 (5.18-6.76)

Table 3: Baseline characteristics of the neoadjuvant chemoradiotherapy rectal cancer cohort. The cohort comprised primary tumours from 90 individuals with a diagnosis of locally advanced rectal cancer.

Within a rectal cancer cohort comprising individuals treated with neoadjuvant chemoradiotherapy (CRT), I initially assessed if molecular subtypes of rectal tumours were associated with neoadjuvant CRT response, determining *Consensus Molecular Subtypes (CMS)* and *Colorectal cancer intrinsic subtypes (CRIS)* from gene expression profiles. *CRIS* subtypes were not significantly linked to tumour regression, however, there was a modest association with *CMS* subtyping ($\chi^2=12.36$, $p=0.006$ for NAR and $\chi^2=10.52$, $p=0.015$ for TRG, Figure 9a-b). *CMS* subtypes 1 (MSI-immune) and 4 (Mesenchymal) were both significantly associated with better tumour regression whereas *CMS* subtypes 2 (canonical), 3 (Metabolic) were associated with poorer outcomes. I additionally observed that higher expression scores of the *Coordinate Immune Response Cluster* signature, comprised of Th1-related immune genes, were associated with better tumour regression, indicative of the beneficial effects of an immunologically active tumour microenvironment ($p<0.05$ for both, Figure 9c).

To complement the subtle differences in the molecular profiles of patients observed based on CRT treatment response, I performed gene set enrichment analysis to provide further insight into the cellular pathways underlying CRT response. Gene sets enriched in good CRT responders were immune-related or immunoregulatory (inflammatory response, Interferon-gamma response, IL6-JAK-STAT3 signalling, TNF-alpha signalling) whereas only moderate correlations with epithelial oncogenic signalling pathways were observed (Figure 9d).

Cancer Immunity Cycle expression profiling

Having identified differences in immune-related gene set expression based on CRT response, I performed a more in-depth a *Cancer Immunity Cycle* immunogenomic analysis to gain a more comprehensive understanding of anti-tumour immune response phenotypes. I quantified expression of twenty previously published gene sets comprising the 7 *Cancer Immunity Cycle* stages (Supplementary Table 1) by calculating the geometric means of every gene within each gene set on an individual sample level. Subsequently, I compared the expression of each stage in the context of neoadjuvant CRT response. I observed that patients with good responses had significant differences in their *Cancer Immunity Cycle* profiles. The baseline biopsies in good responders had enhanced T cell infiltration (stage 5) as expected, ^{228,229}, however, I also observed that good responses to treatment were strongly associated with higher expression of earlier stages of the *Cancer Immunity Cycle*, including higher Cancer Antigen Release (stage 1) and Cancer antigen presentation (stage 2) (Figure 10). More specifically, I identified that a TRG of 3-4 was associated with higher expression of antigen presentation via MHC classes I and II ($p<0.05$ for both). This was strongly suggestive that the early phases of immune initiation are of as much importance as downstream T cell-driven responses.

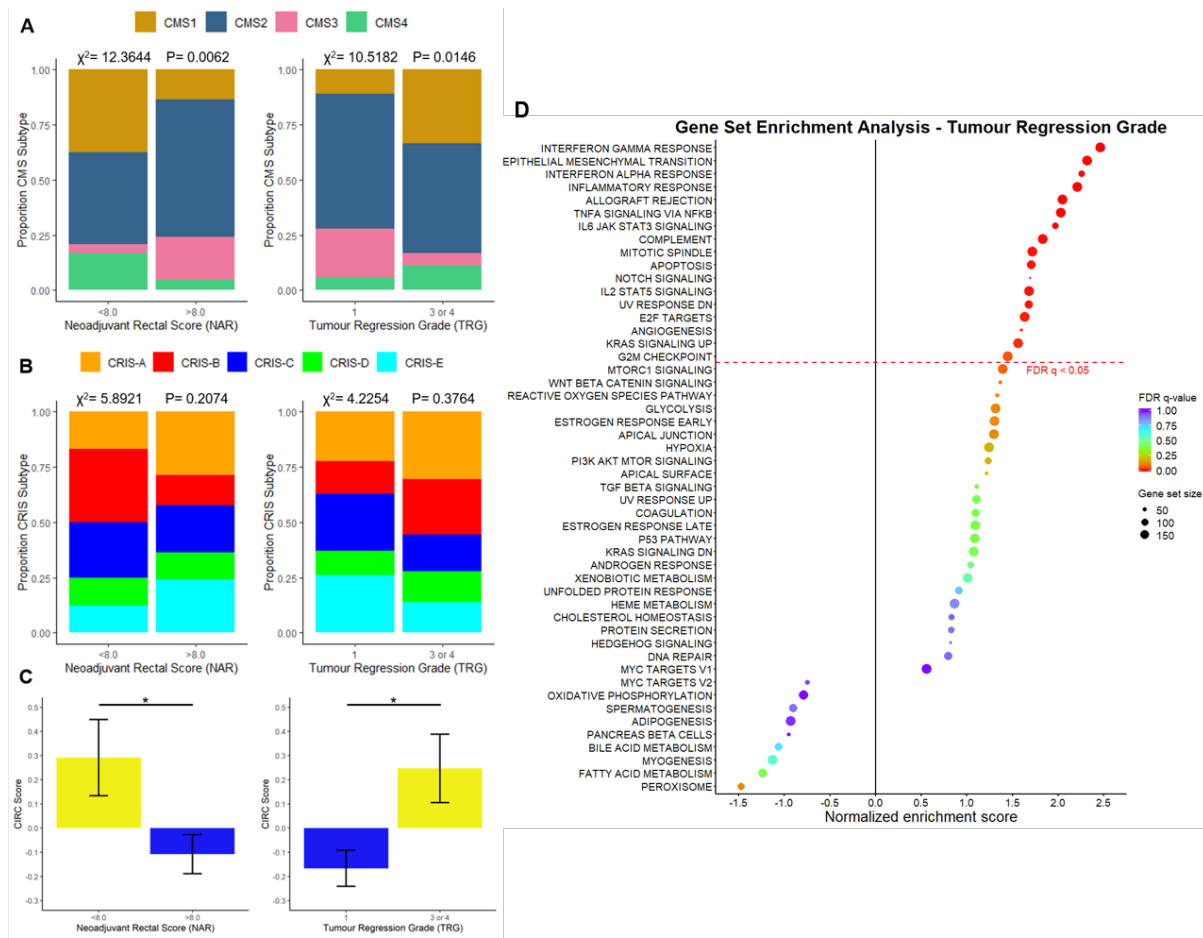


Figure 9: Association of neoadjuvant rectal score (NAR) and tumour regression grade (TRG) with molecular classifications of rectal cancer and differential gene set expression patterns. A-C) Bar charts showing the association between the CMS (A) and CRIS (B) classifiers, and CIRC expression scores (C). D) Dot plot illustrating the enrichment of GSEA hallmark gene sets and association with patients with good tumour regression. * $p < 0.05$.

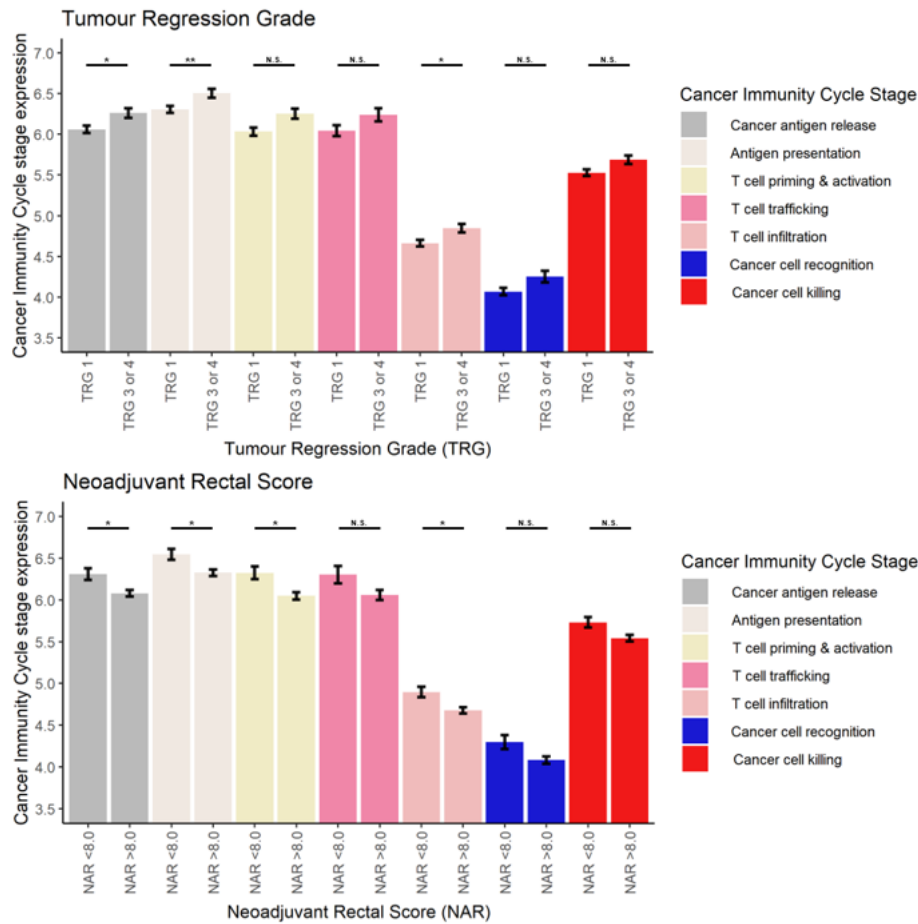


Figure 10: Cancer immunity cycle bar plots comparing expression levels of different stages of the cycle in patients with tumour regression assessed by tumour regression grade and neoadjuvant rectal score following CRT. * $p < 0.05$, ** $p < 0.01$, N.S.=not significant.

Neoadjuvant CRT response and cancer antigen presentation

Cancer antigen presentation is a multi-step process. This process initially involves protein synthesis of the tumour antigen, followed by proteolysis in the proteasome, peptide fragment transportation to the endoplasmic reticulum, MHC1 synthesis, binding to MHC molecules in the Golgi apparatus, vesicle formation and exocytosis to the cellular membrane to be presented to CD8+ T lymphocytes¹⁷². I analysed each of these factors and identified that vesicle formation by the Golgi apparatus (go_golgi_associated_vesicle) and exocytosis of these vesicles to the plasma membrane (go_exocytosis) were significantly associated with antigen presentation via MHC classes I and II, potentially flagging these cellular processes as rate-limiting steps (Figure 11).

To validate this, I utilised a single cell epithelial CRC dataset (GSE81861) to address the issue of epithelial and immune cell admixtures with bulk transcriptomic analyses. In keeping with our previous results, MHC class I presentation was strongly correlated with the degree of Golgi-associated vesicle formation and exocytosis, further adding weight to the importance of these cellular functions (Figure 12a-b). Furthermore, I observed that rectal tumours characterised by higher expression of vesicle formation and exocytosis gene sets were associated with better responses to neoadjuvant CRT treatment (Figure 12c-d) and therefore confirming the importance of early activation of the cancer immunity cycle through vesicle formation and exocytosis.

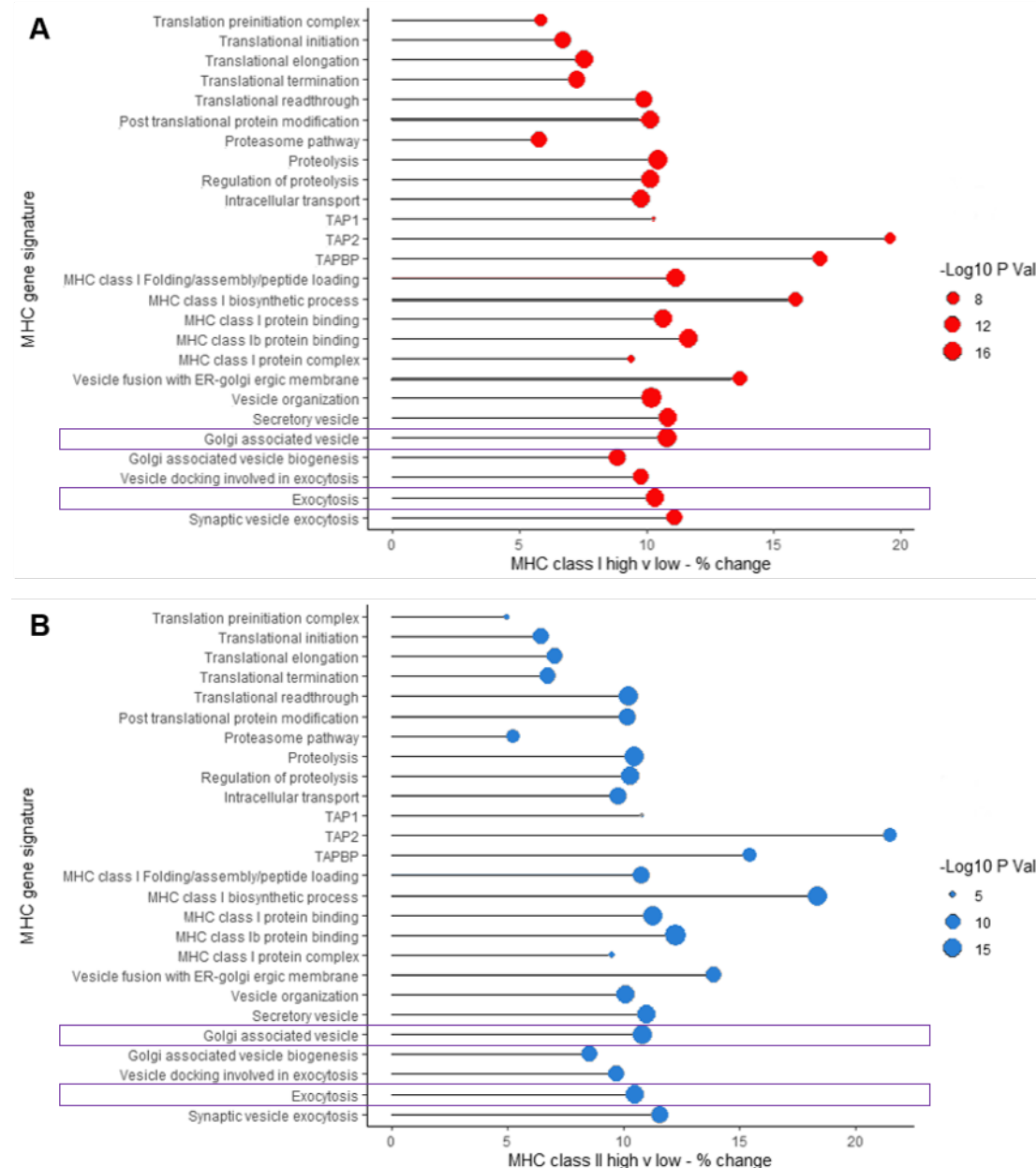


Figure 11: Lollipop plots showing the association between high MHC classes I (A) and II (B) expression with gene sets characterising steps of the antigen presentation pathway. Gene signatures representing golgi-associate vesicle formation and exocytosis are highlighted in purple boxes.

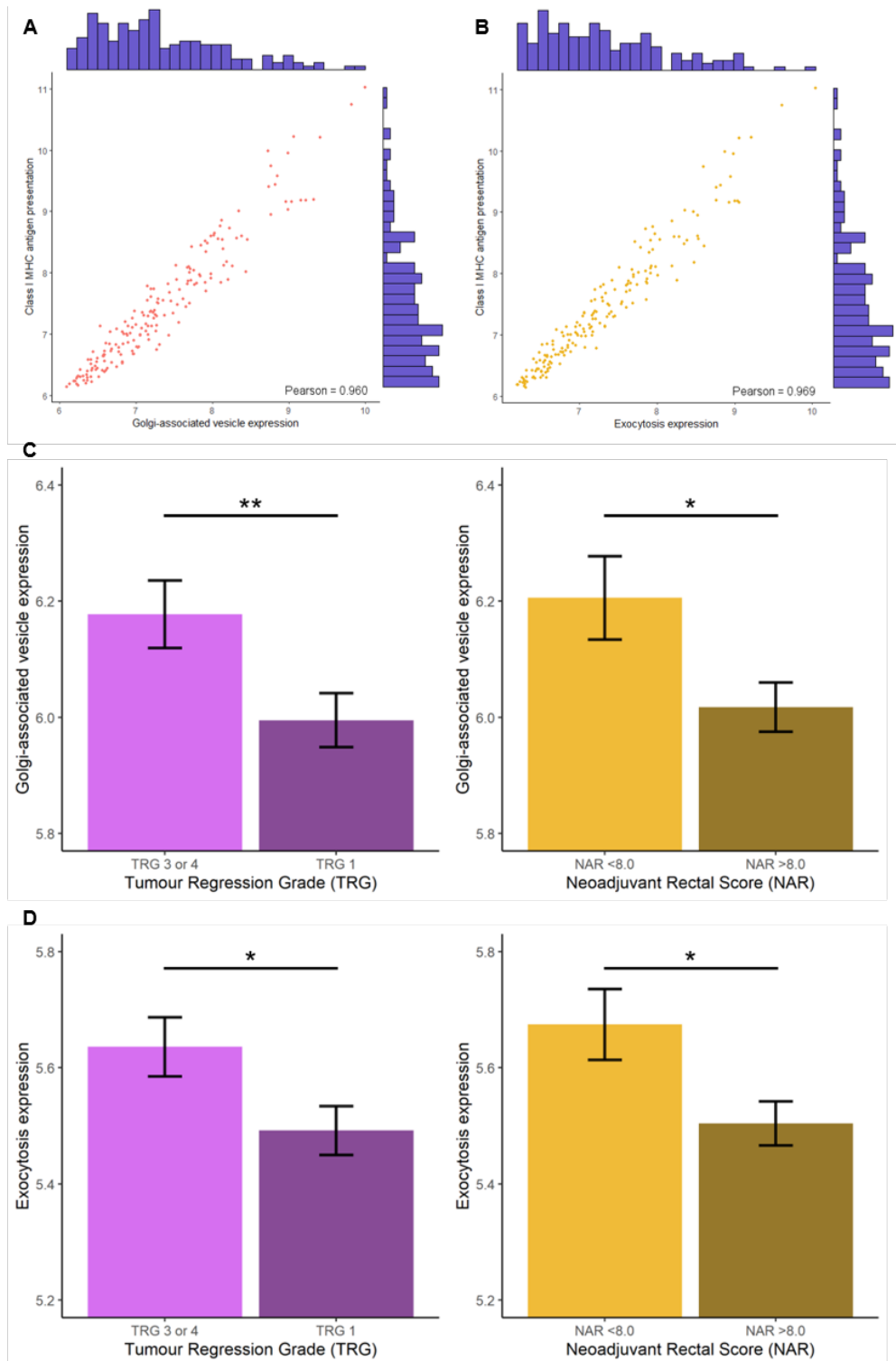


Figure 12: Association of Golgi-associated vesicle and exocytosis gene set expression with MHC class I expression and response to neoadjuvant CRT. A-B) Marginal dot plots showing the association of MHC class I expression with Golgi-associated vesicle (A) and exocytosis (B) gene set expression at a single cell transcriptomic level. C-D) Bar plots showing the relationship between Golgi-associated vesicle formation (C) or exocytosis (D) expression and CRT response. Tumour regression grade or neoadjuvant rectal score determined CRT response. * $p < 0.05$, ** $p < 0.01$.

DISCUSSION

Most studies to date have analysed the host immune response or the tumour epithelial markers in colorectal cancer (CRC) in relative isolation, which may have hindered a more granular analysis of anti-tumour immunity and tumour biology. As part of analysing tumour biology in an epithelial and immune context, I utilised contemporary molecular classifications of rectal cancer including consensus molecular subtyping (CMS) and colorectal intrinsic subtyping (CRIS) ^{45,46}. I have also extended these analyses by developing a novel analysis pipeline to quantify the 7-stage *Cancer Immunity Cycle*, a widely accepted model for characterising anti-tumour immune responses, and functionally characterising processes underlying cancer antigen presentation including vesicle formation (within the golgi apparatus) and exocytosis. I successfully implemented the novel analysis pipeline, also referred to as the Tumour Immunogenomic Report (TIGER), in CRC primary tumours and in the context of neoadjuvant chemoradiotherapy (CRT) response in a cohort of locally advanced rectal cancers.

At the CRC primary tumour level, I have subsequently used the power of TIGER and machine-learning-derived hierarchical clustering to classify CRC tumours based entirely on the Cancer Immunity Cycle for the first time, identifying four subgroups of tumours with distinct immune phenotypes namely hyper-immune, hypo-immune, uninitiated (early-block) and unpropagated (late-block). My immune classification has been linked to epithelial-based CMS tumour classifications with the “hyper-immune” tumours primarily CMS subtyped into phenotypic groups characterised by microsatellite instability (MSI) and increased expression of immune infiltrating cell types. Further analysis of the four *Cancer Immunity Cycle* clusters also identified association of the hyper-immune cluster with MSI phenotype whereas the hypo-immune cluster with a predominantly MSS phenotype. Although corresponding with other features of tumour biology, it is important to note that my *Cancer Immunity Cycle* immune clusters were beyond existing molecular subtypes or tumour features, as none of these were specific or sensitive to each cancer immunity cycle cluster.

Aside from epithelial and *Cancer Immunity Cycle* (CIC) classifications, changes in *Cancer Immunity Cycle* stage and constituent expression profiles were linked to patient and tumour features. In keeping with the literature, MSI tumour phenotype was a major driver of differential immunogenomic profiles and was characterised by increased expression of infiltrating T cells (CIC stage 5) and cytotoxic activity associated with effector CD8+ and natural killer (NK) cells (CIC stage 7) with MSI-high tumours specifically linked to increased class II MHC antigen presentation and CD4+ T cell subset expression. However, most notable was the association of NK and NKT cell expression with MSI phenotype, suggestive of an interaction of MSI phenotype with crosstalk between innate and adaptive immune responses. As a tumour becomes more invasive (i.e., metastatic), it is characterised by lymphovascular invasion as well as invasion into lymph nodes and distant organs. The immunogenomic profiles of

stage 3/4 tumours and/or those with confirmed lymphovascular invasion were markedly different to non-invading tumours with a more immunologically active and inflammatory microenvironment discernible in fully metastatic cancers (stage 4). In addition, my immune profiling of right-sided colonic tumours ascertained increased immune effector cell infiltration and cancer cell killing activity as determined by elevated expression of T lymphocyte trafficking, infiltration, recognition and killing of cancer cells (CIC stages 4-7) when compared to left-sided colonic and rectal tumours as confirmed by increased expression of signatures defining positive regulation of immune responses and immune activation associated with immunologically “hot” tumours ⁴². Conversely, patient survival outcomes were only moderately correlated with differential immunogenomic profiles with activated CD4+ T cell population and T cell trafficking prior to infiltration of effector CD4+ T cells associated with good survival.

Further to patient features, TIGER immunogenomic profiling also corresponded with tumour somatic variation. Mutations in several known cancer-related genes were associated with increased expression of *Cancer Immunity Cycle* stages including PIK3CA and BRAF (CIC stages 5 and 7), and MLH1 and MSH6 (CIC stages 4 and 7) whereas KRAS and APC gene mutations were linked with lower immune activity of CIC stages 2 (cancer antigen presentation) and 5 (T cell infiltration), respectively. Notably, mutations in several of these genes have previously been implicated in immunological changes ^{107,222,230} but have been characterised in the context of the *Cancer Immunity Cycle* for the first time. My subsequent analysis of tumour mutational load (TML), analogous to tumour mutational burden, identified association of high TML with increased expression of innate responses (dendritic cells, NKT cells, innate immune response activation) and adaptive T cell responses (activated CD4 cells, effector memory CD8 cells, gamma delta T cells) as well as T cell infiltration, recognition and killing of cancer cells (CIC stages 5-7) supporting the association of high TML with increased anti-tumour immune activity ¹³¹.

Within my neoadjuvant CRT rectal cancer cohort, I utilised a similar to that of my Tumour Immunogenomic Report (TIGER) whereby I measured anti-tumoural immunity through analysing each stage of the Cancer Immunity Cycle from cancer antigen release to cancer cell killing. My most notable observation is that the degree of cancer antigen presentation (cycle stage 2) is significantly associated with tumour regression. Specifically, for MHC class I presentation, expression differences in this process appears to be modulated and associated with expression activity of basic cell processes underlying peptide trafficking, such as Golgi vesicle formation and vesicle exocytosis. This is important, as cancer antigen presentation via MHC class I is a bridge between tumour cellular biology and the ensuing immune response in a multi-step process occurring within tumour cells. The process is initiated as tumour associated peptides are broken down by the proteasome and results in transportation of peptides bound to MHC class I complexes to the cell surface ¹⁷². As expression changes in peptide-trafficking steps of this pathway were associated with tumour regression, I propose that this innate host capacity for peptide trafficking via Golgi-associated vesicles and exocytosis could be an important determinant of anti-tumoural immune response activation.

Furthermore, I identified that antigen presentation via MHC class II in antigen-presenting immune cells, critical for presenting cancer antigens to prime and activate T cell-driven responses, is enriched in patients with better treatment responses. In combination, our findings have highlighted the importance of cancer antigen presentation with rectal cancer patient outcomes following neoadjuvant CRT. I subsequently provide confirmatory evidence that the degree of lymphocytic infiltration is associated with tumour regression as previously reported^{203–210}. I demonstrated this in the context of the Cancer Immunity Cycle and Th1-centric Coordinate Immune Response Cluster (CIRC) scoring with higher expression of 'T cell priming and activation' and CIRC observed in patients with more complete tumour regression. Finally, I identified enrichment of immune-related signatures including 'inflammatory response' and interferon-gamma signalling' in CRT responders which has recently been reported by Wilkins *et al*²³¹.

There are some potential limitations of my analyses. Firstly, there are differences in tumour cellularity between individual cancers which cannot be ascertained using my analysis stream alone and, as a result, it is not possible to state with certainty that genes enriched in a particular tissue type are directly involved in shaping anti-tumour immunity without subsequent *in-vitro* or *in-vivo* functional validation experiments. This could include validation of observed differences for *Cancer Immunity Cycle* cell subtyping within the TCGA CRC patient cohort. Furthermore, clinical annotation and the mutation spectrum of patients was also not complete for the neoadjuvant CRT cohort and thus I could not describe anti-tumour immune responses at these disease sites in the context of microsatellite instability or individual somatic mutations of cancer-associated genes (e.g., BRAF). However, my analyses utilise experimental approaches not previously used to elucidate the immune microenvironment of colorectal cancers at the primary tumour level and/or within a specific treatment response context such as neoadjuvant chemoradiotherapy.

In summary, my tumour immunogenomic report (TIGER) analysis pipeline provides an opportunity to use RNA expression data to quantify mechanism and functionally characterise the anti-tumour immune response using the *Cancer Immunity Cycle* paradigm. My TIGER analysis pipeline enables a more comprehensive understanding of anticancer immunity in the context of tumour biology and can be utilised to assess the interaction between these processes including epithelial pathway expression, microsatellite instability and somatic variation. Furthermore, I provide evidence of association of host immune profiles with CRT treatment response suggesting that early-stage Cancer Immunity Cycle expression, namely cancer antigen presentation, is strongly associated with patient response to neoadjuvant chemoradiotherapy (CRT) which this has not previously been reported. I therefore propose that *Cancer Immunity Cycle* analysis immune expression profiling in addition to with existing analyses of lymphocyte infiltration can provide a more granular, if not completely different, assessment of the anti-tumour immune response in CRC. Like other genomic or experimental tests such as Immunoscore, CMS or CRIS, its uptake and utility within standard healthcare settings needs to be ascertained. However, it may potentially enable

identification of novel targets for immunotherapy, including targeting cancer antigen presentation pathways, and patients most likely to benefit from treatment.

CHAPTER 3: IMMUNOGENOMIC ANALYSIS OF LOCAL AND DISTANT METASTASES OF COLORECTAL CANCER

INTRODUCTION AND AIMS

Understanding the anti-tumour immune response and interaction with tumour biology is crucial for identifying patients who are likely to respond well to cancer treatment and provide an insight into patient outcomes. In colorectal cancer (CRC), this has frequently been defined in the context of immune cell infiltration, such as T-lymphocyte counts, and immune checkpoint regulation ^{47,144,147,195}. However, tumour biology and the anti-tumoural immune responses may also be considered in parallel through the *Cancer Immunity Cycle*, a 7-stage process initiated by release of tumour neoantigens ^{196,197}. This is followed by cancer antigen presentation, priming and activation of T-cells within the lymph nodes, trafficking and infiltration of T cells into tumours, immune recognition and killing of cancer cells, which in turn leads to more release of tumour neoantigens ^{196,197}. Within this multi-step cycle, individual processes or stages can subsequently interact with tumour biology and vice-versa ¹⁹⁶. Several studies have also shown associations between anti-tumour immune responses and features of tumour biology as well as clinical responses to anticancer treatments ^{31–33,48,49}. However, it has been proposed that for some patients with cancer, different aspects of the Cancer Immunity Cycle may be impaired resulting in an inadequate or impaired anti-tumour response and subsequent tumour progression ¹⁹⁶.

The factors that affect an individual's ability to activate the Cancer Immunity Cycle and an effective anti-tumour immune response are likely to be multifaceted. Primary tumour genomics may drive cancer immunity through the evolution of immune evasion strategies ^{14,47} or an individual's immune profile might shape and direct tumour evolution ²³². The latter of these mechanisms may be determined by variations in germline genetics, the microbiome, the presence of infectious agents, environmental exposures and the use of pharmacological agents ^{233–238}. Differences in the immune response at distant metastatic sites compared to primary tumours have been previously reported ⁶⁸, however, neighbouring 'normal' tumour lymph nodes (LN) are the first site of tumour spread and a site where the initial tumour immune response is initiated. Thus, to gain a greater understanding of the immune system of individuals with locally advanced CRC as well as the epithelial and immune profiles of lymph node metastases, I performed gene expression profiling analyses of sequenced clinically annotated biopsies comprising stage 3 CRC tumours, lymph node metastases and neighbouring tumour lymph nodes.

In the context of secondary CRC metastases and anticancer treatment, a study comprising twenty-four metastatic CRC patients treated with regorafenib and nivolumab identified that patients with lung metastases had response rates to immunotherapy of 50%. This was observed in a cohort of patients with predominantly microsatellite stable (MSS) tumours, a subtype that was previously believed to be immunotherapy resistant ⁴⁸. There is evidence that lung metastases are characterised by a higher inflammatory phenotype compared to brain, bone or liver metastases irrespective of primary tumour origin ⁴⁹, and that T-lymphocyte infiltrates are associated with size of CRC metastases as well as response to anti-epidermal growth factor receptor treatment ⁶⁸. This

underlies the importance of understanding the immunological processes underlying anti-tumour responses, corresponding changes in tumour biology and implications for potential treatment efficacy. However, the underlying biology as to why metastatic CRC patients with lung metastases appear to have a better response to immunotherapy in this cohort is unknown. I therefore aim to describe the molecular subtypes and immunological contexture of lung, liver metastases relative to the primary tumours to understand why differences in response to anticancer treatments might arise.

METHODS

Transcriptomic (RNA sequencing) datasets/patient cohorts

Lymph Node Metastases (METALLIC)

Metastasis and lymph nodes in colorectal cancer cohort (METALLIC) biospecimens were collected from newly diagnosed patients with stage II/III sporadic colorectal adenocarcinoma who had undergone surgical resection. They had received no prior treatment for their disease and were treated at the Oxford University Hospitals between March 2012 and May 2015. All patients had given consent for this study, and this was performed under ethics reference 11/SC/0236. Cases were staged according to the European Society of Medical Oncology (ESMO) guidelines. Patients were excluded from this analysis if there was a previous history of CRC, any predisposing conditions for CRC (such as IBD, polyposis syndromes), patients presenting with synchronous CRC lesions or if there was evidence of inherited CRC predisposition syndromes. The presence or absence of microsatellite instability (MSI) was determined using immunohistochemical staining antibodies to MLH1, MSH2, MSH6 and PMS2. Histological sections were reviewed by a histopathologist to confirm that the tumour was consistent with a diagnosis of colon adenocarcinoma and representative regions were marked on both the primary tissue and metastatic deposits in lymph nodes that contained more than 60% tumour cell nuclei, with less than 20% necrosis. Histopathology review was also performed to identify uninvolved mesenteric lymph nodes.

Distant Metastases

The distant metastasis cohort is comprised of 153 unpaired specimens from 129 patients with metastatic colorectal cancer consisting of 79 liver metastases, 33 lung metastases and 41 primary colorectal tumours. All patients had given consent for this study and samples were obtained under national research ethics committee (NRES) references 11/SC/0236 (South Central – Oxford C, July 2017) and 13/WM/0339 (West Midlands, November 2013). A publicly available distant metastasis cohort, CaArray_notte-00422, was utilised for validation and comprised microarray data of 390 samples from 284 patients including colon cancer (CRC) primary tumours, CRC liver metastases and lung metastases, and adjacent normal tissue at each disease site. The CaArray_notte-00422 patient cohort data was obtained from the NCBI GEO accession database under reference GSE68468 (Affymetrix array HG-U133A).

RNA extraction, sequencing, and data processing

Lymph Node Metastases and Distant Metastases

Specimens were processed and embedded in Formalin-Fixed Paraffin-Embedded (FFPE) blocks, sectioned, and RNA was extracted using the High Pure FFPE RNA Isolation Kit (Roche) following the manufacturer's protocol. Column purified RNA was analysed on the TapeStation (Agilent) for determination of the RNA integrity Number (RIN). Library preparation for 3' RNA sequencing was performed using the Quantseq 3' mRNA seq library kit (Lexogen) and sequencing was performed on the HiSeq 4000 platform (Illumina).

The bioinformatics pipeline for analysis of 3' RNA sequencing reads involved sample quality control, sequence alignment to the human genome assembly and calculation of read counts. Sample quality control was performed before and after sequence alignment to the human genome using FastQC (version 0.11.9) and Qualimap 2, respectively. The lymph node metastasis cohort sequencing data was aligned to the human genome assembly GRCh37 using HiSat2 (version 2.0.4) and the distant metastasis cohort to assembly GRCh38 using STAR (version 2.6.0). Transcript read counts were generated using HTSeq (version 0.6.1 for lymph node metastases, 0.9.1 for distant metastases). HTSeq gene counts extracted from 3' RNA sequencing data were subsequently normalised by performing a $\log_2+1$ transformation using the 'edgeR' R package (version 3.28.1) in R version 3.6.3. RNA sequencing data processing required access to the High Performance Computer (HPC) cluster hosted by the Compute and Storage for Life Science (CaStLeS) infrastructure as part of the Birmingham Environment for Academic Research (BEAR)²³⁹.

CaArray_notte-00422 patient cohort

Affymetrix probe IDs were background corrected and normalised using the 'gcrma' R Package (version 2.56.0), and subsequently converted to gene symbols using the biomaRt R package (version 2.42.0). Affymetrix probe intensities were averaged using the limma R package (version 3.42.2) where multiple probes existed for an individual gene. Potential batch effects between sample or array batches were not directly assessed as Affymetrix batch information at the sample level was not publicly available.

Sequencing batch correction

Sequencing batch effect adjustment for all datasets was performed using the ComBat function built into the 'car' (version 3.0-7), 'survival' (version 3.1-11) and sva (version 3.34.0) R packages to account for sequencing batch effects²⁴⁰. Hugo gene symbols for all cohorts were converted to gene entrez ids using the org.Hs.eg.db R

package (version 3.10.0). Tumour purity within each sample was subsequently assessed using the ESTIMATE R package version 1.0.13²²⁰.

Data analysis and visualisation – RNA sequencing

Differential gene expression and geneset enrichment analysis

I performed Differential gene expression analyses using the DESeq2 R package (version 1.26.0)²¹⁹. Analyses were performed on gene counts using the Wald test method and applying p value correction using the Benjamini-Hochberg method. Significant differences in gene expression were defined as a Benjamini-Hochberg P value of below 0.05. I generated volcano plots for differential expression analyses using the ggplot2 (version 3.3.0) and calibrate (version 1.7.5) R packages. I performed gene set enrichment analyses for 50 Hallmark genesets from the Molecular Signatures Database (MSigDB) using the GSEA software package version 4.0.1 (Broad Institute). Gene list inputs were pre-ranked by log2 fold change from DESeq2 analyses. For analyses beyond DESeq2 differential gene expression and geneset enrichment, gene count normalisation was performed using the 'edgeR' r package (version 3.28.1) in R version 3.6.3.

Hierarchical clustering and heatmap generation

I performed unsupervised hierarchical clustering and heatmap figure generation in R Studio (R version 3.6.1) using the pheatmap (version 1.0.12) and RColorBrewer (version 1.1-2) R packages. The 'complete' clustering linkage method and 'euclidean' clustering distance algorithm were utilised. Heatmaps were also generated using pheatmap where hierarchical clustering was not performed.

Molecular subtyping and co-ordinate immune response cluster (CIRC)

I performed molecular subtyping using two different classifiers. Consensus molecular subtyping (CMS) and CRC intrinsic subtyping (CRIS) were performed using the CMSclassifier (version 1.0.0, random forest machine-learning algorithm) and CRISclassifier (version 1.0.0) R packages, using methods as previously described^{45,46}. I also calculated co-ordinate Immune Response Cluster (CIRC) scores for each sample using methods previously described²¹⁸. In brief, z scores were calculated for the 28 CIRC genes²¹⁸ using the scale() function of R version 3.6.1 (center = TRUE, scale = FALSE) and averaged to generate CIRC scores.

Cancer Immunity Cycle (CIC) Stage profiling

I assessed tumour immune responses by quantifying each stage of the Cancer Immunity Cycle. I achieved this through the quantification of 47 genesets representative of the 7 stages of the Cancer Immunity Cycle¹⁹⁶, or immune regulation and activation and subsequently used to calculate geneset expression scores for each cycle stage. Full information on Cancer Immunity Cycle genesets, also referred to as TIGER (Tumour Immunogenomic Report), are available in Supplementary Tables 1-3.

xCell analysis pipeline and other methods of immune characterisation

For xCell analysis, I inputted normalised, sequencing batch corrected gene counts into the xCell R package (version 1.1.0)²⁴¹. In brief, xCell is a gene-signature based method that enables the characterisation of immune, stromal and tumour microenvironment cell types within individual cancers using bulk transcriptomic data. This is achieved through compilation of gene expression profiles, cell types and gene signatures from multiple studies including ENCODE and the Human Primary Cell Atlas to curate gene signatures for the 64 different cell types with multiple gene signatures defining individual cell types²⁴¹. xCell scores are initially calculated as an average single-sample gene set enrichment analysis (ssGSEA) score of all gene signatures corresponding to the specific cell type which are subsequently transformed to a linear score with a spillover compensation correction that accounts for individual genes within multiple signatures²⁴¹. Implementation of functions from the xCell R package enabled the calculation of fractions of sixty-four immune, stromal and tumour microenvironment cell types within each sample in my analyses.

For other methods of immune characterisation, I calculated geneset expression scores using a pipeline requiring the FSA R package (version 0.8.30) (Supplementary Table 6). Normalised gene counts and genesets in entrez id format were inputted and the geometric mean calculated for each geneset. The genesets quantified are outlined in Supplementary Table 7. I also assigned aggressiveness (proliferative capacity) scores to samples within the lymph node metastasis cohort, which was defined by expression of a cell proliferation geneset (GO: 0008283), and calculated geneset expression using the FSA R package. I categorised lymph node metastases (LNM) as being “aggressive” or “non-aggressive” for each patient whereby a LNM was considered to be aggressive if the metastasis aggressiveness score was higher than that of the paired primary tumour.

Statistical methods

I performed statistical analyses in R versions 3.6.1 and 3.6.3, and the following R packages were utilized: broom, car, digest, dplyr, foreign, ggplot2, MASS, peperr, survival, survminer, sva, tidyverse. T tests and Wilcoxon Rank-Sum tests were performed using the `t.test` and `wilcox.test` base R functions, respectively, for directly comparing gene signature expression between two groups and assessing statistical significance. I fitted patient survival data using the `survfit` function of the 'survival' R package and corresponding Kaplan-Meier plots generated using the `ggsurvplot` function for performing and visualising time-to-event analyses for survival and how this corresponds with differences in gene signature expression .

Data visualisation

Data visualization and generation of all figures utilized several packages within R versions 3.6.1 and 3.6.3. I used functions from the ggplot2 R package to generate bar plots (`geom_bar`) with or without standard error bars (`geom_errorbar`), boxplots (`geom_boxplot`, `stat_boxplot`), dotplots (`geom_dotplot`, `geom_line`), and spaghetti plots (`geom_point`, `geom_line`). I generated marginal plots using the `ggmarginal` function of the `ggextra` R package. I also used the `Stars` (`grDevices`) and `SankeyNetworks` (`NetworkD3`) functions to generate star plots and sankey plots, respectively. I generated correlograms by applying a Pearson correlation coefficient between continuous variables and using the `corrplot` function of the 'corrplot' R package.

RESULTS

Immune and epithelial transcriptomic profiling of CRC lymph node metastases

Lymph nodes (LN) are the first site of CRC tumour spread and is where anti-tumour immune responses are initiated and previous work has identified differences in the immune response at metastatic sites from primary tumours⁶⁸. However, to date, immunogenomic profiling in CRC has almost exclusively focussed on primary CRC tumours and therefore may have limited a more comprehensive molecular analysis into the role of host anti-tumour immune responses. Therefore, to gain a greater understanding of the host immune system and to describe the wider immune architecture of LN metastases, I utilised RNA sequencing data from a clinical cohort comprising 728 clinically annotated biopsies encompassing mesenteric lymph nodes (LN), LN metastases and corresponding primary CRC tumours from 221 patients. Of these, 74 of which had stage 2 CRC, 147 of which had stage 3 CRC and performed gene expression profiling analyses. The baseline characteristics of the lymph node metastasis (METALLIC) cohort are outlined in Table 4.

Firstly, I performed gene expression profiling for normal LN stroma from 74 patients (3 LN per patient) with early stage 2 CRC to discern the drivers of gene expression changes in the uninvolved LN stroma, and therefore the features of an immunologically primed or active lymph node. I analysed a number of previously validated immune genesets representing the cancer immunity cycle to enable a measurement of the degree of immune cell infiltration and activity, including those characterising the Cancer Immunity Cycle. I observed that the major drivers of variability in the LN shown in PC1 were “Activated CD4”, “yDT” (gamma-delta T), “memory B”, and “Tregs” as determined using the principal component analysis method (Figure 13, Supplementary Table 8).

I subsequently sought to identify the extent to which immune responses in lymph nodes were impacted by tumour proliferation. I determined the degree of tumour proliferation from expression of the Gene Ontology (GO) cell proliferation geneset (GO: 0008283), which was used as a surrogate to represent CRC tumour aggressiveness. I observed that primary tumour proliferation was important for driving initiation of immune responses in lymph nodes and that several immune genesets were upregulated in these immunologically primed or active lymph nodes in response to aggressively proliferating primary tumours. This included expression of MHC class 1, immune cells of the lymphoid lineage (activated CD4 and CD8, Th1 and yDT (gamma delta T) cells) and cell types associated with immunological memory (effector memory CD4 and effector memory CD8). I also observed a smaller effect for immune cells derived from the myeloid lineage (Neutrophils, Macrophages, and Eosinophils). (Figure 14, Supplementary Table 9).

Characteristic	Stage 3 CRC (n=147)	Stage 2 CRC (n=74)
<i>Age at diagnosis</i>		
Age median (IQR)	68.0 (58.0-76.0)	75.0 (66.0-80.0)
<i>Gender</i>		
Female	71 (48.3%)	32 (43.2%)
Male	76 (51.7%)	42 (56.8%)
<i>Tumour differentiation</i>		
Well	3 (2.0%)	2 (2.7%)
Moderate	113 (76.9%)	66 (89.2%)
Poor	25 (17.0%)	6 (8.1%)
Unknown	6 (4.1%)	0 (0.0%)
<i>Microsatellite status</i>		
MSI	20 (13.6%)	15 (20.3%)
MSS	87 (59.2%)	57 (77.0%)
Unknown	40 (27.2%)	2 (2.7%)
<i>Mutation status</i>		
BRAF mutation	18 (12.2%)	7 (9.5%)
KRAS mutation	32 (21.8%)	23 (31.1%)
PI3KCA mutation	6 (4.1%)	11 (14.9%)
PTEN mutation	3 (2.0%)	3 (4.1%)
TP53 mutation	14 (9.5%)	6 (8.1%)
None or other	30 (20.4%)	13 (17.6%)
Unknown	62 (42.2%)	24 (32.4%)
<i>Disease recurrence</i>		
Yes (DFS Event)	56 (38.1%)	7 (9.5%)
No (No DFS Event)	91 (61.9%)	64 (86.5%)
Unknown	0 (0.0%)	3 (4.0%)
<i>Death during follow-up period</i>		
Yes (OS Event)	39 (26.5%)	6 (8.1%)
No (No OS Event)	108 (73.5%)	68 (91.9%)
<i>Median follow-up period (days)</i>		
All patients (IQR)	887.0 (655.0-1187.5)	1035.5 (844.25-1250.0)
Patients with death event (IQR)	539.0 (258.0-776.0)	820.5 (350.75-1054.75)
Patients with no death event (IQR)	949.0 (792.75-1332.5)	1044.5 (875.75-1262.5)
<i>Median follow-up period (years)</i>		
All patients (IQR)	2.43 (1.79-3.25)	2.84 (2.31-3.42)
Patients with death event (IQR)	1.48 (0.71-2.12)	2.25 (0.96-2.89)
Patients with no death event (IQR)	2.60 (2.17-3.65)	2.86 (2.40-3.46)

Table 4: Baseline characteristics of the lymph node metastasis (METALLIC) cohort. The cohort comprises primary tumours from stage 2 and stage 3 colorectal cancer patients with corresponding mesenteric lymph nodes and lymph node metastases (stage 3 only or metastases).

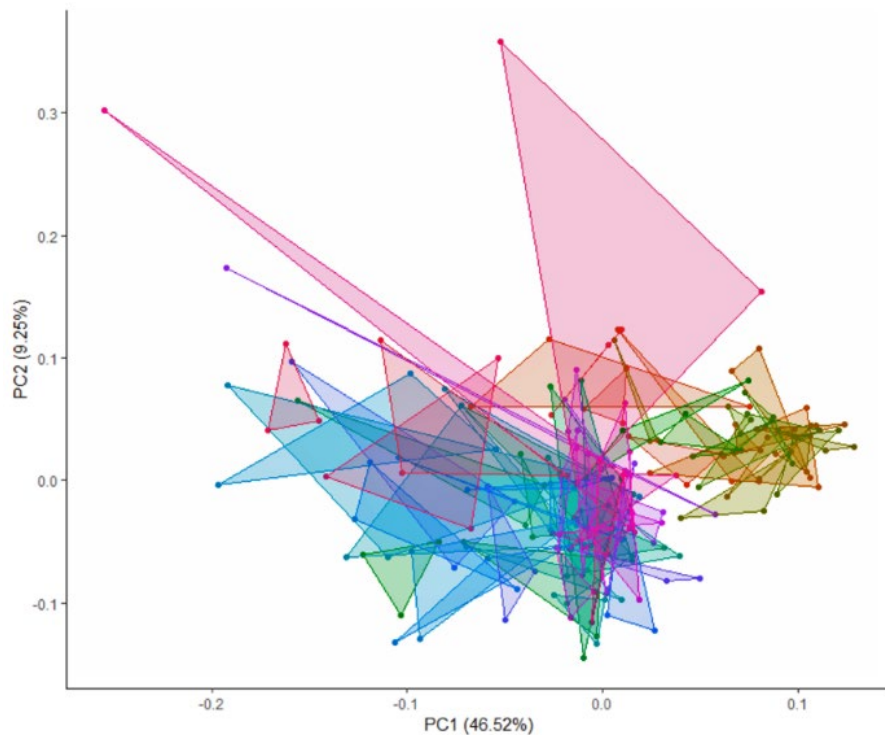


Figure 13: Principal Component Analysis (PCA) of normal CRC lymph nodes. Each coloured dot represents a single lymph node. Colour-coded dots and triangles represent the 3 lymph nodes from a single patient and variation between them based on PCA analyses. Major drivers of variability in the immunogenomic signatures of lymph nodes (PC1) are determined by “activated CD4”, “yDT” (gamma-delta T), “memory B cells”, “Mast cells” and “T-regs”.

The interplay between epithelial biology and tumour metastatic spread has previously been assessed for multiple cancer metastases ^{68,69}. However, to better understand the epithelial gene alternations that drive metastatic spread to lymph nodes (LN) in CRC, I analysed RNA sequencing data from a cohort of stage 3 CRC patients (n=147) with paired primary tumours and LN metastasis biospecimens with the aim of elucidating transcriptomic alternations that characterised lymph node metastases compared to the primary tumours they arose from. I determined tumour purity and immune content using the ESTIMATE R package ²²⁰. Reflecting differences in the tumour microenvironment, I observed differences in immune infiltrate and tumour content in LN metastases (Figure 15).

I then utilised primary tumour and LN proliferation as determined by expression of the GO_CELL_PROLIFERATION gene signature (GO: 0008283) to determine primary tumour and LN aggressiveness. I found that LN metastases not to be uniformly more aggressive than primary tumours, identifying that 42% of LN metastases had a higher aggressiveness score than the paired primary tumour. I then hypothesised that the degree of LN metastasis aggressiveness might be shaped by a combination of tumour biology and patient features including patient age, positive lymph node count, and the presence or absence of CRC-associated mutations (TP53, KRAS, BRAF, PI3K, and PTEN). Of the patient features analysed, I identified that the presence of microsatellite instability (MSI) ($p=0.025$), BRAF mutations ($p=0.003$) and poor tumour differentiation ($p=0.040$) were associated with the development of a more aggressive lymph node metastases relative to the primary tumour (Figure 15).

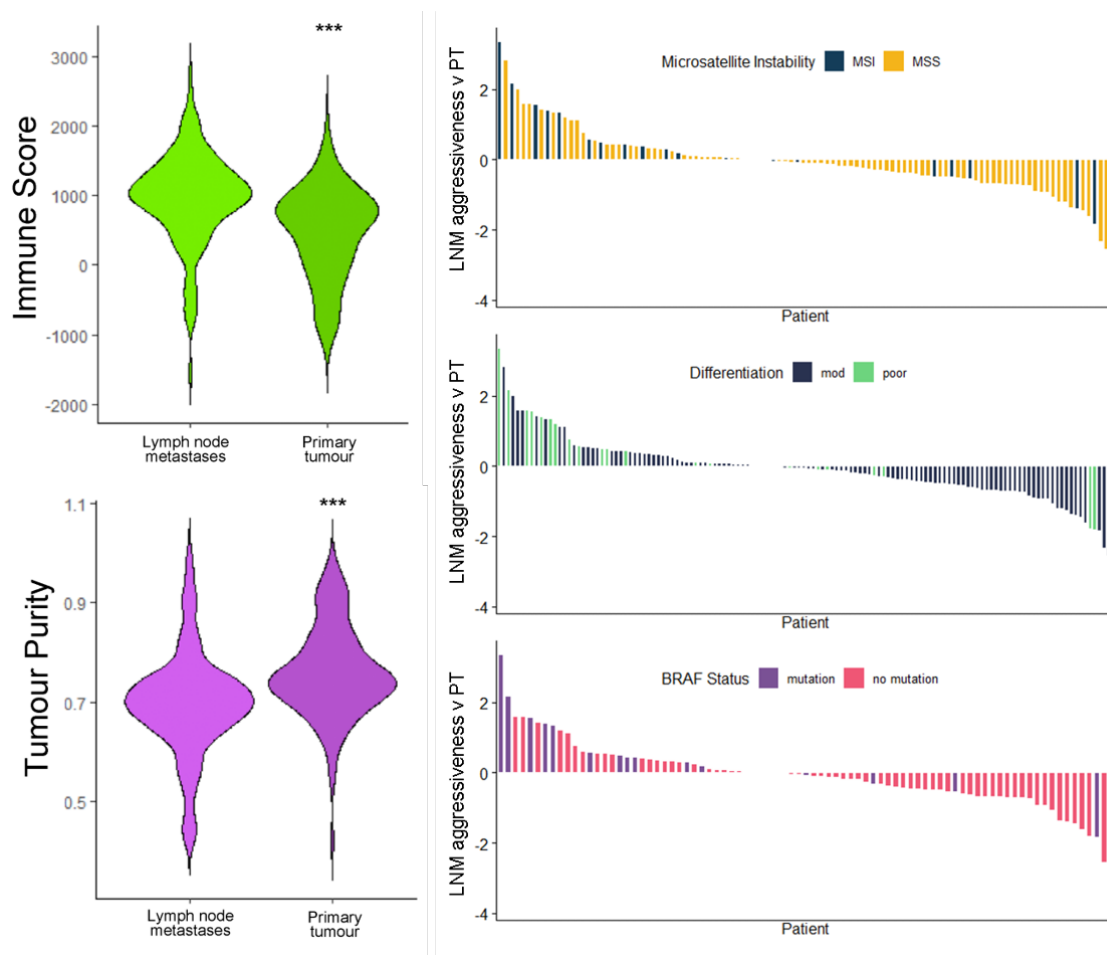


Figure 15: Immune and tumour composition of lymph node metastases and primary tumours (left) and waterfall plot showing aggressiveness of the lymph node metastases compared to the primary tumour (right). Left) Immune and tumour composition were determined using the ESTIMATE algorithm. Lymph node metastases had marginal higher levels of immune infiltrates. Right) The presence of a primary tumour with microsatellite instability, poorly differentiation or a BRAF mutation were significantly associated with more aggressive lymph node metastases. * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

In addition, I identified notable differences at the transcriptomic level in the lymph node metastases compared to their corresponding primary tumours. I observed switching of intrinsic molecular subtypes (CMS and CRIS) of tumours in a considerable proportion of patients (Figure 16) and identified 2527 genes which were differentially expressed in paired LN metastases compared to primary tumours. Of these, 2098 were up-regulated in the lymph node metastases (negative binomial model, samples n=148, adjusted p<0.05, Supplementary Table 10) and several of these have been previously described to play an important role in anti-tumour immune responses including CD72, CLEC4C, ADAM23, TAPBPL and SHC1^{242–246}. Moreover, I observed upregulation of a number of cancer hallmark genesets including MYC, TGF-beta, Notch, Wnt beta-catenin, VEGF-C and PI3K AKT MTOR signalling in LN metastases (Figure 16, all p<0.05).

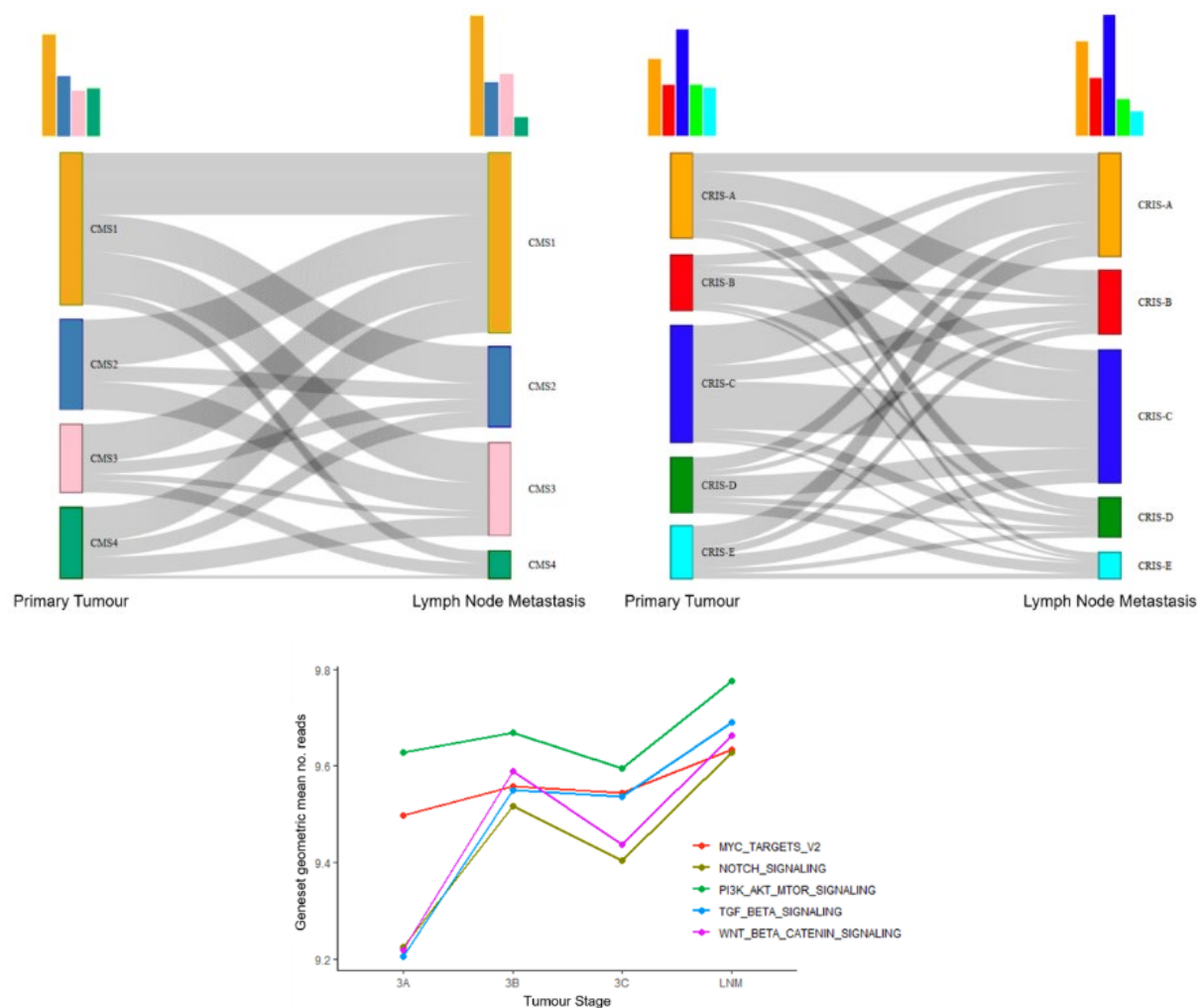


Figure 16: Sankey plots demonstrating molecular subtypes of CRC primary tumours and lymph node metastases (top) and spaghetti plot showing changes in cancer hallmark genesets in lymph node metastases compared to the tumour stage of the primary (bottom). Top) There is switching of CMS and CRIS molecular subtypes of CRC in lymph node metastases. Bottom) Spaghetti plot showing upregulation of MYC, Notch, PI3K, TGF-beta and Wnt signalling in lymph node metastases.

Previous studies have disputed the role of the host immune system in metastasis development. Some studies have demonstrated the importance of an immunosuppressive microenvironment for LN metastasis development^{247–249}, whereas others showed that the immune system might promote the development of metastases^{250,251}. I therefore assessed the tumour immune response within lymph node metastases using the Co-ordinate Immune Response Cluster (CIRC) scoring system, a validated Th1-specific CRC immune scoring system comprised of 28 genes²¹⁸. From this analysis, I found that the CIRC score was higher in LN metastases than primary tumours, suggestive of a more pro-inflammatory microenvironment in the LN metastasis setting (Figure 17).

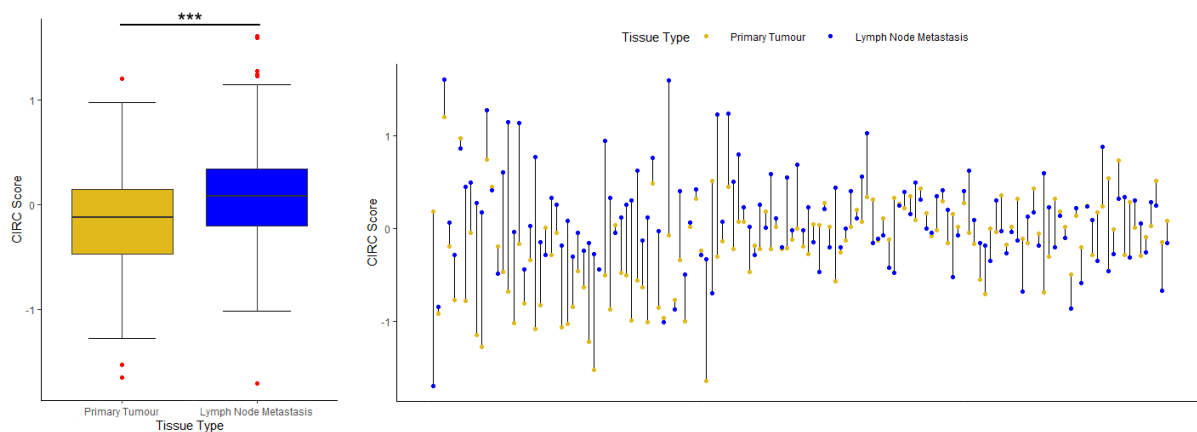


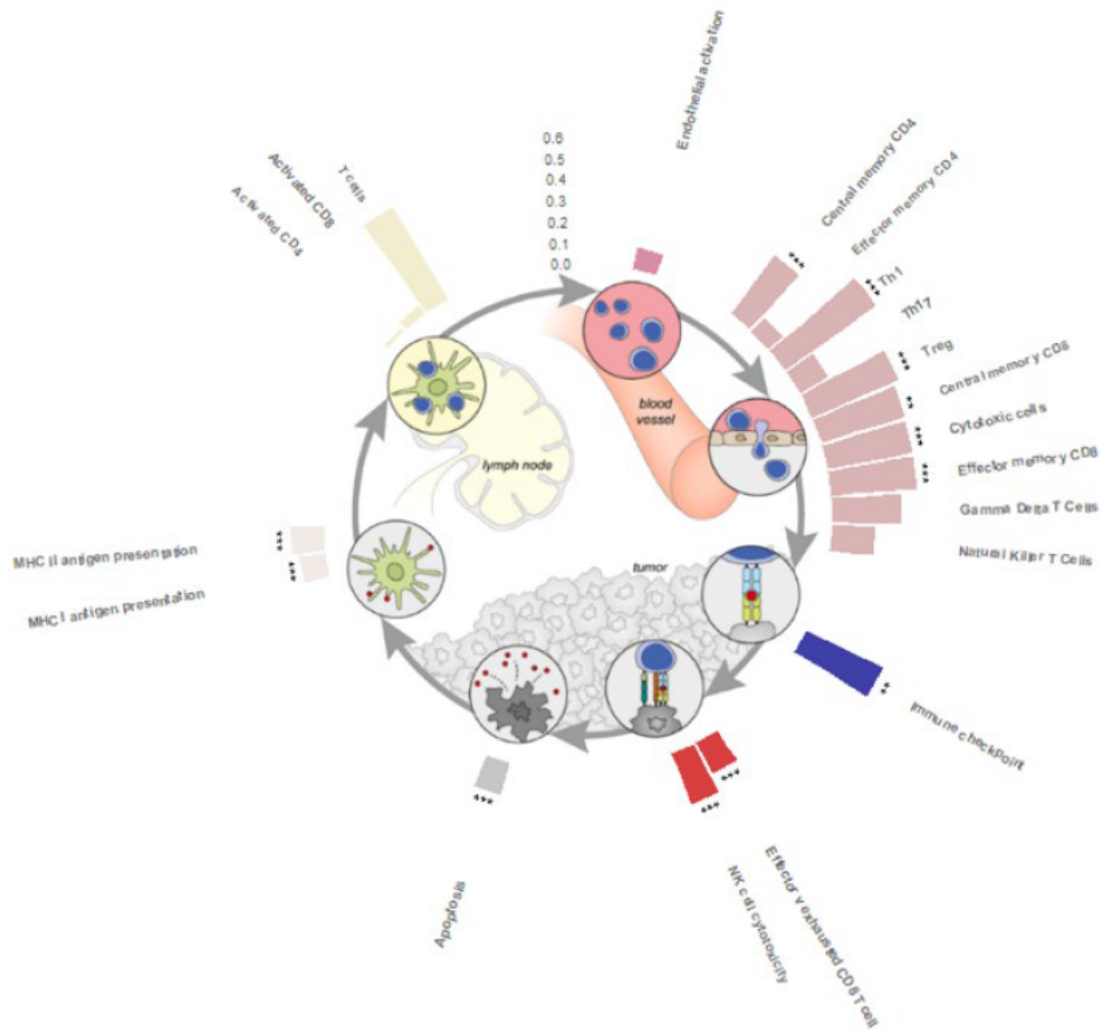
Figure 17: Co-ordinate Immune Response Cluster (CIRC) scores in primary tumours and LN metastases.

*Left) Boxplot of average CIRC score for LN metastases compared to primary tumours ($p < 0.001$). Right) Vertical dotplot of CIRC scores by individual primary tumour and LN metastasis sample pairs. The majority of paired CIRC scores are higher in LN metastases. *** $p < 0.001$.*

Subsequently to gain an understanding of the immune phenotype in LN metastases relative to the primary tumour, I quantified the levels of immune activation of different stages of the Cancer Immunity. I observed that all Cancer Immunity Cycle stages were more active in LN metastases compared to primary tumours, including immune initiation and cancer antigen presentation (stages 1-2), immune propagation (stage 5) and immune activation (stages 6-7) (Figure 18). Having noted these significant differences, I analysed expression of previously validated immune regulation and activation genesets in LN metastases relative to the primary tumour as these processes are known to be responsible for modulating innate and adaptive cytokine immune responses^{147,252}. I identified distinct immunoregulation profiles for LNMs compared to the primary tumour and clustered separately (Figure 19). Furthermore, I observed significantly higher expression of innate and adaptive immune regulation genesets in LNMs, with a lesser, though still significant effect also observed on immune regulation of cytokine responses (Supplementary Table 11).

It is thought that a cancer patient's ability to activate the Cancer Immunity Cycle and an effective anti-tumour immune response is likely to be multifactorial for which two competing theories have been proposed. The first, the Tumour Ascendant model, proposes that the tumour genome is the predominant driver of cancer immunity, namely through the evolution of immune evasion strategies^{14,47} whereas the Immunosurveillance model theorises that a patient's immune profile shapes and directs tumour evolution²³². The Immunosurveillance model implies that metastases develop from a failure of immune control and/or activation at the lymph nodes, enabling tumour establishment²³². If this is the case, there should be evidence of failure to activate anti-tumour immune responses. Therefore, I directly compared the immune profiles of the most aggressive LN metastases to less aggressive ones in patients with stage 3 CRC. From Cancer Immunity cycle stage analyses of aggressive LN metastases, I identified increased expression of immune initiation and antigen presentation genesets (stages 1-2), immune propagation (stage 5) and immune activation (stages 6-7) (Supplementary Table 12). On a single geneset level, I observed higher expression of several immune genesets in the aggressive LN metastases including apoptosis, MHC1, Activated CD4, γ DT (gamma-delta T) cells, immune checkpoint, NK mediated cytotoxicity and evidence of enhanced immunological memory geneset expression (effector memory CD8). Thus, the association of LN metastasis aggressiveness with greater immune activation is contrary to the Immunosurveillance model.

A



B



C

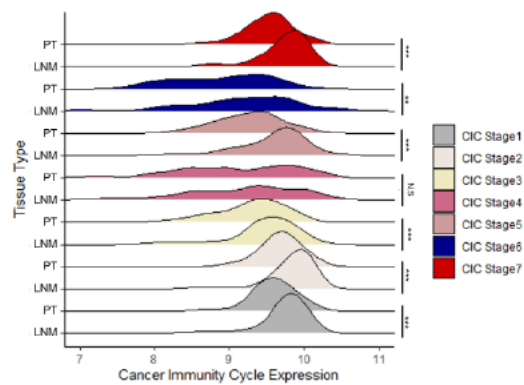
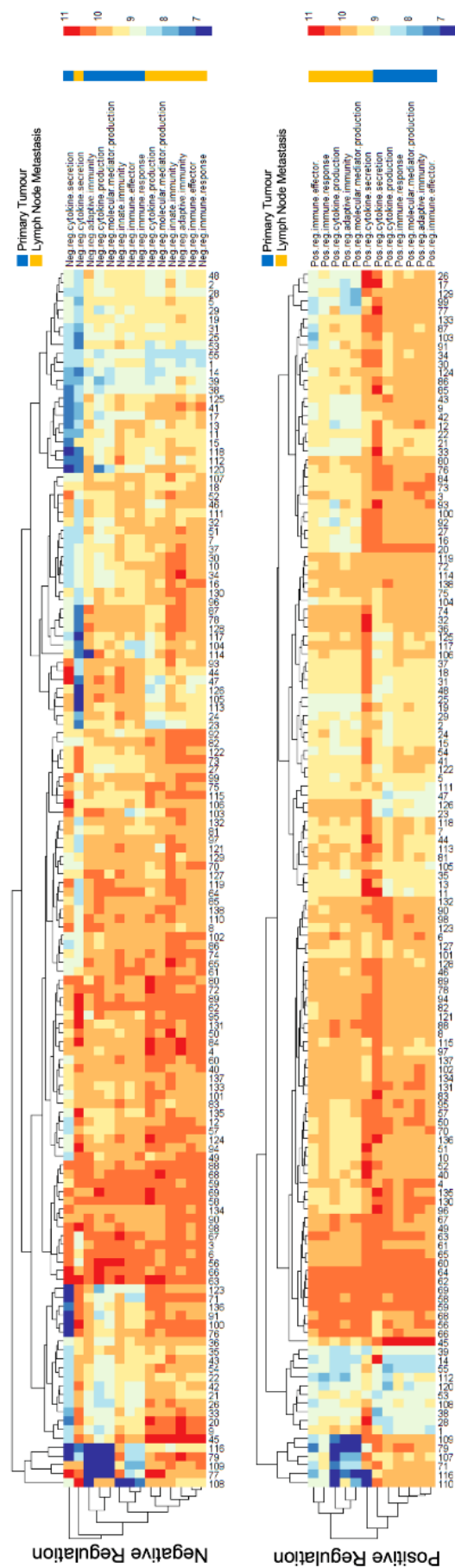


Figure 18: Circular bar chart (A), star plot (B) and ridgeline plot (C) of Cancer Immunity Cycle expression changes in LN metastases compared to the primary tumour. All cycle stages show significantly increased activity in LN metastases except for stage 4. * $p < 0.05$, ** $p < 0.01$, * $p < 0.001$.**

As distant metastases may arise from LN metastases ²⁵³ and/or from a failure of immune activation at lymph nodes, the possibility that patient outcomes from early stage CRCs might be associated with transcriptomic signatures of the draining lymph nodes or LN metastatic lymph nodes was considered. Therefore, I assessed the epithelial and immune gene expression profiles of stage 3 CRC primary tumours, lymph node metastases and normal lymph node stroma to determine association to clinical and patient survival outcomes. At the primary tumour level, I observed that only the high expression of the mast cell immune geneset was associated with poorer disease free survival, confirming evidence from the literature ^{254,255} ($p=0.02$). However, my interrogation of the gene expression profiles of the LN metastases identified that poorer clinical outcomes were associated with evidence of metabolically active tumours (Oxidative phosphorylation $p=0.002$) and higher PI3K-AKT-MTOR signalling ($p=0.039$), findings in keeping with published literature ²⁵⁶.



In addition, in terms of immune genesets, I observed significant association of increased expression of MHC class I antigen presentation ($p=0.029$) with poorer clinical outcomes. From subsequent analysis of the normal lymph node stroma in patients with stage 3 CRC, I also identified that disease free survival and overall survival were associated with immune geneset expression changes. This included MHC class II antigen presentation and the establishment of central memory CD8, with overall survival additionally associated with NK, Treg, Th1, central memory CD8, Cytotoxic Cells gene expression signatures. ($p<0.05$ for all).

Finally, I assessed the possible association of clinical outcomes with activation of different stages of the Cancer Immunity cycle. I found that the degree of activation of the cancer immunity cycle in the primary tumours was not linked to clinical outcomes, however, the degree of antigen presentation in lymph node metastases and in the normal lymph node stroma was linked with clinical outcomes ($p<0.05$). Furthermore, in lymph node normal stroma, there was evidence that better clinical outcomes, namely disease-free survival (DFS) and overall survival (OS) were linked with a higher degree of priming and activation in lymph nodes (stage 3), greater infiltration of T-cells (stage 5) and more cancer cell killing (stage 7) (Figure 20a, Figure 20b).

Disease Free Survival by Cancer Immunity Cycle Stage



DFS Event = Yes



DFS Event = No

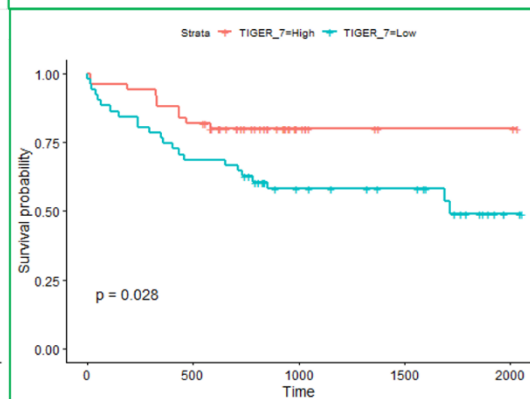
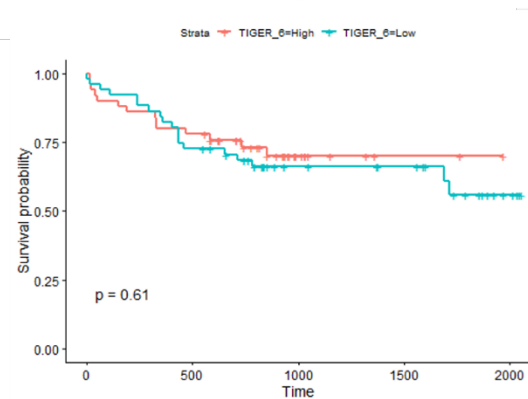
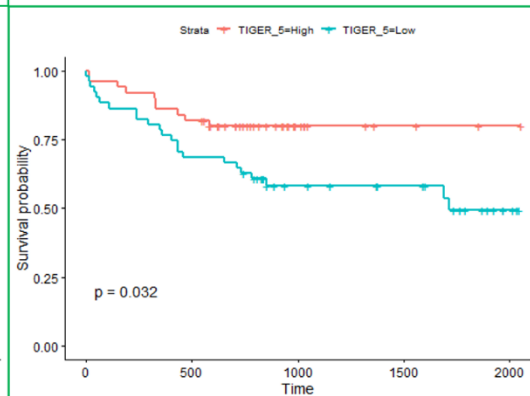
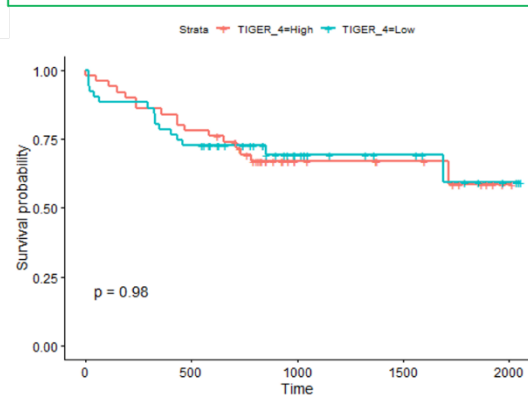
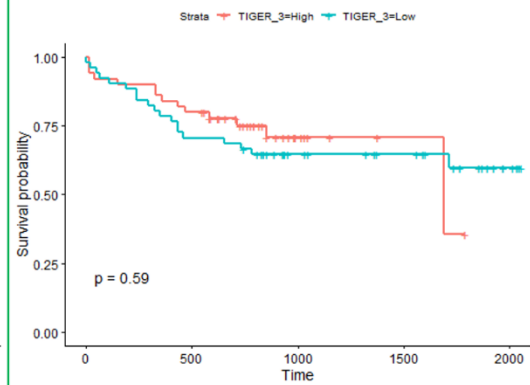
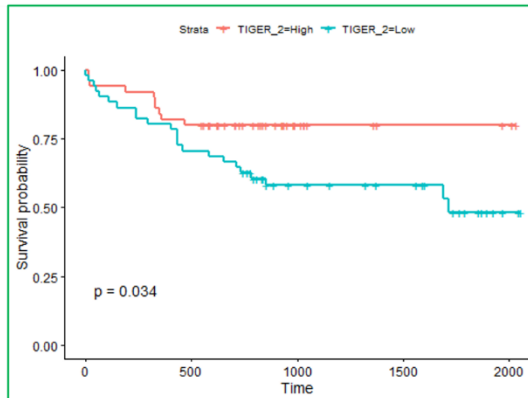
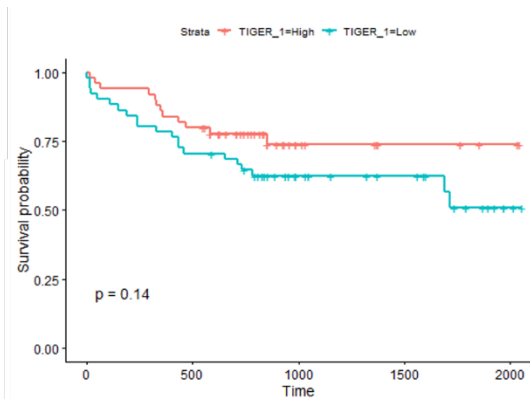


Figure 20a: Kaplan-Meier plot showing effect of Cancer Immunity Cycle stage expression profiles and disease-free survival (DFS, no disease recurrence) in CRC normal lymph nodes. Green boxes denote $p < 0.05$.

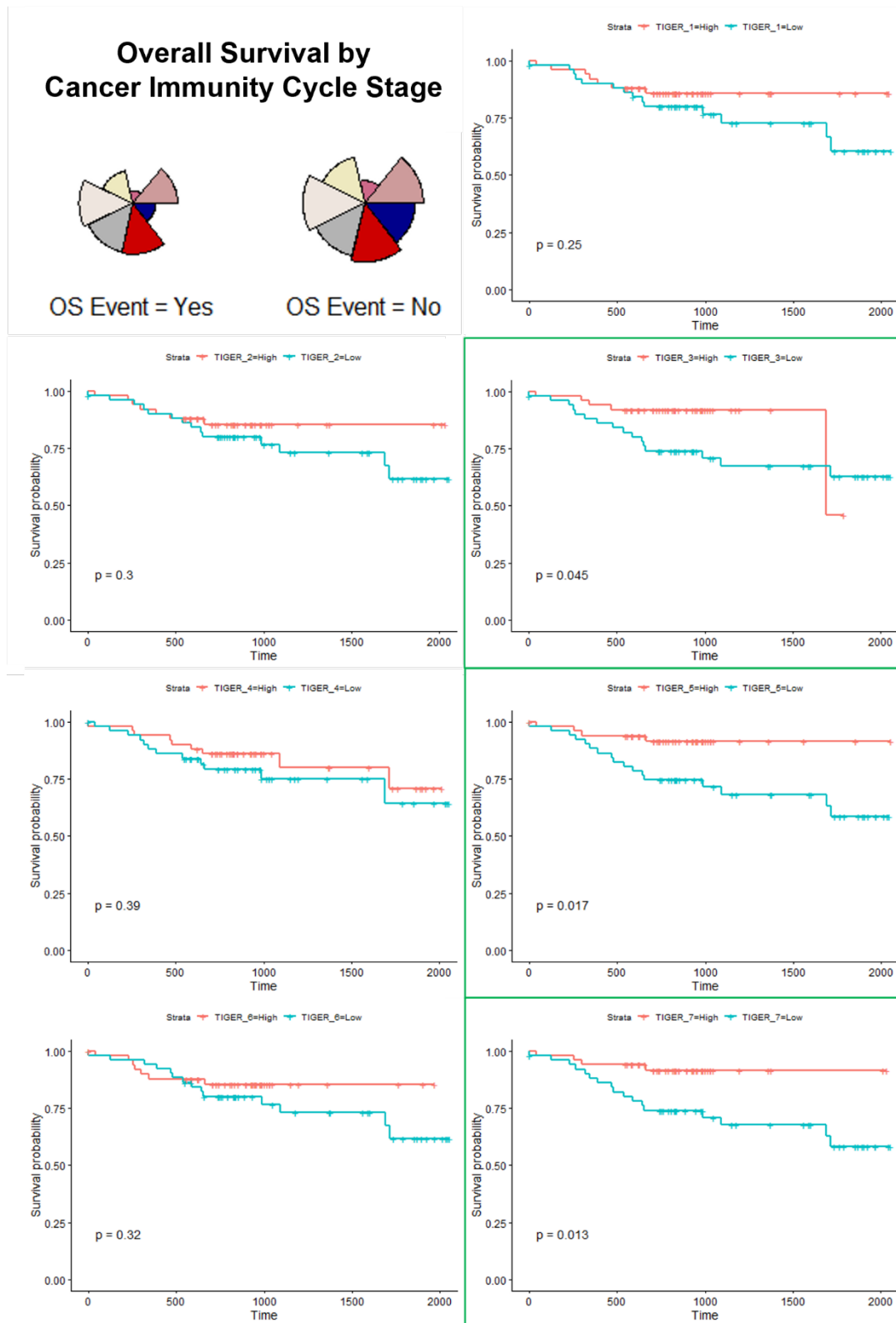


Figure 20b: Kaplan-Meier plot showing effect of Cancer Immunity Cycle stage expression profiles and overall survival (OS) in CRC normal lymph nodes. Green boxes denote $p < 0.05$.

Immune architecture of distant metastases of colorectal cancer

In addition to local metastases of colorectal cancer, the immune architecture of distant metastases has not been fully defined on a transcriptomic level. Distant CRC metastases, including in the liver and lung, are characteristic in patients with advanced CRC and the phenotypic nature of which have been associated with differences in response to cancer treatments such as a better response for patients with microsatellite instability (MSI) or lung metastases treated with immune checkpoint blockade agents ^{34,37,48,49}. However, on an immunogenomic level, the reason for why CRC patients characterised by metastases at different sites appear to respond differently to immunotherapy is unclear. To better understand as to why differences in response to anticancer treatments might arise, I extracted RNA and analysed bulk transcriptomic data from a clinical cohort, termed the 'discovery' cohort, of 153 unpaired specimens from 129 patients consisting of 79 liver metastases and 33 lung metastases and 41 primary tumours from CRC patients with metastatic disease to describe the molecular subtypes and immunological contexture of lung, liver metastases relative to the primary tumours. I also used a previously published gene expression CRC metastasis dataset comprised of 390 samples from 284 patients for experimental validation of findings from the discovery clinical cohort (GSE68468) ²⁵⁷.

Firstly, I performed an analysis of three major molecular subtyping methodologies for CRC to identify epithelial and immunogenomic features at each distant metastasis site, namely consensus molecular subtyping (CMS), colorectal intrinsic signatures (CRIS) and coordinate immune response cluster (CIRC). I observed that a greater proportion of liver (31.0%) and lung (18.2%) metastases were classified as "CMS4" compared with CRC primary tumours (9.8%) whereas a lower proportion of distant metastases were classified as "CMS3" (Figure 21a), of which the proportion of liver metastases classified as "CMS4" was significantly higher than CRC primary tumours ($p=0.012$, Wilcoxon rank-sum test). This result is in keeping with previous observations of a higher proportion of metastatic samples classified as "CMS4" characterised by epithelial-mesenchymal transition (EMT) ²³⁰ and lower proportion of metabolic-phenotype "CMS3" phenotype in CRC metastases ²²². However, I observed no significant differences from CRIS subtyping between distant metastases and primary CRC tumours (Figure 21b). Most interestingly, I observed that the expression of CIRC comprised of twenty-eight immune-related genes (predominantly Th1) was significantly higher in liver metastases and lung metastases compared to CRC primary tumours (Figure 21c), indicating a more pro-inflammatory microenvironment in the metastatic setting, and corresponding with previous findings for lymph node metastases.

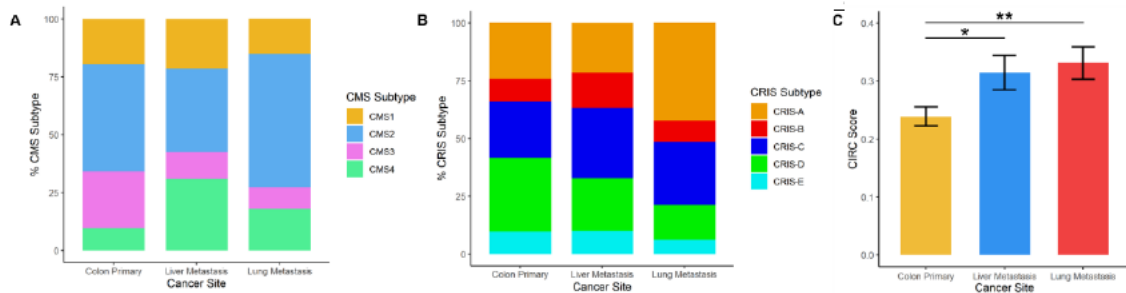


Figure 21: Barplots showing the distribution of molecular subtypes by cancer site. A) Consensus Molecular Subtype (CMS). B) Colorectal Intrinsic Signatures (CRIS). C) Co-ordinate Immune Response Cluster (CIRC). * $p < 0.05$, ** $p < 0.01$, * $p < 0.001$. Statistical comparisons for panel C were performed using Wilcoxon rank-sum tests.**

To determine the immune and stromal compositions of CRC primary tumours and distant metastases, I then performed immunophenotyping using the xCell analysis pipeline in order to deconvolute the tumour and immune microenvironment of CRC primary tumours and secondary metastases which may correspond to previously reported differences cancer treatment responses. Implementation of xCell enables this through calculating the proportions of 64 immune and stromal cell types from bulk transcriptomic data from individual cancers²⁴¹ (Supplementary Table 5 for more information about the 64 cell types). From my xCell analysis, I observed no discernible differences in the proportions of tumour infiltrating lymphocytes, including CD8, CD4, NK and B-cells, in liver or lung metastases compared to primary tumours (data not shown). In subsequent analysis of the expression levels of CD44, a T-cell activation marker and GZMB, a serine protease expressed by NK cells and cytotoxic T lymphocytes, I did not observe any significant differences between the primary tumours and metastases. However, I identified that expression of CD69, a T cell activation marker, was more highly expressed in lung metastases compared to primary tumours ($p = 0.040$). Overall, the tumour infiltrating lymphocyte profiles were similar for CRC primary tumours, liver metastases, and lung metastases.

Having identified that proportions of TILs were not significantly different, I used the principal component analysis method to identify possible drivers of immunophenotypic variation between primary tumours, liver metastases, and lung metastases. I inputted the xCell fractions of all sixty-four immune and stromal cell types for each primary tumour and metastasis with variation outputted using the PCA method. From this analysis, I found that two PCA axes (PC1 and PC2) accounted for 44.24% of sample variation based on xCell immune and stromal compositions (Figure 22). Moreover, I found that the PC1 axis accounting for 32.38% of the immunophenotypic variation was principally driven by different classes of dendritic cells including activated DCs, conventional DCs and immature DCs (Supplementary Table 13), thus possibly indicating importance of immune cell lineages involved early in anti-tumour immune responses (and the cancer immunity cycle) for characterising the immune microenvironment in CRC primary tumours and distant metastases.

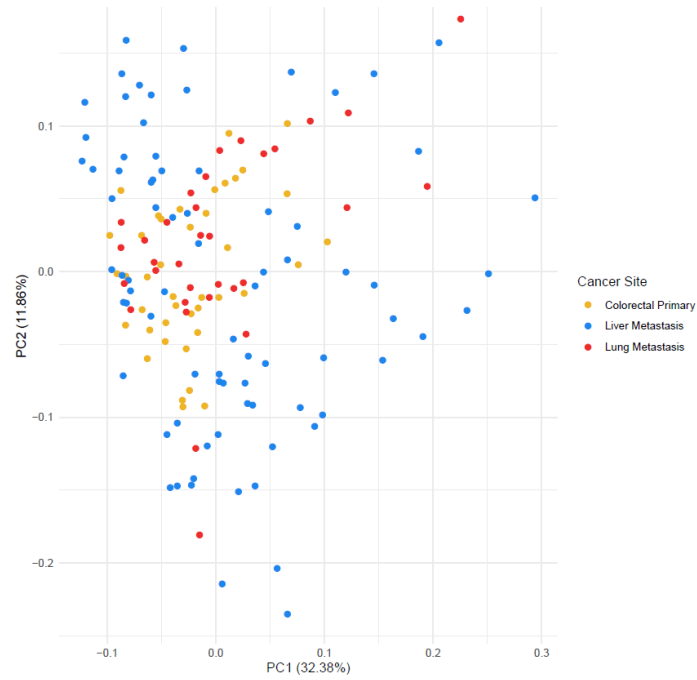


Figure 22: Principal components Analysis (PCA) plot visualising primary tumour, liver metastases, and lung metastatic patterns of expression of immune cell gene expression signatures identified by xCell immunophenotyping.

After identifying dendritic cells as a major driver of variation, I assessed the proportions of dendritic cells in primary tumours and distant metastases using the xCell analysis pipeline. I found that lung metastases had significantly higher proportion of activated dendritic cells (aDC) compared to CRC primary tumours (Figure 23). To confirm this finding, I utilised an independently published metastatic transcriptomic dataset (GSE68468) for validation and showed a significantly higher proportion of aDCs in lung metastases in this dataset (Figure 23).

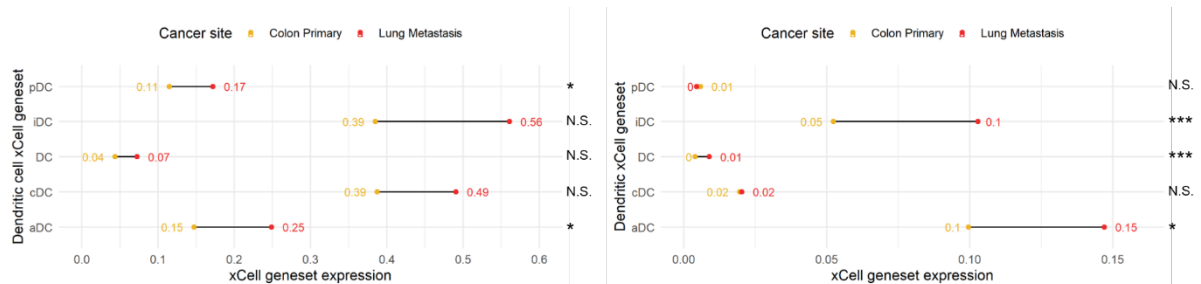


Figure 23: Horizontal dot plots showing differences in dendritic cell (DC) subsets in CRC primary tumours and lung metastases. Left panel correspond with discovery cohort, right panel with validation cohort. * $p < 0.05$, ** $p < 0.01$, * $p < 0.001$. Statistical comparisons were performed using Wilcoxon rank-sum tests.**

Following the finding that the proportion of activated DCs was significantly higher in lung metastases, I compared the relationship between expressions of dendritic cell signatures with priming of tumour infiltrating lymphocytes (TILs) between liver metastases and lung metastases. The dendritic cell expression profiles of lung metastases were significantly different to colon primary tumours and liver metastases, noting significant changes in several DC gene signatures (Figure 24). This included increased DC surface marker expression in lung metastases compared with other tissues and baseline steady-state migratory DCs (prior to maturation) compared with liver metastases. Furthermore, I observed increased expression of inflammatory dendritic cells in lung metastases compared to liver metastases, which in turn corresponded with increased pro- inflammatory IL6-JAK-STAT3 signalling (Figure 24) and therefore possibly indicative of more inflammatory microenvironment within lung metastases compared to liver metastases.

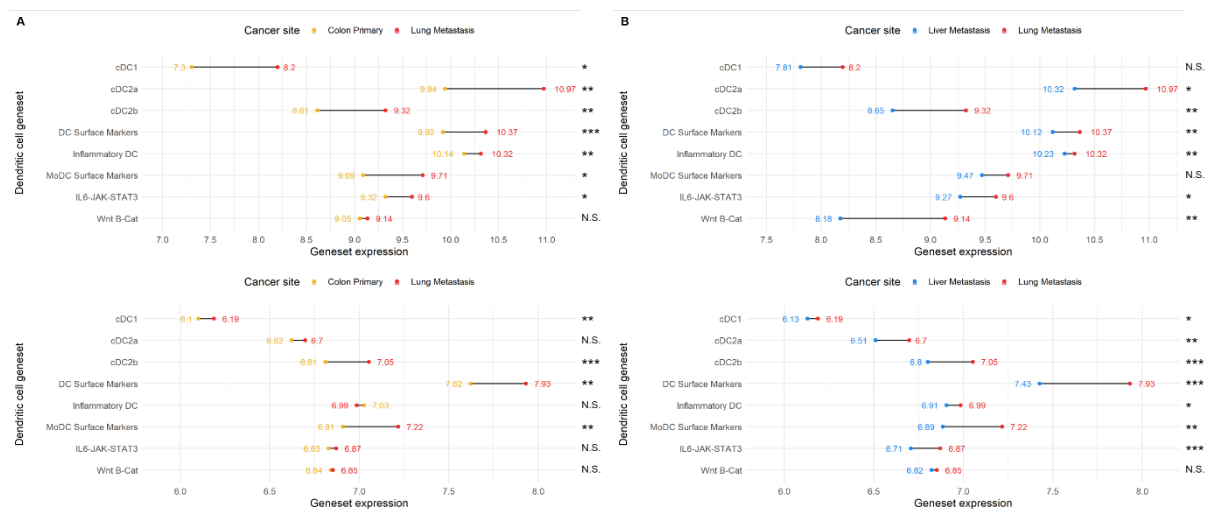


Figure 24: Horizontal dot plots showing average expression of dendritic cell subsets and associated pathway genesets. A) Lung metastases v CRC primary tumour. B) Lung metastases v liver metastases. Top panels correspond with discovery cohort, bottom panels with validation cohort. * $p < 0.05$, ** $p < 0.01$, * $p < 0.001$. Statistical comparisons were performed using Wilcoxon rank-sum tests.**

In keeping with this observation, I also observed increased expression of inflammatory DC precursors, monocyte derived DCs (MoDCs), in lung metastases compared with CRC primary tumours. However, expression of the Wnt signalling pathway associated with immune exclusion from tumours²⁵⁸ was not significantly perturbed in distant metastases. Most interestingly, signatures comprising conventional DC subsets implicated in CD8+ (CDC1) and CD4+ (CDC2A/2B) T cell priming and activation^{259,260} were also enriched in lung metastases compared with the primary tumours (CDC1 & CDC2B) and liver metastases (CDC2A/2B), perhaps indicating a

more immune-active microenvironment where different dendritic cell subsets drive priming and activation of immune effector cells.

In the context of immune checkpoint blockade, Garriss et al proposed that interleukin 12-producing (IL-12) type 1 conventional dendritic cells (cDC1s) are required to sustain the anti-tumour immune response through priming of CD4+ and CD8+ T cells and that anti-PD1 immunotherapy can drive this process through facilitating interferon gamma immune responses (i.e. priming the immune response against PD-L1)²⁵⁹. Therefore, I further interrogated expression of cDC1 signatures in lung metastases and molecular pathways driving DC activity enabling T cell priming. I found that expression of a gene signature comprised of cDC1 surface markers²⁶¹ was significantly higher in lung metastases compared with both advanced CRC primary tumours and liver metastases (Figure 25a). Furthermore, I observed that IL-12 signalling facilitating DC and T cell crosstalk, and NFkB signalling associated with DC activation was enriched in lung metastases compared to liver metastases, corresponding with increased proportions of activated DCs in lung metastases (Figure 25b-c). However, interferon-gamma (IFN-g) signalling involved in DC activation was not significantly perturbed in lung metastases (Figure 25d), perhaps indicating a greater role for NFkB signalling for DC activation. Overall these findings indicate increased DC presence and activity in lung metastases, and corresponds with metastatic CRC patients with lung metastases responding well to immune checkpoint blockade, namely anti-PD1 therapy²⁵⁹.

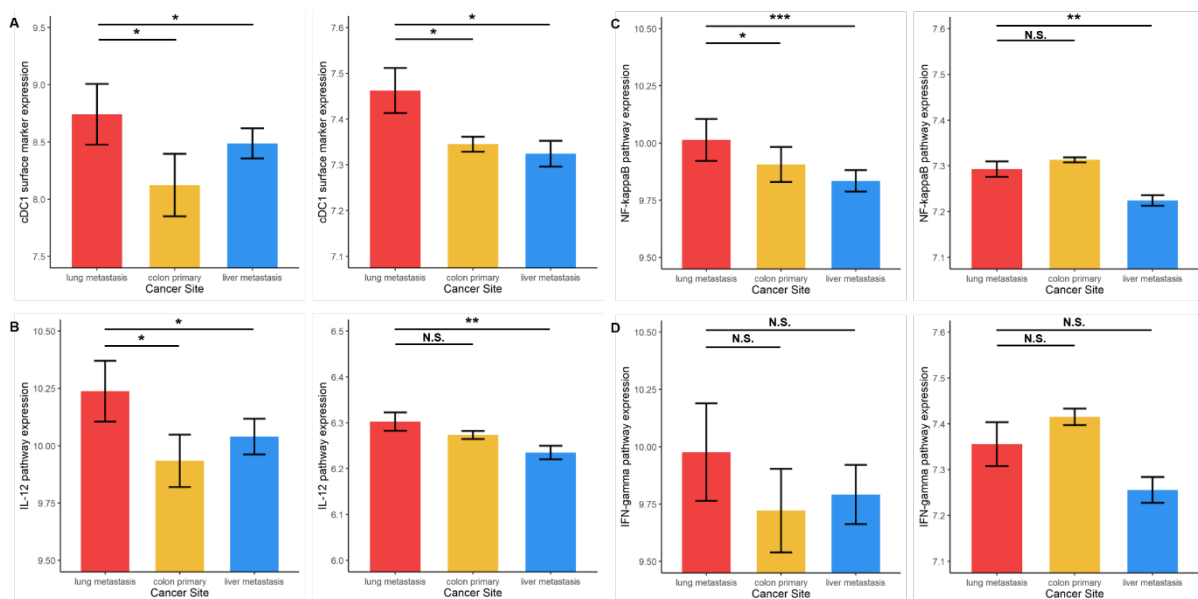


Figure 25: Barcharts showing differences in cDC1 surface marker (A), IL-12 pathway (B), NF-kappaB pathway (C) and IFN-gamma (D) geneset expression at lung metastases compared to the primary tumour or liver metastases. Left panels correspond with discovery cohort, right panels with validation cohort. * $p < 0.05$, ** $p < 0.01$, * $p < 0.001$. Statistical comparisons were performed using Wilcoxon rank-sum tests.**

The increased expression of dendritic cell populations and associated markers in lung metastases is also suggestive of increased dendritic cell activity including antigen presentation. To determine if this is the case, I determined expression levels of several genesets defining dendritic cell activity and antigen presentation and compared expression in primary tumours and distant metastases. In lung metastases, I found that expression of dendritic cell process genesets was significantly higher including “Dendritic cell differentiation,” “Dendritic Cell Antigen processing and presentation” and “Dendritic cell migration” in both metastatic transcriptomic datasets (Figure 26).

These findings corresponding with increased expression of mature dendritic cell types (e.g., CDC1 and CDC2) in lung metastases and implications for T cell priming as well as a higher base level of non-mature steady-state migratory DCs. Moreover, in keeping with these observations, a number of class II MHC antigen presentation genesets used as a surrogate to describe antigen presentation by “professional antigen-presenting cells” (e.g., DCs) were also enriched in lung metastases compared to liver metastases or primary tumour. These included genesets defining positive regulation of class II MHC, and class II MHC protein binding, complex expression, and receptor activity (Supplementary Table 14).

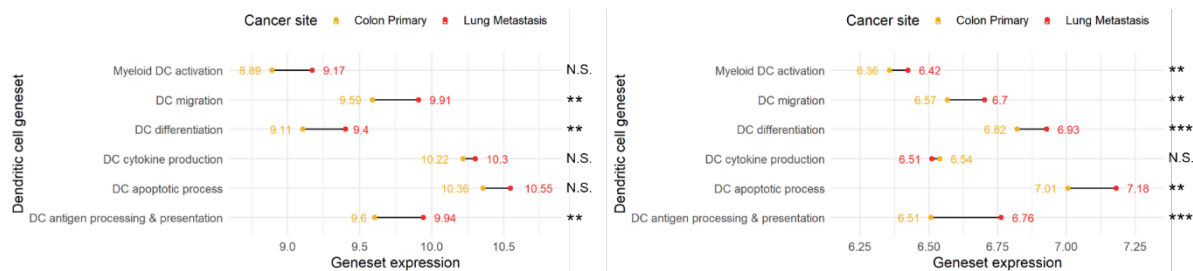


Figure 26: Horizontal dot plots showing differences in dendritic cell (DC) processes in CRC primary tumours and lung metastases. Left panel correspond with discovery cohort, right panel with validation cohort. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Statistical comparisons were performed using Wilcoxon rank-sum tests.

Having identified that lung metastases were characterised by a higher expression levels of dendritic cell process and antigen presentation, my next step was to determine if this was a metastasis-specific phenomenon, or a tissue-specific effect observable in normal lung tissue. I obtained normal tissue gene expression data from the colon, liver, and lung from The Cancer Genome Atlas (TCGA) and used this for comparative analyses. With the exception of MHC class Ib protein binding ($p = 0.007$) and MHC class II biosynthetic processes ($p = 0.009$), which had higher baseline expression in normal lung tissues compared to normal gastrointestinal tissue, I found that differences in expression levels of other MHC class 1 or 2 genesets were not significantly marked between these normal tissue specimens. Furthermore, I observed no significant differences in expression of any dendritic cell

genesets between these normal tissues, suggesting that elevated DC populations are a characteristic of lung metastases in CRC.

From these focussed analyses of expression of cell type by tissue site, I identified notable metastasis-specific differences in dendritic cell activity and antigen presentation. However, I also observed a number of other salient differences with relevance to immunosuppression and immune checkpoint blockade. Most interestingly, regulatory T cells (Tregs) were significantly enriched in liver metastases compared to lung metastases and subsequently validated using the GSE68468 dataset (Figure 27a). As Tregs inhibit T cell cytokine production and have immunosuppressive properties downregulation the proliferation of effector T-cells^{262,263}, this finding may be indicative of a more immunosuppressive environment in liver metastases.

In addition, I identified significant differences in the expression of immune checkpoint signalling by tumour site. Lung metastases expressed significantly higher levels of PD1 and CTLA4 signalling than primary tumours in both the discovery and validation cohorts, a phenomenon I did not observe in liver metastases (Figure 27b-c). Lastly, further analyses of lung metastases also identified increased angiogenesis in these lesions compared to primary tumours or liver metastases whereby the expression of both lymphatic and microvascular endothelial cells ($p=0.029$ and $p=0.009$, respectively) was increased in lung metastases in both metastatic datasets.

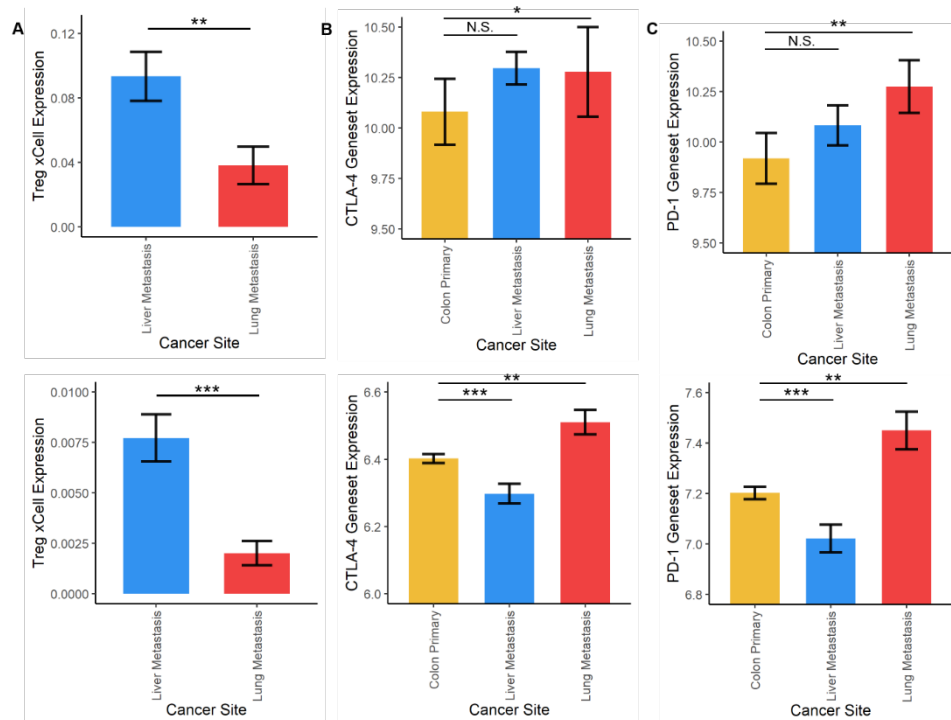


Figure 27: Barcharts showing differences in regulatory T cell (Treg), CTLA-4 and PD-1 expression. A) Treg expression is significantly increased in liver metastases compared to lung metastases. B and c) CTLA-4 and PD-1 geneset expression is significantly increased in lung metastases compared to CRC primary tumours or liver metastases. Top panels correspond with discovery cohort, bottom panels with validation cohort. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Statistical comparisons were performed using Wilcoxon rank-sum tests.

DISCUSSION

The majority of studies to date have analysed the host immune response or the tumour epithelial markers in colorectal cancer (CRC) in relative isolation, which may have hindered a more holistic analysis of anti-tumour immunity and tumour biology. In the context of cancer progression, this is notable as the immune biology of secondary CRC metastases have been described relative to primary tumours⁶⁸ but typically not within lymph nodes in the proximity of primary tumours which are the first site of tumour spread and where anti-tumour immune responses are initiated. Moreover, immune profiling of metastatic lesions associated with advanced CRC have generally considered liver metastases and lung metastases as a single entity^{68,264}, despite a reported link between metastatic profile and response to immunotherapy⁴⁸. It was on this basis that I analysed the immune expression profiles of an early-stage CRC cohort comprising lymph node metastases, primary tumours, and neighbouring lymph nodes as well as advanced stage 4 CRC primary tumours, liver metastases and lung metastases.

Firstly, consistent with the observations I described at the primary tumour level in chapter 2, I observed noticeable interaction of epithelial tumour biology and anti-tumour immune responses in lymph node metastases, frequently the first metastatic site of cancer including CRC. This is notable as lymph nodes comprise reservoirs of lymphocytes and myeloid cells which can be recruited and activated during anti-tumour immune responses as well as typically being the first site of CRC metastatic spread. Corresponding with increasing epithelial proliferation, determined by a cell proliferation gene signature, I observed that an immune response is initiated in “pre-primed” or immunologically “hot” lymph nodes comprising activated MHC signalling and antigen presentation, the development of immunological memory and early features of immune activation. Furthermore, I found that lymph node metastases are not universally more aggressive or have higher proliferative capacity than their corresponding primary tumours and that metastases are characterised by frequent switching of molecular subtypes (CMS), increases in known oncogenic signalling pathways such as MYC, TGF-beta and Wnt, and the formation of an immune-stimulatory microenvironment. Importantly, I showed that these large-scale changes in molecular features occur in stage 3 CRC and therefore at an earlier stage than previously described.

Furthermore, interaction between cancer biology and the anticancer immune response was evident in both the primary tumour and lymph node metastasis contexts whereby more aggressive primary tumours and more aggressive lymph node metastases are both associated with increased immune response expression. These findings are perhaps consistent with a *Tumour Ascendant* model whereby epithelial biology shapes the resulting anti-tumour immune response^{14,47}. That said, I identified that the immunoregulatory profiles of lymph node metastases were markedly different to primary tumours, suggesting that either immune cells in lymph nodes are pre-primed by their microenvironment to respond differently compared to primary tumours and/or that lymph node

immune cell populations have intrinsically different immune capabilities than those found in the primary tumour. This in turn underlies importance of understanding immunological heterogeneity at different metastatic sites and that only a partial understanding of the immune response can be gained by studying immune biology at the primary tumour level.

Following the identification of distinct immune profiles in late-stage primary tumours and local metastases (lymph node), I performed immune profiling of distant liver metastases and lung metastases in metastatic stage-4 CRC patients. Previous molecular studies of liver metastases and lung metastatic lesions have either primarily focused on epithelial gene expression alterations ^{265,266} or in instances where immune infiltration analysis has been performed, liver metastases and lung metastases have been considered individually ^{68,69}. However, CRC metastasis phenotype and anatomical site have been associated with patient response to anticancer treatments whereby patients with lung metastases respond better to treatment with cytotoxic ²⁶⁷ or immunotherapy ⁴⁸ agents which, if validated, would represent a significant shift in the utilisation of targeted treatments. My analysis of the tissue-specific immune architecture of CRC metastases identified higher expression of regulatory T cells in liver metastases whereas lung metastases were characterised by increased expression of activated dendritic cells as well as enrichment of the Th1-driven Coordinate Immune Response Cluster (CIRC) signature. Moreover, I identified higher expression of immune checkpoint inhibition gene and genesets and validated in lung metastases with corresponding evidence of higher MHC class II expression activity, implying that higher expression of these markers in lung metastases may determine their response to immunotherapy.

A body of evidence from the literature that tissue-specific differences in immune responses may exist ²⁶⁸, however, the exact underlying reasons for these observations as to why the immunogenomic profiles of lung and liver metastases might differ amongst each other and CRC primary tumours are unclear. I observed that expression of these genesets do not appear to differ in normal liver metastases and lung tissue but arise following development of metastases potentially suggestive that metastatic stromal tissue expresses stimulatory and inhibitory factors which direct both epithelial evolution and immune infiltration in response to metastases. In normal tissues, I found that expression of activated dendritic cells (DC) is greater in lung tissue than other sites, a finding confirmed in lung metastases where markedly increased expression levels of DC migration, DC antigen presentation and class II MHC antigen presentation (expressed on professional antigen-presenting cells) were also observed. The increased MHC class II pathway expression in lung metastases is notable and corresponds with findings that metastatic CRC patients with lung metastases treated with immune checkpoint blockade respond better to treatment ²⁶⁹.

There are some potential limitations of my analyses. Firstly, there are differences in tumour cellularity between individual cancers as well as comparative differences of primary tumours and metastases which cannot be ascertained using my analysis stream alone. This includes the inability of 'bulk' tissue RNA sequencing to differentiate tissue compartments such as the epithelium and stroma. Cellular admixture and spillover of adjacent

normal tissue into also cannot be completely ruled out, even though primary and metastatic tumour regions were carefully selected during RNA extraction. Furthermore, as for all gene expression studies, it is not possible to state with certainty that genes enriched in a particular tissue type (e.g., metastases) are directly involved in shaping anti-tumour immunity without subsequent in-vitro or in-vivo functional validation experiments.

One specific example of bulk transcriptomics presenting a limitation is through the use of a cell proliferation signature (GO_CELL_PROLIFERATION) for defining tumour cell aggressiveness in my lymph node metastases analyses. As each sample may contain some stromal and immune cells, it is possible that this measure of tumour and metastasis aggressiveness is confounded by underlying immune cell proliferation and not solely driven by tumour biology. This potential confounding effect in turn may be more pronounced in MSI tumours which typically have more immunologically active tumour microenvironments and therefore cell proliferation signature expression may consequently be driven by infiltration of proliferating lymphocytes into these tumours. However, protein-level methods exploring the expression of Ki67, a cell proliferation marker ²⁶⁴, within tumour cells (e.g., using immunohistochemistry) and/or measuring proliferation signature expression using spatial transcriptomic methods could be implemented to validate my tumour aggressiveness analyses ^{270,271}.

Beyond assessing the proliferative capacity of tumour cells, spatial transcriptomic methods could also be used to assess the spatial distribution of immune infiltrates and signature expression within primary tumours, lymph node metastases (LNMs) and the areas immediately surrounding each disease site (e.g., peritumour). This would enable geographic validation of my comparison of Cancer Immunity Cycle expression between paired primary tumours and LNMs including the identification of differences in immune signature expression in cells within the tumour and adjacent peritumour, with the intratumoural and peritumoural microenvironments reported to be distinct for colorectal cancers ²⁷². In turn, this may overcome multiple limitations of bulk RNA sequencing analyses including cellular admixtures and potential discrepancies in the relative quantities of RNA sequenced from the tumour and peritumour for each sample. It is also important to note that differences in Cancer Immunity Cycle expression between primary tumours and LNMs could be driven by an immune-exclusionary CRC tumour phenotype and that this could be determined using spatial transcriptomic analyses comparing expression within intratumour and peritumour microenvironments ^{258,271,273}.

It is also important to note that some of the differences in data processing steps between the lymph node metastasis and distant metastasis cohorts. Perhaps most notably, transcriptomic data from each cohort was aligned to different human reference genome versions (GRCh37 and GRCh38) using two different aligner tools (HISAT and STAR). The consequence of this is that the results from each analysis cannot be directly compared with each other due to the reported differences between aligners and human genome assemblies ^{274–276}. As a result, any future direct comparison of the lymph node metastasis and distant metastasis datasets would require RNA sequencing data processing using the same aligner and human genome assembly version.

In addition to data processing, analyses performed using the distant metastasis datasets incorporated unpaired (unmatched) stage 4 CRC primary tumours and secondary metastases with the rationale underlying this approach being that the distant metastasis dataset comprised sequencing data from either primary tumour or metastatic samples for a number of individuals. A large proportion of samples would therefore have been excluded from my analyses using a paired approach within this cohort with this outweighing the benefit of a direct isogenic comparison between primary and metastatic disease using only paired samples. Thus, further evaluation using paired transcriptomic data from primary tumours and secondary metastases to overcome the potential issue of inter-individual variation across primary tumours and metastatic disease may be warranted.

Further to the limitations in data processing differences and unmatched sample comparisons, validation of observed differences in primary tumour and lymph node metastases in other patient cohorts may also be required, most notably for comparative analyses between stage 3 primary tumours and lymph node metastases. Clinical annotation and the mutation spectrum of patients was also not complete for the liver metastases and lung metastasis patient cohort and thus I could not describe anti-tumour immune responses at these disease sites in the context of microsatellite instability or individual somatic mutations of cancer-associated genes (e.g., BRAF). However, I have validated observations within my analysis of distant metastases with a corresponding stage 4 CRC bulk transcriptomic dataset. More importantly, my analyses utilise experimental approaches not previously used to elucidate the immune microenvironment of local and distant metastases of stage 3 and 4 CRCs.

In summary, I provide evidence of the importance of greater understanding of colorectal cancer tissue-specific immune microenvironments. My interrogation of anti-tumour immunity in early stage CRCs potentially shows that the immune response is shaped by tumour biology and an understanding of lymph node biology, particularly around cancer antigen presentation will be crucial in gaining a more comprehensive understanding of the anti-tumour immune response. Furthermore, evidence of increased dendritic cell infiltration and class II MHC expression in lung metastases may present a mechanism of immune pre-initiation which is deterministic of better CRC patient response to immunotherapy, a finding which may warrant further functional analyses. Taken together, immune expression profiling of CRC metastases may potentially enable identification of novel targets for immunotherapy and patients most likely to benefit from treatment including through the modulation of cancer antigen presentation.

CHAPTER 4: OVERVIEW OF THE CHARACTERISTICS OF SARS-COV-2 AND IMPACT OF COVID-19 AT THE POPULATION LEVEL

COVID-19 in people living with cancer in early 2020 – a brief overview.

Coronavirus disease 2019 (COVID-19) is an acute respiratory syndrome caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus which was first identified in China's Hubei province in late 2019 ²⁷⁷. Since the start of the COVID-19 pandemic, over 676 million cases and 6.88 million COVID-19 related deaths worldwide reported by the Johns Hopkins University (JHU) coronavirus resource centre as of 10th March 2023 when data collection was halted (<https://coronavirus.jhu.edu/map.html>) ²⁷⁸. COVID-19 disease presentation and course following a SARS-CoV-2 infection is variable with many individuals developing only mild symptoms or none at all whereas others develop severe disease characterised by respiratory and immunological sequelae, such as acute respiratory distress syndrome (ARDS), requiring inpatient hospital and/or intensive care admission ^{279–281}. The prevalence of severe disease was also identified to be overrepresented in already clinically vulnerable individuals such as the elderly and those with other comorbidities including people living with cancer with the importance of suppressing SARS-CoV-2 transmission strongly emphasised ^{282–286}. Clinical guidelines and social isolation policies were subsequently implemented in healthcare settings and the community to suppressing SARS-CoV-2 transmission ^{277,287,288}. Nevertheless, these observations were reported from small cohorts or population subgroups thereby highlighting the urgency of evaluating the impact of COVID-19 disease in larger patient cohorts including the monitoring of outcomes of individuals with cancer who contract COVID-19.

For people living with cancer, emerging evidence in early 2020 reported high susceptibility to SARS-CoV-2 infections when compared to the wider population and elevated risk of severe disease, intensive care admission and death ^{289–292}. This was first reported by *Liang et al* in February 2020 in a cohort of 1590 patients admitted to hospital with COVID-19 of which 18 had a cancer diagnosis ²⁸⁹. Nine (50.0%) of these patients with COVID-19 and cancer had a severe disease presentation, including three out of four (75%) of patients who had received recent chemotherapy or surgery, in comparison to 245/1572 (15.6%) of individuals without a diagnosis ²⁸⁹. These results were also corroborated in other small-scale studies, with lung cancer diagnoses overrepresented in the cancer cohorts of all studies ^{289–291}. A higher risk of severe COVID-19 following recent antitumour treatment reported by *Zhang et al.* in a cancer cohort comprising 28 individuals, with the authors recommending that immunosuppressive anticancer treatments should be stopped, or dosages reduced ^{291,292}. Guidelines for managing patients with cancer in healthcare settings were rapidly altered as a result of these early studies, including the reduction of anticancer treatment in hospital settings and cessation of systemic anticancer therapy (SACT) treatment in some instances ^{293,294}. However, changes to cancer disease management and treatment on the basis of such small scale studies was controversial, with more granular assessments of the impact of COVID-19 required in larger cohorts comprising individuals with cancer ^{295,296}.

Considering this early evidence describing the impact of COVID-19, understanding the biology of the SARS-CoV-2 virus, disease transmission (spread) and courses, and outcomes of infected individuals therefore became paramount better protect clinically vulnerable individuals and the wider population. National health and research organisations in the UK and worldwide have sought to shed light on all these key areas among others with a notable focus on protecting vulnerable individuals such as people living with cancer. Here, COVID-19 within the United Kingdom (UK), features of SARS-CoV-2 viral biology and links to therapeutic development, disease phenotype at the population-level and people living with cancer will be explored.

A brief outline of COVID-19 in the United Kingdom in 2020

The first cases in the United Kingdom (UK) were recorded on 30th January 2020 with 22,241,942 SARS-CoV-2 positive tests and 229,272 SARS-CoV-2 related deaths, defined as COVID-19 mentioned on the death certificate, reported as of 31st August 2023 (<https://coronavirus.data.gov.uk/>)²⁹⁷. In the absence of therapeutic interventions when SARS-CoV-2 first emerged in the UK, significant changes were mandated to suppress SARS-CoV-2 transmission on a population level. This included policies to significantly restricting social interactions and re-allocating of healthcare resources to prevent infection spread in primary and secondary care settings^{298–301}. As part of this strategy, the National Health Service (NHS) Test and Trace Programme was set up to record and manage SARS-CoV-2 transmission in the community and healthcare settings^{302,303}.

Population-level testing of individuals presenting with symptoms associated with SARS-CoV-2 infection was implemented, namely of individuals with a new persistent cough, fever (temperature greater than 38.5°C) and/or loss of smell and taste, to identify cases^{302,303}. Asymptomatic testing in healthcare settings and the community was subsequently introduced, as infected individuals not presenting with symptoms have also been shown to shed the SARS-CoV-2 virus and thereby infecting other individuals^{302,304–306}. As a result, provision of both symptomatic and asymptomatic testing has been maintained in some healthcare settings and for severely immunocompromised individuals^{302,307}. These measures have subsequently aided health systems and scientific research community to better understand disease transmission while also reducing disease spread, though clinically vulnerable individuals remain at elevated risk including people living with cancer.

Biology of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)

SARS-CoV-2 transmission and structure

The causative agent of COVID-19, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), is a *Betacoronavirus* ³⁰⁸. Though the origin of the virus is disputed, SARS-CoV-2 is proposed to be of zoonotic origin and is closely related to the SARS-CoV-1 and MERS-CoV viruses which caused the severe acute respiratory syndrome (SARS) and Middle East respiratory syndrome-related (MERS) coronavirus epidemics, respectively ^{308,309}. It is a highly contagious pathogen which is transmitted through transfer of respiratory droplets during coughing and sneezing, either directly or indirectly via contact with contaminated surfaces ^{306,310}. This is a consequence of viral shedding during SARS-CoV-2 infection in symptomatic and asymptomatic individuals, primarily within sputum and/or from the upper and lower respiratory tract ^{310,311}, which facilitates transmission. To this end, multiple studies conducted in the early stages of the COVID-19 pandemic identified detectable viral shedding by infected individuals at least seven days following infection, though can be detected several weeks after symptom onset in some individuals ^{310,311}. Furthermore, individuals with severe SARS-CoV-2 infections are reported to have higher viral loads and have a longer viral shedding time ³¹¹. In light of viral shedding and mechanism of transmission, strict infection control measures such as deployment of personal protective equipment (PPE) and social distancing or isolation measures were therefore proposed by the World Health Organisation (WHO) during the early stages of the SARS-CoV-2 in 2020 and implemented by governments worldwide ³¹².

Structurally, SARS-CoV-2 is a positive-sense single-stranded RNA (ssRNA) virus which is characterised by spike glycoproteins that enable viral binding to the surface of human cells via the angiotensin-converting enzyme 2 (ACE2) receptor ^{313,314}. ACE2 receptors are near-ubiquitously expressed on cells within the human body including cells of the upper and lower respiratory tract and lungs, thereby providing a key entry point for SARS-CoV-2 infection at the cellular level ^{314–317}. The SARS-CoV-2 virus also comprises envelope glycoproteins that are important for fusion of the virus to the host cell membrane during viral replication, membrane proteins which facilitate anchoring of spike proteins to the envelope and nucleocapsid protein containing the ssRNA viral genome ^{313,314}. An RNA-dependent RNA polymerase (RdRp) subsequently drives transcription of viral RNA following infection of host cells, enabling viral replication and propagation ³¹⁸. Viral replication is followed by packaging of the replicated ssRNA genome and translated viral proteins within a nucleocapsid in the host cell cytoplasm, formation of infectious virus particles (virions) in the endoplasmic reticulum (ER)–Golgi intermediate compartment and release from the cell via exocytosis ³¹⁸.

The genome of SARS-CoV-2 viral genome has been fully sequenced and was first made available on 10th January 2020 by the Chinese Center for Disease Control and Prevention to the scientific research community worldwide ³¹⁹. It is approximately 30 kilobases (kb) and comprises 14 open reading frames that encode structural

proteins and non-structural proteins which are involved in the virus life cycle or replication ³¹⁸. The four key *SARS-CoV-2* structural proteins (spike, envelope, membrane, nucleocapsid) are encoded by the c-terminus of the viral genome whereas the remainder encodes non-structural proteins involved in viral replication and processes within the host ³¹⁸. This includes the RNA-dependent RNA polymerase and helicase encoded by non-structural proteins 12 and 13, respectively (nsp12/13) as well as proteins that are involved in evading the host immune system including Nsp15 and Nsp16 ^{320–323}. In turn, these components provide viral structural integrity, enable viral infection of host cells and replication which subsequently facilitates transmission ^{318,320–323}. Therefore, elucidating the biology and genomics of *SARS-CoV-2* is important for understanding the impact of viral infections in human populations and identifying potential therapeutic targets.

Physiological impact and implications of *SARS-CoV-2* infection in humans

The biology and genomics of *SARS-CoV-2* have significant implications for infection course and severe acute disease in humans. Firstly, entry of the *SARS-CoV-2* virus through binding of ACE2 receptors of cells in bronchial and alveolar epithelia of the lower respiratory tract results in the onset of characteristic symptoms of COVID-19 such as persistent coughing and fever ^{314–316}. This infection is reported to occur through a clathrin-mediated endocytic pathway upon the *SARS-CoV-2* virus binding to the extracellular domain of cell-surface ACE2 receptors, enabling viral entry and release of genomic components into the host cell cytoplasm prior to viral replication ³²⁴. Infection of cells in the lower respiratory tract can subsequently result in viral entry into the peripheral bloodstream through the lungs and onset of viraemia where virus particles are circulating in the blood ³²⁴. This process can result in significant epithelial and endothelial cell death, causing vascular leakage and production of pro-inflammatory cytokines and chemokines which in turn trigger inflammation and increase vascular permeability ^{325–328}. Within the respiratory tract, viraemia and resulting inflammation has the potential to lead to more adverse sequelae, such as shortness of breath, which in the most severe cases can develop into acute respiratory distress syndrome (ARDS) or acute phase pneumonia requiring intensive care and/or oxygen treatment ^{279–281,327,329}. Thus, the mechanism of *SARS-CoV-2* entry via ACE2 enables viral spread throughout the respiratory system and into the bloodstream with immune dysfunction and respiratory distress observable in individuals with severe infections ^{279,326}.

Further to triggering and amplifying respiratory sequelae, the onset of viraemia can consequently lead to systemic infection of cells and organs expressing ACE2 such as the heart, liver and gastrointestinal (GI) tract ³¹⁶. This is theorised to result from *SARS-CoV-2* infection of ACE2-expressing cells such as epithelial cells and absorptive enterocytes of the ileum and colon ³¹⁶. In turn, this can be reflected in the disease phenotype observed in individuals with *SARS-CoV-2* infections. GI symptoms such as loss of appetite, diarrhoea, nausea and vomiting have been frequently observed at the population level with multiple cohort studies reporting relatively high symptom

presentation in infected individuals as well as being more prevalent in individuals who are admitted to hospital or critical care with severe COVID-19 ^{316,330,331}. Abnormal liver function has also been reported in individuals with more severe disease which is characterised by changes in skin pigmentation as a consequence of COVID-related liver injury resulting from systemic inflammation ^{316,332}. Aside from the GI tract and hepatobiliary system, several cardiovascular pathologies following a SARS-CoV-2 infection have also been reported ranging from chest pain, heart palpitations and arrhythmia to myocardial injury and organ failure in the most severe instances ^{314,333}. This is thought to be a consequence of SARS-CoV-2 infection of ACE2-expressing myocytes and pericytes resulting in cellular damage and local activation of pro-inflammatory immune responses in addition to systemic inflammation, resulting in organ injury and dysfunction ^{333–335}. Moreover, SARS-CoV-2 binding to ACE2 can downregulate the degradation of angiotensin II whereby presence of excessive angiotensin II can activate the p38 Mitogen-Activated Protein Kinase (MAPK) pathway that can result in thromboembolic events (blood clots) ^{333,336}. Thus, it's unsurprising that multiple studies have observed that individuals presenting with cardiovascular symptoms were more likely to require admission to hospital or intensive care as well as having poorer mortality outcomes from acute COVID-19 ^{333–335}.

In the context of cancer, upregulation of ACE2 expression in lung cancer tissue has been reported in multiple studies ^{315,337}. This enrichment of ACE2 expression is hypothesised to be driven by upregulation of multiple epigenetic factors and contribute to slowing down tumour growth through mas receptor activation via catalysis of angiotensin (1-7) peptides and resultant inhibition of tumour-promoting MAP kinases ^{338,339}. As ACE2 acts as the entry point for SARS-CoV-2 infection, increased expression may increase susceptibility to infection and disease severity. Indeed, multiple studies reported overrepresentation of COVID-19 disease in lung cancer patients compared to other cancer types during the early stages of the COVID-19 pandemic ^{289,290,292}. Beyond lung cancer, overexpression of ACE2 has been observed in colorectal tumours compared to other solid tumour subtypes within RNA sequencing data from The Cancer Genome Atlas (TCGA) ^{340–342}. This overexpression is reported to result from hypomethylation of the ACE2 gene promoter ^{340–342} and shown to correspond with changes in immune cell infiltrates in CRC tumours including lower levels of CD8+ T cell infiltrate ³⁴⁰ known to be a key driver of cancer cell killing ¹⁹⁶. While this change in ACE2 expression does not significantly affect CRC disease prognosis ³⁴⁰, a more severe COVID-19 disease phenotype has been observed in CRC patients in a small-scale study comprising 25 patients ³⁴². This severe disease phenotype is noted to be characterised by a higher likelihood of lymphopenia, higher C-reactive protein levels and increased respiratory rate when compared to propensity score-matched COVID-19 patients without a cancer diagnosis ³⁴². Taken together, this indicates an association of ACE2 expression observed within some cancers with changes in tumour biology and anti-tumour immunity as well as susceptibility to SARS-CoV-2 infection and COVID-19 disease severity.

Beyond the varying physiological impact and presentation of acute phase COVID-19, a notable proportion of individuals report ongoing symptoms and health issues several weeks or months following a SARS-CoV-2 infection often termed as long COVID-19 or post-acute COVID-19 syndrome ^{343–346}. Several definitions for the condition exist with the World Health Organisation defining long COVID-19 as occurring *“in individuals with a history of probable or confirmed SARS CoV-2 infection, usually 3 months from the onset of COVID-19 with symptoms and that last for at least 2 months and cannot be explained by an alternative diagnosis”* ³⁴⁷. Long COVID-19 can impact multiple organ systems such as the respiratory, cardiovascular, gastrointestinal and neuropsychological systems with commonly observed symptoms including post-viral fatigue, shortness of breath and impaired cognitive function which can continue to have an impact on everyday life ^{343–345}. Moreover, individuals hospitalised with COVID-19 are at higher risk of experiencing dysfunction of the endocrine system with higher rates of hyperglycaemia and diabetes mellitus observed despite no pre-existing history or diagnosis of diabetes mellitus ^{348,349}. Among those who are at a higher risk of experiencing long COVID are the elderly, females and/or those with existing comorbidities, with obesity and smoking also considered to be risk factors ³⁴⁴.

On a physiological level, the disease course and presentation of long COVID is likely underpinned by multiple pathogenic mechanisms of which continued SARS-CoV-2 virus presence, immune dysfunction and unrepaired tissue damage have been theorised as contributing factors ^{333,350–353}. The first of these, persistent viral infection, may result in chronic inflammation and immune dysfunction through long-term elevation of pro-inflammatory interferons and cytokines as well as by reactivating other viruses within the body including the Epstein-Barr virus (EBV) ^{354,355}. Furthermore, autoantibodies generated as a result of autoimmune dysfunction and circulating in the bloodstream have been hypothesised to drive and/or exacerbate respiratory symptoms experience during long COVID by some individuals ^{355,356}. In combination, this evidence demonstrates that the physiological impact of SARS-CoV-2 infection goes beyond acute disease with long term health implications.

SARS-CoV-2 viral evolution

Further to understanding the biology of SARS-CoV-2 and its physiological impact in humans during the initial stages of the pandemic, both have evolved throughout the course of the COVID-19 pandemic. Coronaviruses, like most RNA viruses, evolve rapidly through gene-level alterations arising from mutation accumulation and/or genetic recombination that can result in observable changes in viral transmission, disease phenotype or presentation and immune evasion or escape ^{357,358}. Some of these changes in the SARS-CoV-2 genome can be introduced during viral replication in the form of replication errors or by host-mediated genome editing by host innate cell defence mechanisms such as the D614G mutation in the spike protein which enhances viral

transmission ^{357–360}. Furthermore, recombination events can combine genetic changes from genetically distinct *SARS-CoV-2* lineages to create novel hybrid variants, with evidence that individuals can be infected with multiple *SARS-CoV-2* variants simultaneously ³⁵⁷.

The first notable *SARS-CoV-2* variant, the Alpha (Pango lineage B.1.1.7) variant of concern (VOC), was identified in the UK in December 2020 with the Beta (B.1.351) and Gamma (P.1) VOCs reported shortly afterwards in South Africa and Brazil, respectively ³⁶¹. Each of these comprised genetic changes which increased transmissibility and/or immune escape including mutations in the gene encoding the *SARS-CoV-2* spike protein which mediates viral entry into host cells ³⁶¹. Subsequently, the Delta VOC (B.1.617.2) was recognised in May 2020 and quickly replaced previous VOCs as the predominant circulating variant globally with a notable uptick in infections ³⁵⁷. Subsequently, the first Omicron VOCs (BA.1-BA.5) emerged in southern Africa in late 2021 and subsequently enhanced transmissibility further, initiating new waves of *SARS-CoV-2* infections worldwide ³⁵⁷. These variants all comprised growth advantages over previous VOCs. Multiple hypotheses exist for the emergence of new *SARS-CoV-2* variants of concern, however a number of genetic changes observed between emerging VOCs appear to have evolved independently of each other in different parts of the world, indicating convergent evolution ^{357,361}.

In addition to changes in transmission, *SARS-CoV-2* viral evolution has also been reported to have implications for disease severity, with a trade-off indicated between transmission and disease virulence ³⁵⁷. Evidence of this has been reported when comparing the co-existence of Omicron and Delta lineages circulating within the population whereby individuals infected with Omicron presented with less severe disease ³⁶². This is thought to be reflective of higher infection rates of Omicron within individuals who had previously been infected with earlier *SARS-CoV-2* variants and an intrinsically lower virulence ^{362–365}. While this in part could be attributed to prior exposure to *SARS-CoV-2* through previous infection and/or vaccination, other explanations of this change in disease severity associated with Omicron infections. These are reported to include reduced fusogenicity (binding or fusion affinity) of the spike protein to the host cell during viral entry, resulting in less tissue damage, and a greater affinity for infecting the upper respiratory tract as opposed to the lungs ^{366–369}. However, this may not necessarily be true of any new *SARS-CoV-2* VOCs or lineages that emerge in the future ³⁶². Therefore, continuing to understand *SARS-CoV-2* biology and evolution is important for evaluating how this affects and is likely to affect *SARS-CoV-2* infections and disease phenotype within the population going forward.

Therapeutic potential of understanding SARS-CoV-2

Diagnostic testing

While SARS-CoV-2 has implications for human health, understanding the structural and genomic features of the virus has also enabled the development of diagnostic tests to better understand and control viral transmission. Reverse transcription-polymerase chain reaction (RT-PCR) was the first of these to be rolled out routinely in the UK. Using the RT-PCR method, SARS-CoV-2 RNA fragments are reverse-transcribed to DNA which is then amplified in real-time by quantitative PCR (qPCR)^{370–372}. A positive result is subsequently determined where a set quantity of amplified DNA derived from SARS-CoV-2 RNA fragments is obtained following PCR amplification in order to determine if an individual has been infected with SARS-CoV-2^{370–372}. This method is regarded as a gold-standard test as it is highly specific and sensitive for the virus, with specificity and sensitivity recorded as being greater than 99% for SARS-CoV-2 RT-PCR assays^{370–373}. However, the RT-PCR method requires specialist laboratory equipment and trained personnel to run each test thereby limiting testing capacity due to available resources and cost.

Subsequently, lateral flow device (LFD) immunoassays were designed and deployed on a large scale within healthcare and community settings in order to address issues with capacity and cost^{303,373,374}. These have been specifically designed to detect SARS-CoV-2 fragments (antigens) present in samples taken from nasopharyngeal or throat swabs, whereby individuals are able to collect samples, run the tests and obtain a result within 30 minutes without requiring specialist equipment^{303,373,374}. Antigen-based LFD immunoassays comprise a sample pad where the sample is applied, a conjugation pad coated with conjugated antibodies, a membrane (e.g., nitrocellulose) and strips indicating test result (test strip) and validity (control strip)³⁷⁵. Once applied, the liquid sample transfer to the conjugation pad where it binds to antibodies with a conjugate tag such as gold or latex and migrates across the nitrocellulose membrane by capillary flow which comprises 'test' and 'control' strips³⁷⁵. The 'test' and 'control' strips are coated with antibodies against the antigen of interest (SARS-CoV-2 viral proteins) and abundant control analyte, respectively, with a positive test result confirmed by the appearance of coloured test and control lines³⁷⁵. Systematic reviews and meta-analyses of LFD SARS-CoV-2 antigen tests performance have subsequently indicated high test specificity (>95%) and reasonable sensitivity (>70%) of LFDs for identifying SARS-CoV-2 infections in symptomatic and asymptomatic individuals^{374,375}. Moreover, many of the challenges associated with RT-PCR testing methods have been overcome with LFD deployment in community and healthcare settings.

Applying a similar immunoassay methodology, LFDs have also been manufactured for detection of immunoglobulin G and M (IgG/IgM) antibodies against SARS-CoV-2 proteins produced as part of the host immune response to viral infection^{376–379}. The conjugate- and test strip-coating antibodies bind to anti-IgG/IgM antibodies

produced against *SARS-CoV-2* nucleocapsid proteins (COVN) or spike glycoproteins (COVS) which may be present in human serum (blood) samples ^{376–379}. A positive test result can indicate previous *SARS-CoV-2* infection (COVN and COVS) and/or a host response following *SARS-CoV-2* vaccination (COVS only) in an individual ^{376–379}. Furthermore, laboratory-based electro-chemiluminescence immunoassays (ECLIA) such as the Roche Elecsys® *anti-SARS-CoV-2* antibody test have been developed, allowing for antibody quantification within specified concentration ranges ³⁸⁰. This has in turn enabled quantitative assessment of host B cell-associated responses following *SARS-CoV-2* and/or vaccination in order to assess the efficacy of vaccines as well as identify immunodeficient population subgroups ³⁸¹.

Vaccines

Further to developing diagnostic tools, understanding the biology of *SARS-CoV-2* has facilitated the design and roll-out of vaccines against *SARS-CoV-2* with the aim of preventing viral transmission and related clinical events such as hospitalisation through acquired immunity. These have predominantly been developed with the *SARS-CoV-2* spike protein (and variants of) used as the target as this is known to elicit good B-cell and T-cell immune responses upon infection ^{382–386}. To this end, several vaccines against have been designed, manufactured and deployed worldwide with 151,248,820 doses administered from the start of mass immunisation in the United Kingdom alone (between 8th December 2020 and 4th September 2023) ²⁹⁷.

A large proportion of the administered vaccines in the United Kingdom (UK) have either been messenger RNA vaccines, such as BNT162b2 (Pfizer-BioNTech) and mRNA-1273 (Moderna), or adenoviral vector vaccines (e.g., ChAdOx1, Oxford-AstraZeneca) ^{387–389}. The first of these, messenger RNA (mRNA), involves the delivery of mRNA constructs encoding the *SARS-CoV-2* spike protein in lipid nanoparticles into human tissue which is taken up by human cells and translated into spike protein ^{328,390}. Expressed spike proteins are then either secreted from the cell or presented by major histocompatibility complex (MHC) molecules on the cell surface which in turn elicits innate and adaptive immune responses including B and T cell-driven immunity ^{328,390}. On the other hand, adenoviral vector (AAV) *SARS-CoV-2* vaccines comprise a double-stranded DNA (dsDNA) non-enveloped adenovirus containing DNA encoding the spike protein ^{391,392}. The AAV infects host cells via Coxsackie and Adenovirus Receptors (CAR) on the surface of human cells, enabling cell entry ³⁹¹. Messenger RNA encoding *SARS-CoV-2* spike proteins are subsequently transcribed in the host cell nucleus which is then translated into spike protein by harnessing host cell ribosomes and replication machinery ³⁹¹. However, the adenovirus (AAV) itself is non-replicating with only spike protein and non-infecting adenoviral proteins synthesised during the process ³⁹¹. Spike proteins are then either secreted (or shedded) from the host cell or presented on cell-surface MHC molecules to induce host immune responses ³⁹¹. For both vaccine technologies, the underlying aim is to provide acquired

immunity against future *SARS-CoV-2* infections whereby systemic immune responses are triggered in response to viral infection ^{393,394}.

Both mRNA and AAV vaccines have both successfully provided acquired immunity against *SARS-CoV-2* infections as well as reducing incidence of viral transmission and severe COVID-19 disease since the start of the population-wide vaccine programmes. Going forward, it is likely that *SARS-CoV-2* vaccines will continue to be updated to account for evolution of variants circulating in the community. Furthermore, other *SARS-CoV-2* viral proteins are also reported to be potential targets for vaccine development. This includes targeting of nucleocapsid glycoproteins which, like the spike protein, can induce strong T-cell responses and potentially has the advantage of being more genetically conserved whereby recombination events are less frequent than in genes encoding the spike protein ^{383,395,396}. As a result, vaccines targeting conserved genomic regions of *SARS-CoV-2* are more likely to remain effective at providing protection against COVID-19 infections and severe disease as the *SARS-CoV-2* virus evolves.

Antiviral agents and anti-*SARS-CoV-2* monoclonal antibodies

In addition to *SARS-CoV-2* vaccines, better understanding *SARS-CoV-2* virus and course of disease has enabled the implementation of therapeutic agents to reduce disease severity and improve clinical outcomes. This includes antiviral agents such as nirmatrelvir/ritonavir (Paxlovid), remdesivir (Veklury) and molnupiravir (Lagevrio) which are now standard-of-care NHS treatments for clinically vulnerable individuals infected with *SARS-CoV-2* for managing COVID-19 symptoms and to help reduce the risk of severe disease ³⁹⁷. These antiviral drugs target different parts of the *SARS-CoV-2* viral replication process, reducing RNA synthesis and virus production, in order to prevent or reduce systemic infection ^{398–401}. The first of these, nirmatrelvir/ritonavir, is a combination of protease inhibitors which target *SARS-CoV-2* and *HIV-1* virus proteases, respectively, whereas molnupiravir inhibits the viral replication process by introducing copying errors during RNA synthesis ^{399–401}. In brief, molnupiravir is metabolised into β -D-N4-hydroxycytidine 5'-triphosphate (NHC-TP) which is subsequently incorporated instead of cytidine bases during viral RNA synthesis by RNA-dependent RNA polymerase (RdRp) ^{402,403}. Both of these drugs are administered orally with nirmatrelvir/ritonavir typically offered as first-line treatment before molnupiravir ³⁹⁷. Remdesivir on the other hand is an intravenous antiviral agent which undergoes biotransformation into GS-441524 triphosphate that in turn interferes with RdRp activity during viral replication and causes arrest of RNA synthesis ^{398,404}. This is a broad-spectrum anti-viral agent which was first developed to treat hepatitis C, while also subsequently used for treating individuals with Ebola ^{405–407}. While these antiviral agents have been successfully deployed to reduce COVID-19 severity, individuals receiving antiviral therapies can experience toxicity and adverse side-effects including hypersensitivity reactions and gastrointestinal distress as well as potential interactions with

other drugs associated with mechanism(s) of action ^{399,408–411}. Furthermore, these agents are only effective following viral infection and therefore do not provide protection against initial SARS-CoV-2 infection.

As for developing and using antiviral agents, understanding the biology of SARS-CoV-2 has also enabled the development of novel prophylactic treatments to reduce the likelihood of infection in the first place. This includes the development of multiple monoclonal antibodies (mAbs) which specifically bind to different parts of the SARS-CoV-2 spike protein in order to prevent virus binding and entry into human cells via the ACE2 receptor ^{412,413}. Some of these are also designed with modified Fc receptors (Fc fragment of immunoglobulin G) which increases antibody half-life and therefore retain activity for a longer duration following administration ^{412,413}. Perhaps the well-known anti-SARS-CoV-2 mAbs are tixagevimab and cilgavimab, which given in combination under the brand name Evusheld (AZD7442) have been shown to reduce the risk of COVID-19 infection by over 75% at three and six months post-treatment in the phase III PROVENT trial ⁴¹⁴. Importantly, this was demonstrated in a cohort of severely immunocompromised individuals who are unlikely to generate a sufficient immune response following SARS-CoV-2 vaccination ⁴¹⁴.

Subsequent studies and meta-analyses of tixagevimab/cilgavimab and other antibodies such as sotrovimab given prophylactically have demonstrated substantial clinical efficacy for preventing SARS-CoV-2 infection, hospitalisation and mortality ^{415–417}. This includes real-world effectiveness against early Omicron variants such as BA1, BA2, BA4 and BA5 ^{415–417}. While initial laboratory studies predict that these antibodies may be less effective against newer variants, including BQ1.1 and XBB1.5 ^{418–421}, the potential exists for updating antibodies to target different parts of the SARS-CoV-2 spike protein as these are modified during viral evolution. To this end, next-generation monoclonal antibodies targeting newer variants are in development ^{422,423}. In combination, the use of therapeutic agents and approaches developed during the COVID-19 pandemic have helped to reduce disease severity and improve COVID-19-related outcomes as well as reduce the likelihood of SARS-CoV-2 infection in populations of immunocompromised individuals. This in turn has implications for SARS-CoV-2 infections and transmission as well as COVID-19 disease presentation in the population.

Population-based and cohort-based studies of COVID-19 disease

COVID-19 disease at the population level

Understanding the biology of the *SARS-CoV-2* virus has enabled the development and deployment of effective diagnostic testing methods and therapeutic agents for preventing or at least reducing *SARS-CoV-2* infection, transmission, and risk of more severe disease. The impact of these have been evaluated through population-scale studies monitoring *SARS-CoV-2* infections, disease transmission, symptom profiles and clinical outcomes throughout the COVID-19 pandemic. Indeed, multiple studies identified in early 2020 that *SARS-CoV-2* can be transmitted rapidly, and that COVID-19 disease is phenotypically diverse with many individuals experiencing mild symptoms consistent with other common respiratory viral infections or not presenting with symptoms at all ^{424–427}. However, numerous individuals presented with very severe disease phenotypes including the development of acute respiratory distress syndrome (ARDS), cytokine release syndrome and organ failure ^{279–281}. Individual clinical features and comorbidities have been linked to this more severe COVID-19 disease phenotype and poorer health outcomes ^{424,425,428–434}. This was theorised to be a consequence of intrinsic immunological deficits or dysfunction and includes people with diagnoses of autoimmune or haematological disorders, organ transplant recipients and those with respiratory or cardiovascular comorbidities. Studies undertaken in early 2020 also indicated that population demographics were also shown to be associated with higher risk of *SARS-CoV-2* infection, severe COVID-19 disease resulting in hospitalisation, and death ^{428,430,433,435} with higher rates of COVID-19 related hospital admissions and deaths noted in elderly individuals ^{287,433,436}. Taken together, early studies indicated that COVID-19 disease phenotypes and clinical outcomes were heterogeneous within specific patient cohorts and at the population level.

COVID-19 disease in patients with cancer

As for immunocompromised individuals such as those with autoimmune diseases and/or having received organ transplants, the risk of *SARS-CoV-2* infection and adverse clinical outcomes appears to be higher for people living with cancer. Initial analyses at the start of the COVID-19 pandemic in early 2020 identified that patients with cancer were more susceptible to *SARS-CoV-2* infections and at higher risk of severe disease, intensive care admission and death ^{289,290,292}. This is likely to be a consequence of impaired immune function resulting from cancer and/or anticancer treatments, with these individuals more likely to present with an extreme disease phenotype ²⁹². As a result, guidelines for managing patients with cancer in clinical care settings were rapidly altered with changes to systemic anticancer treatment (SACT) plans implemented ^{293,294}.

Further to heterogeneity in individualised risk during the initial stages of the pandemic, there is also likely to be differences within patients with cancer when evaluating acquired immunity following SARS-CoV-2 vaccination. Small cohort studies earlier in the COVID-19 pandemic demonstrated that patients with cancer often had lower or non-detectable humoral and cellular immune responses following vaccination compared to healthy non-immunocompromised individuals ^{437–441}. However, individuals with a cancer diagnosis and/or receiving cancer treatment were often excluded from larger-scale vaccine trials resulting in limited evidence of efficacy before population-level SARS-CoV-2 immunisation programmes were initiated ^{442,443}.

Lastly, while it is widely acknowledged that the COVID-19 disease phenotype in the general population has evolved over the course of the pandemic, less is understood about how this has changed for patients with cancer. These changes at the population level have been driven by the use of therapeutic interventions such as SARS-CoV-2 vaccines and antivirals to reduce disease severity as well as external factors such as the virulence of SARS-CoV-2 variants ^{444–446}. This has been demonstrated within a number of small-scale studies indicates changes in clinical severity and outcomes linked to vaccine uptake and the emergence of the Omicron SARS-CoV-2 variant compared to clinical outcomes of individuals earlier in the COVID-19 pandemic ^{447–450}. It is therefore likely that the interactions of these factors with cancer-associated features and impact on COVID-19 disease phenotype have changed over time for patients with cancer which requires further evaluation.

Evolution of COVID-19 disease phenotype

Following the first SARS-CoV-2 infection wave, widespread evidence has indicated that COVID-19 disease phenotype has evolved at the population level ³⁶⁵. These phenotypic changes have coincided with the development and deployment of therapeutic interventions on a population-scale, most notably the rollout of SARS-CoV-2 vaccinations globally. Clinical trials and cohort studies have demonstrated the power of vaccinations for providing acquired immunity for providing protection against SARS-CoV-2 infections and severe clinical outcomes, resulting in reduced disease transmission and a typically milder disease phenotype for infected individuals ^{451–453}. Further to therapeutic interventions, the SARS-CoV-2 virus has also evolved throughout the course of the COVID-19 pandemic with notable genetic changes within predominant circulating variants appearing to increase viral transmission while reducing virulence. This has corresponded with lower rates of COVID-19 related deaths relative to the number of recorded SARS-CoV-2 infections across different infection waves driven by different circulating SARS-CoV-2 variants ^{362,365,454}. However, severe COVID-19 disease phenotype and associated clinical outcomes are still reported to be higher within immunocompromised individuals and specific population demographic subgroups ^{349,445,450}, highlighting the importance of continued SARS-CoV-2 infection and disease monitoring within clinically vulnerable populations.

Role of COVID-19 cancer patient registries

To address the emerging challenge of COVID-19, registries were set-up in the UK and internationally to evaluate its impact on patients with cancer. This included the collection, processing, analysis and dissemination of data in real-time to inform clinician decision-making and government policy to safeguard patients with cancer. These registries sought to assess potential factors affecting COVID-19 disease severity and outcomes beyond those influencing observed in the general population including primary cancer disease site (tumour type) and active systemic anti-cancer treatments (SACT - chemotherapy, immunotherapy, targeted treatment and hormone therapy) ⁴⁵⁵. One of the largest of these cancer-specific COVID-19 registries is the COVID-19 and Cancer Consortium (CCC-19), comprising 126 cancer centres predominantly in the United States, that was set-up to collate reports of adult patients with an active or historical invasive solid tumour or haematological malignancy diagnosed with COVID-19 ^{456,457}. To date, 22 studies evaluating the impact of COVID-19 in patients with cancer using registry data have been published (up to April 2024) ⁴⁵⁷. However, several smaller-scale registries also emerged during the first wave of the pandemic including the Dutch oncology COVID-19 consortium (DOCC, Netherlands), SAKK 80/20 CaSA (Switzerland) and GCO-002 CACOV-19 (France) which saliently informed oncologists of the impact of COVID-19 within national cancer centres and healthcare settings ^{458–460}. Furthermore, in a UK context, registries including OnCovid were set-up to collect data across multiple cancer centres aiming to describe the features of COVID-19 infection, investigate disease severity and evaluate long-term outcomes in patients with cancer at the beginning of and over the course of the pandemic ⁴⁶¹. In combination, COVID-19 cancer patient registries rapidly developed during the early stages of the COVID-19 pandemic successfully provided insights into the risk factors and prognoses of cancer patients infected with SARS-CoV-2 as well as highlighting areas requiring further evaluation to safeguard patients with cancer.

Concluding remarks and remaining challenges in patients with cancer

In summary, the emergence of SARS-CoV-2 has had profound implications for the wider population in the United Kingdom and worldwide. This has included patients with cancer, where clinical guidelines for anticancer treatment pathways were rapidly developed and altered to reduce the risk of SARS-CoV-2 exposure and severe disease. However, these guideline changes were made principally based on protecting patients from COVID-19 without underlying evidence of the benefits weighed up against how this could have impacted cancer progression. Understanding SARS-CoV-2 viral biology and the physiological impact of resulting COVID-19 disease has therefore been important to develop diagnostic and therapeutic strategies for reducing initial SARS-CoV-2 as well as the risk of severe disease at the population level. This includes the development and deployment of community

symptomatic and asymptomatic testing, vaccine development and initiation of mass vaccination programmes, and introduction of novel or repurposed therapeutic agents into clinical practice to reduce *SARS-CoV-2* transmission, susceptibility to infection and severe disease. In combination, this has shown marked improvements in clinical outcomes in the wider population with fewer individuals presenting with a severe or critical COVID-19 phenotype.

However, when evaluating patients with cancer, the impact of *SARS-CoV-2* has been less well characterised when compared to the wider population. Firstly, many of these initial changes made in clinical settings in the early stages of the COVID-19 pandemic were based on very small-scale retrospective studies ^{289,290}, with many of these changes often not accounting for cancer patient heterogeneity characterised by patient demographics and cancer features (subtype, stage and treatment(s)) which might affect individual risk. Furthermore, when assessing the potential of *SARS-CoV-2* vaccines, individuals with cancer were not included in vaccine trials resulting in limited evidence of efficacy before the rollout of population-level *SARS-CoV-2* immunisation programmes ^{442,443}. It is therefore pertinent to better characterise differences in patient heterogeneity and association with clinical outcomes, assess the impact of *SARS-CoV-2* vaccines within subgroups of patients with cancer for reducing COVID-19 infections and severe disease and how this is associated with quantifiable acquired immunity. Moreover, potential changes in COVID-19 disease phenotype across the pandemic require evaluation on a population-scale to understand how individual risk has evolved for people living with cancer.

In order to address a number of these challenges, the UK Coronavirus Cancer Programme (UKCCP) (<https://ukcovidcancerprogramme.org/>) was set up to better understand the impact of COVID-19 within people living cancer in the UK through population-scale analyses and evaluations of the impact of COVID-19 in patients with cancer ^{462,463}. This was initially launched on 18th March 2020 as the UK Coronavirus Cancer Monitoring Project (UKCCMP), a public health surveillance registry incorporating clinician-reported data on a 24/7 basis from patients with cancer contracting COVID-19 within UK cancer centres ⁴⁶³. COVID-19 records added to the registry were subsequently processed and insights detailing the ongoing impact of COVID-19 in cancer centres disseminated in real-time to UK oncologists as previously described ⁴⁶³. This process was facilitated by rapid approval from the UK Health Research Authority and Professor David Adams (then Pro-Vice Chancellor, University of Birmingham), enabling timely allocation of personnel and computational resources to the project in the face of rapidly-rising COVID-19 hospital admissions ⁴⁶³. Data collection, processing and subsequent dissemination to oncologists in the form of regular reports were coordinated by the Centre for Computational Biology at the University of Birmingham ^{239,463}. However, to deliver the project in full, academics and clinicians based at the University of Birmingham and other UK academic institutions including myself were redeployed in light of the substantial emerging morbidity and mortality of patients with cancer after contracting COVID-19 ^{462,463}. Moreover, and most importantly, the programme sought to empower clinicians to make informed decisions for individual patient care as well as improving cancer care during the COVID-19 pandemic.

CHAPTER 5: MONITORING THE IMPACT OF CORONAVIRUS DISEASE 2019 (COVID-19) IN PATIENTS WITH CANCER AT THE START OF THE PANDEMIC

INTRODUCTION AND AIMS

Understanding the impact of Coronavirus Disease 2019 (COVID-19) in people living with cancer is important for developing strategies to evaluate and mitigate the risks of severe acute respiratory syndrome coronavirus 2 (*SARS-CoV-2*) infections, inform anticancer treatment decisions and prevent severe clinical events related to COVID-19 such as inpatient hospitalisation and mortality. At the start of the pandemic in early 2020, little was known about the risks of *SARS-CoV-2* infection in patients with cancer and how this subsequently corresponded to clinical outcomes, especially across different cancer patient subgroups. Initially, there were no therapeutic options existed for treating the symptoms of COVID-19 in the clinical setting, thus potentially further increasing the risk posed to individuals living with cancer. My primary aim of this chapter was therefore to evaluate COVID-19 in cancer patients at the pandemic outset through analysis and interpretation of clinician-reported data in order identify the most clinically vulnerable patient subgroups following a *SARS-CoV-2* infection during the early stages of the COVID-19 pandemic and demonstrate the usefulness of clinician reported data for COVID-19 disease monitoring.

METHODS

UK Coronavirus Cancer Monitoring Project (UKCCMP) overview (2020)

The United Kingdom Coronavirus Cancer Monitoring Project (UKCCMP) was launched on 18th March 2020 in response to assess the emerging impact of COVID-19 on patients with cancer ⁴⁶³. Involving the collection of data across 55 UK cancer centres, the UKCCMP aimed to analyse and disseminate data in real-time about SARS-CoV-2 infections and COVID-19-related clinical outcomes in patients with cancer to better understand disease impact and inform rapid oncologist decision-making ⁴⁶³. The project database was set-up as a Public Health Surveillance registry with approval from the UK Health Research Authority and in accordance with the UK Policy Framework for Health and Social Care Research, the UK National Research Ethics Service and the UK Governance Arrangement for Research Ethic Committees", as previously noted ^{284,463}. Data reporting was subsequently performed by a Local Emergency Response Reporting Group (LERRG) at each participating UK cancer centre and real-time data processing and dissemination performed through a dedicated pipeline coordinated at the University of Birmingham as previously reported ^{284,463}. The UKCCMP was designed as a Public Health Surveillance registry, with infectious disease monitoring and research studies performed in accordance with the UK Policy Framework for Health and Social Care Research, the UK National Research Ethics Service and the UK Governance Arrangement for Research Ethic Committees ^{284,463}.

Data collection and processing

To support rapid clinical decision-making, a live clinical data reporting and dissemination system was set-up for reporting COVID-19 cases in patients with cancer which incorporated cancer centres throughout the United Kingdom. This prospective data collection was led by a COVID-19 Emergency Response Reporting Individual (ERRI) and subsequently supported by a Local Emergency Response Reporting Group (LERRG) at each centre to promote COVID-19 case reporting on a real-time basis following a positive SARS-CoV-2 test as previously reported ^{284,463}. Corresponding information relating to patient demographics, treatment, COVID-19 disease progression and cancer features were subsequently collected and updated in real-time within each cancer centre, with all information entered into a Research Electronic Data Capture (REDCap) software system set-up by the Institute of Translational Medicine at the University of Birmingham ^{464,465}. The information collected included tumour subtype (by ICD-10 code ⁴⁶⁶), patient age, ongoing cancer treatment, and clinical outcomes following a positive SARS-CoV-2 test. Data anonymisation, processing, analysis, and real-time dissemination to UK cancer centres in the form of weekly reports were then coordinated by the Centre for Computational Biology at the University of

Birmingham and hosted by the Compute and Storage for Life Science (CaStLeS) infrastructure as part of the Birmingham Environment for Academic Research (BEAR) ^{239,463}.

Participant inclusion criteria

All study participants enrolled were adult (age 18+) patients with cancer who presented at a UK cancer centre within the UKCCMP reporting network and in receipt of a positive *SARS-CoV-2* Real-Time Reverse Transcription Polymerase Chain Reaction (RT-PCR) test confirming COVID-19 disease ⁴⁶³. Study recruitment commenced on 18th March 2020 with endpoints for data analyses implemented on 8th May 2020. Individuals with cancer were defined as having metastatic disease, in receipt of any anticancer treatment in any setting (e.g., curative, radical, adjuvant, neoadjuvant) or had been treated with cytotoxic chemotherapy, radiotherapy and/or surgery within 12 months of the study commencing. In keeping with clinical guidelines at the time, RT-PCR testing was only performed where individuals presented as symptomatic of COVID-19 disease ⁴⁶⁷. Individuals with a diagnosis of skin cancer (ICD-10 codes C43-C44) were excluded from the UKCCMP dataset as the registry was updated primarily by oncologists whereas skin cancers are regularly diagnosed by dermatologists rather than oncologists.

Office for National Statistics (ONS) comparison dataset

In addition to the UKCCMP dataset, I used a comparator UK cancer registration dataset obtained from the UK Office for National Statistics (ONS). The “Cancer Survival In England” dataset, which comprised information about adults patients up to 2017, is publicly available from the ONS website and includes patient demographics, including patient age and sex, and primary tumour subtype by ICD-10 codes ⁴⁶⁸. This enabled comparison of patients with cancer and COVID-19 with those without a COVID-19 diagnosis. Patient demographics and primary cancer subtypes corresponded with those recorded within the UKCCMP dataset, with age provided in 10-year age categories (e.g., 40-49, 50-59, 60-69, 70-79).

Statistical analyses and data visualisation

I performed statistical analyses in R version 3.6.3 utilising the `glm()` (family = binomial(link = "logit")) and `fisher.test()` functions, respectively. Data processing and visualisation utilised R (version 3.6.3) packages including `broom`, `dplyr`, `forestplot`, `ggplot2`, `ggsci`, `pheatmap`, `RColorBrewer`, `robustbase` and `viridis`. Data subsetting was performed using the `subset()` function of ‘robustbase’ and data reshaping for visualisation involved the use of the `tidy()` function of ‘broom,’ and `group_by()` and `melt()` functions of ‘dplyr.’ I used functions from the `ggplot2` R package

to generate multiple plots including barplots (`geom_bar`) and lineplots (`geom_line`). The `pheatmap()` and `forestplot()` functions of the 'pheatmap' and 'forestplot' R packages was also used to generate the heatmap and forest plots, respectively.

RESULTS

Outcomes in patients with cancer by demographics and primary tumour subtype in 2020

UKCCMP cohort characteristics

The UKCCMP cohort comprised 1044 patients with cancer and a positive SARS-CoV-2 PCR test (infection) registered within the UKCCMP database between 18th March 2020 and 8th May 2020, with patient outcomes censored at this point. Within the cohort, 57.0% of individuals were male (595/1044) and median age was 70.0 years (IQR 60.0-77.0). Each patient was followed up from the point of a confirmed SARS-CoV-2 infection up to hospital discharge or mortality.

UKCCMP comparison with ONS cancer census

I first compared the demographics and primary cancer subtype of the UKCCMP SARS-CoV-2 positive cancer cohort with the overall population of individuals living with cancer represented in the ONS cancer census which I used as a historical control group. This comparison identified that individuals with a cancer diagnosis and SARS-CoV-2 positive test were significantly more likely to be male than within the ONS cancer census (57.0% v 51.3%, OR: 1.26 (95% CI: 1.12-1.43), $p=0.0002$), though the age distributions within both cohorts were not significantly different (median age group of 70-79 years for both) (Table 5, Figure 28, Supplementary Figure 1).

When assessing primary tumour subtypes, I found that patients with haematological malignancies were overrepresented within the SARS-CoV-2 positive cancer cohort including those with a diagnosis of leukaemia (OR: 2.82, 95% CI: 2.21-3.55, $p<0.001$), myeloma (OR: 2.03, 95% CI: 1.42-2.83, $p<0.001$) or lymphoma (OR: 1.63, 95% CI: 1.28-2.06, $p<0.001$) (Table 5). Conversely, patients with a diagnosis of lung or prostate cancer were underrepresented in the UKCCMP dataset compared to the ONS cancer census. Lung cancer diagnoses comprised 10.7% of patients within the UKCCMP dataset compared to 13.7% of ONS cases (OR: 0.75, 95% CI: 0.61-0.91, $p=0.003$). Similarly, prostate cancer diagnoses made up 11.0% of individuals in the UKCCMP cohort compared to 14.6% of the ONS census (OR: 0.72, 95% CI: 0.59-0.88, $p<0.001$).

Patient characteristics	UKCCMP cases (%)	ONS cases (%)	Odds Ratio (95% CI)	p value
Age				
Median age (years)	70.0	NA*		
Sex**				
Male	595 (57.0%)	145034 (51.3%)	1.26 (1.12-1.43)	0.0002
Female	445 (42.6%)	137844 (48.7%)		
Other	4 (0.4%)	0 (0.0%)		
Solid tumour malignancy (ICD-10 subtype)				
Lip, oral cavity, and pharynx (C00-C14)	33 (3.2%)	7558 (2.7%)	1.19 (0.82-1.69)	0.334
Non-colorectal gastrointestinal (C15-C17, C22-C26)	95 (9.1%)	30096 (10.6%)	0.84 (0.68-1.04)	0.118
Colorectal gastrointestinal (C18-C21)	124 (11.9%)	36039 (12.7%)	0.93 (0.76-1.12)	0.456
Lung (C34)	111 (10.7%)	38878 (13.7%)	0.75 (0.61-0.91)	0.003
Respiratory and intrathoracic (C30-C33, C35-C39)	11 (1.1%)	2780 (1.0%)	1.08 (0.53-1.94)	0.752
Bone and articular cartilage (C40-C41)	4 (0.4%)	376 (0.1%)	2.90 (0.78-7.50)	0.053
Mesothelial and soft tissue (C45-C49)	16 (1.5%)	4682 (1.7%)	0.93 (0.53-1.52)	0.903
Breast (C50)	143 (13.7%)	46109 (16.3%)	0.82 (0.68-0.98)	0.026
Female gynaecological (C51-C58)	56 (5.4%)	17969 (6.4%)	0.84 (0.63-1.10)	0.226
Prostate (C61)	114 (11.0%)	41200 (14.6%)	0.72 (0.59-0.88)	<0.001
Male urological (C60, C62-C63)	4 (0.4%)	2435 (0.9%)	0.44 (0.12-1.14)	0.126
Urinary tract (C64-C68)	77 (7.4%)	19333 (6.8%)	1.09 (0.85-1.38)	0.460
Central nervous system (C69-C72)	25 (2.4%)	5038 (1.8%)	1.36 (0.87-2.02)	0.127
Endocrine glands (C73-C75)	4 (0.4%)	3374 (1.2%)	0.32 (0.09-0.82)	0.010
Haematological malignancy (ICD-10 subtype)				
Lymphoma (C81-C85)	79 (7.6%)	13537 (4.8%)	1.63 (1.28-2.06)	<0.001
Myeloma (C90)	37 (3.6%)	5033 (1.8%)	2.03 (1.42-2.83)	<0.001
Leukaemia (C91-C95)	79 (7.6%)	8018 (2.8%)	2.82 (2.21-3.55)	<0.001
Other Haematological (C86, C88, C96)	29 (2.8%)	423 (0.1%)	19.14 (12.59-28.05)	<0.001

Table 5: Demographics and primary tumour subtype representation of patients with cancer in the UKCCMP COVID-19 cohort compared to the ONS cancer census. Univariate comparisons, not adjusting for age and sex, were performed using Fisher's Exact testing for determining Odds Ratios and p values. * Individual ages were not available within the ONS dataset. ** Statistical comparison for sex only included individuals listed as 'male' or 'female.'

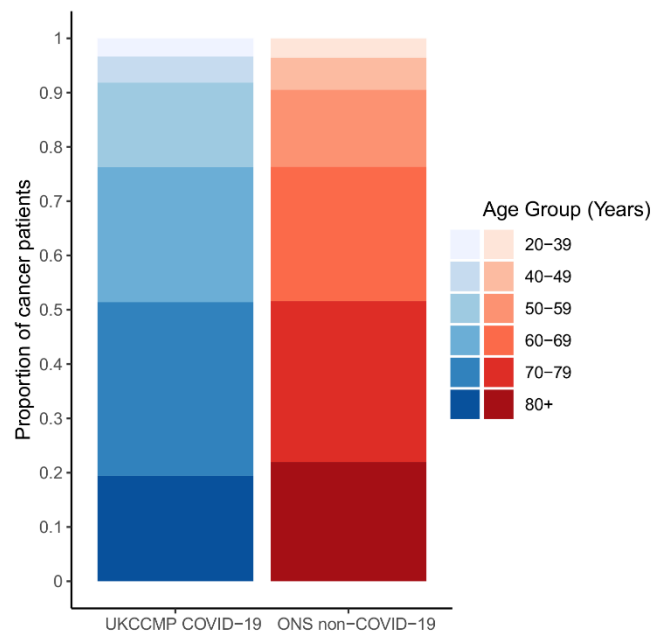


Figure 28: Stacked bar chart showing the age distributions of patients with a recorded cancer diagnosis within the UKCCMP SARS-CoV-2 positive and ONS cancer control population. The median age group for both cohorts was 70-79 years whereby individual ages were not available within the ONS dataset.

Case-fatality rates by patient demographics

Having identified differences in patient demographics between the UKCCMP and ONS datasets, I evaluated how this corresponds with outcomes of patients with cancer by comparing case-fatality rates (CFRs) of individuals with or without a confirmed SARS-CoV-2 infection. In the UKCCMP dataset, 319 out of 1044 (30.6%) patients died within the study period, of which the cause of death was reported to result from COVID-19 in 295/319 (92.5%) cases. This is in contrast to reported case-fatality rates of between 1% and 15% within non-cancer populations upon contracting COVID-19 ^{280,281,286,424}. I subsequently found that differences in all-cause CFRs following a SARS-CoV-2 infection were significantly linked to increasing age, whereby the case fatality rates in the 40-49, 50-59, 60-69, 70-79 and over 80 age groups being 0.10, 0.17, 0.28, 0.35 and 0.48 respectively, with no deaths recorded in individuals under 40 (Figure 29, Supplementary Figure 1). When comparing patients with cancer and a confirmed SARS-CoV-2 infection with ONS cancer census data, age was confirmed to be a significant risk factor for death following COVID-19 with fatalities overrepresented in over 70-year olds in the UKCCMP cohort (Figure 30). Additionally, the all-cause CFR in patients with cancer once they had contracted COVID-19 also appeared to be higher in males than females at 0.34 and 0.23, respectively (OR: 1.92, 95% CI: 1.51-2.45, $p < 0.001$).

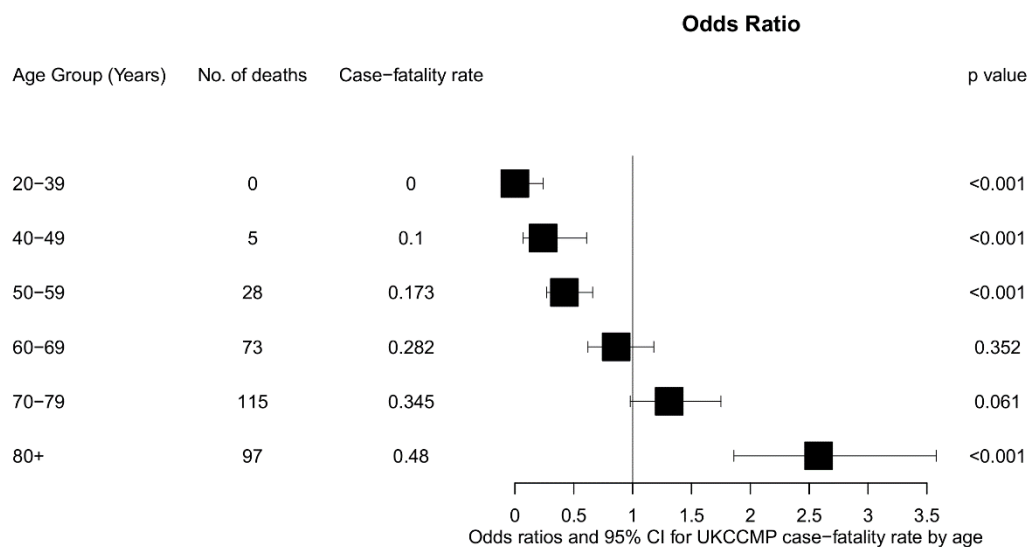


Figure 29: Forest plot showing the distribution of age groups of patients with cancer who died in the UKCCMP and corresponding case-fatality rates. All Odds Ratio are relative to the whole UKCCMP population within the study period for calculating Odds Ratios and 95% confidence intervals.

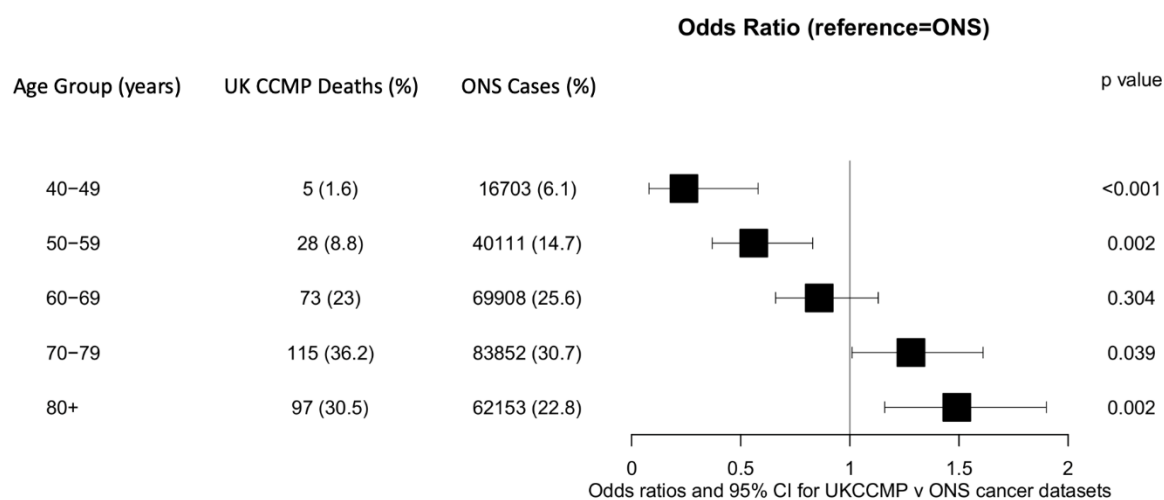


Figure 30: Forest plot showing distribution of age groups of patients with cancer who died in the UKCCMP relative to the age distribution of the ONS cancer census control population. The ONS cancer census was used as the reference group for calculating Odds Ratios and 95% confidence intervals.

Case-fatality rates by primary tumour subtype

Noting differences between the UKCCMP and ONS datasets when comparing the composition of primary tumour subtypes, I assessed case-fatality rates (CFRs) by primary cancer diagnosis within the UKCCMP cohort. The highest CFRs following a SARS-CoV-2 infection were observed in patients with a diagnosis of myeloma (0.432), prostate cancer (0.430) or leukaemia (0.418) whereas the CFRs were lowest in patients with gynaecological (female genital organ) (0.125) or breast (0.182) cancer diagnoses (Table 6, Supplementary Figure 2).

To compare relative differences in CFRs between different cancer subtypes for patients proceeding a SARS-CoV-2 infection, I used individuals with non-colorectal gastrointestinal cancer diagnosis (digestive organ, ICD-10 C15-C17 and C22-C26) as a reference as this was the tumour subtype with the median CFR (0.295). Through univariate analyses, I observed that patients with a diagnosis of prostate cancer (OR: 2.14, 95% CI: 1.17-3.96, $p=0.014$) or leukaemia (OR: 2.03, 95% CI: 1.04-3.97, $p=0.038$) were both at significantly higher risk, while also identifying lower risk in patients with a gynaecological (female genital organ) (OR: 0.36, 95% CI: 0.13-0.87, $p=0.031$) or breast (OR: 0.53, 95% CI: 0.28-1.00, $p=0.049$) cancer diagnosis (Figure 31, Table 6).

I subsequently performed a corresponding multivariable adjusted analysis to account for age and sex as clinically relevant confounders, observing that patients with leukaemia remained at significantly higher risk of death than individuals with other cancers (OR 2.25, 95% CI: 1.13-4.57, $p=0.023$), (Figure 31, Table 6, Supplementary Figure 3). However, following multivariable adjustment, prostate cancer was no longer significantly associated with a higher CFR, and gynaecological and breast cancer diagnoses were no longer associated with a lower CFR, highlighting the impact of patient age and sex on the likelihood of death following or resulting from COVID-19 (Figure 31, Table 6). Similarly, following multivariable adjustment, the CFR for patients with lung cancer following a SARS-CoV-2 infection was not significantly higher compared to other primary cancer diagnoses within the UKCCMP population (OR: 1.41, 95% CI: 0.75-2.67, $p=0.285$) compared to the rest of the UKCCMP population.

Malignancy (ICD-10 subtype)	COVID-19 deaths		Univariable		Multivariable adjusted	
	Total	CFR	OR (95% CI)	p value	OR (95% CI)	p value
Prostate (C61)	49	0.430	2.14 (1.17-3.96)	0.014	1.09 (0.51-2.33)	0.824
Lung (C34)	43	0.387	1.62 (0.89-3.00)	0.118	1.41 (0.75-2.67)	0.285
Mesothelial and soft tissue (C45-C49)	6	0.375	1.18 (0.37-3.51)	0.772	1.52 (0.43-5.30)	0.505
Urinary tract (C64-C68)	23	0.299	1.08 (0.54-2.13)	0.834	0.87 (0.41-1.81)	0.715
Colorectal gastrointestinal (C18-C21)	35	0.282	1.03 (0.56-1.90)	0.934	0.85 (0.44-1.64)	0.627
Central nervous system (C69-C72)	7	0.280	1.15 (0.39-3.18)	0.797	1.87 (0.57-6.05)	0.292
Respiratory organs (C30-C33, C35-C39)	3	0.273	0.84 (0.17-3.29)	0.813	0.96 (0.18-4.10)	0.954
Lip, oral cavity, and pharynx (C00-C14)	8	0.242	0.75 (0.28-1.85)	0.542	0.77 (0.25-2.27)	0.644
Breast (C50)	26	0.182	0.53 (0.28-1.00)	0.049	0.97 (0.40-2.52)	0.942
Female gynaecological (C51-C58)	7	0.125	0.36 (0.13-0.87)	0.031	0.79 (0.24-2.63)	0.704
<i>Myeloma (C90)</i>	16	0.432	1.85 (0.81-4.22)	0.142	1.65 (0.71-3.85)	0.241
<i>Leukaemia (C91-C95)</i>	33	0.418	2.03 (1.04-3.97)	0.038	2.25 (1.13-4.57)	0.023
<i>Lymphoma (C81-C85)</i>	25	0.316	1.60 (0.80-3.19)	0.184	1.72 (0.81-3.68)	0.156
<i>Other Haematological (C86, C88, C96)</i>	7	0.241	0.81 (0.28-2.12)	0.675	0.81 (0.26-2.33)	0.702
Non-colorectal gastrointestinal (C15-C17, C22-C26)	28	0.295	Reference	Reference	Reference	Reference

Table 6: All-cause case fatality rate following COVID-19 by primary tumour subtype and assessment of relative risk before and after multivariable adjustment by patient age and sex. Non-colorectal gastrointestinal (C15-C17 and C22-C26) diagnoses were used as a baseline for calculating univariable and multivariable Odds Ratios for assessing relative risk.

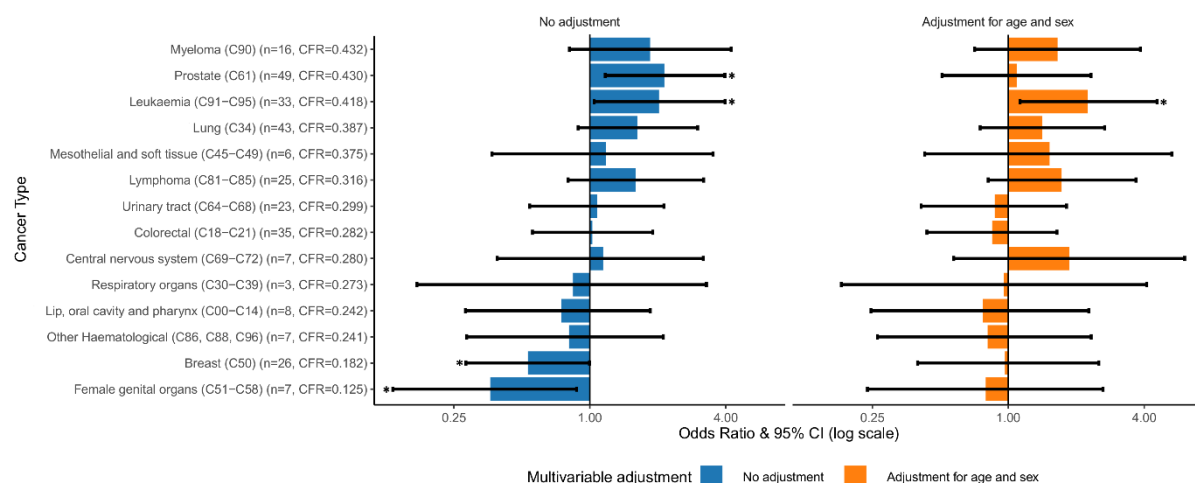


Figure 31: Waterfall plot showing the relative risk of death for each primary tumour before and after multivariable adjustment for patient age and sex. Non-colorectal gastrointestinal (digestive organ, C15-C17 and C22-C26) diagnoses were used as a baseline for calculating univariable and multivariable Odds Ratios. Statistical significance before or after multivariable adjustment ($p < 0.05$) is denoted by *.

Clinical features and outcomes of patients with haematological and solid tumour malignancies

Having identified higher case-fatality rates or risk of death of patients with haematological malignancies, I performed a more detailed comparison of patients diagnosed with haematological malignancies (n=227) versus those with solid tumour malignancies (n=817) following COVID-19 within the UKCCMP cohort. Initially, I found that patients with haematological or solid tumour malignancies presented with similar COVID-19 symptoms, however after adjusting for age and sex, individuals with haematological malignancies were significantly more likely to require high flow oxygen (OR 1.82, 95% CI: 1.11-2.94, $p=0.015$), non-invasive ventilation (OR: 2.10, 95% CI: 1.14-3.76, $p=0.014$), intensive care admission for ventilation (OR: 2.73, 95% CI: 1.43-5.11, $p=0.002$) and have a severe or critical disease presentation (OR: 1.57, 95% CI: 1.15-2.15, $p=0.004$) (Supplementary Table 15). Moreover, I confirmed that all-cause and COVID-19 attributed deaths were significantly higher for individuals with any haematological malignancy compared with patients diagnosed with a solid tumour malignancy before and after multivariable adjustment for age and sex (Table 7).

Patient outcome	Haematological (n=227)	Solid tumour (n=817)	Univariable		Multivariable adjusted	
			Odds Ratio (95% CI)	p value	Odds Ratio (95% CI)	p value
Death (all cause)	82 (36.1%)	237 (29.0%)	1.61 (1.15-2.24)	0.005	1.74 (1.21-2.48)	0.002
Death (COVID-19)	80 (35.2%)	215 (26.3%)	1.77 (1.27-2.48)	0.001	1.93 (1.35-2.77)	<0.001
Death (Cancer)	1 (0.4%)	18 (2.2%)	0.21 (0.01-1.02)	0.129	0.22 (0.01-1.06)	0.138
Death (other)	1 (0.4%)	4 (0.5%)	0.96 (0.05-6.54)	0.972	1.12 (0.06-7.79)	0.923
Hospitalisation	5 (2.2%)	36 (4.4%)	0.52 (0.18-1.23)	0.178	0.53 (0.18-1.26)	0.192

Table 7: Univariable and multivariable-adjusted analyses of clinical outcomes of patients with primary diagnoses of haematological and solid tumour malignancies. Univariable analysis was conducted with presence compared to absence (reference for each category) in haematological malignancies vs. solid-tumour malignancies. Multivariable analyses adjusted for patient age and sex.

Finally, in the context of anticancer treatments, I identified that 47.6% of patients with haematological malignancies had received chemotherapy within 4 weeks of COVID-19 diagnosis compared to 29.5% of those with solid tumour cancers (OR: 2.15, 95% CI: 1.57-2.95, $p<0.0001$) (Supplementary Table 15). Noting previous studies reporting contrasting findings of the association between chemotherapy treatment and death following COVID-19 ^{284,291,292,456}, I subsequently compared outcomes of patients with haematological malignancies based on chemotherapy treatment status. I observed that recent use of chemotherapy was associated with an increased risk of death following a SARS-CoV-2 infection following multivariable adjustment for patient age and sex (OR: 2.09, 95% CI: 1.09-4.08, $p=0.028$). However, this finding did not reflect the whole cancer cohort (Table 8), with the opposite seen for individuals with solid tumour malignancies. This is in keeping with previous studies describing

COVID-19 cancer cohorts of predominantly solid tumour malignancies which reported no significant excess risk of death following recent chemotherapy ^{284,456}.

Malignancy type	Case-fatality rates (deaths/total)			Univariable		Multivariable adjusted	
	Overall	Chemotherapy	No Chemotherapy	Odds Ratio (95% CI)	p value	Odds Ratio (95% CI)	p value
All Cancers	0.306 (319/1044)	0.258 (90/349)	0.329 (229/695)	0.64 (0.47-0.86)	0.003	0.95 (0.68-1.31)	0.747
Haematological	0.361 (82/227)	0.407 (44/108)	0.319 (38/119)	1.54 (0.86-2.80)	0.149	2.09 (1.09-4.08)	0.028
Solid Tumour	0.290 (237/817)	0.191 (46/241)	0.332 (191/576)	0.41 (0.28-0.59)	<0.001	0.63 (0.42-0.93)	0.023

Table 8: Univariable and multivariable-adjusted analyses of all-cause mortality of patients based on receipt of chemotherapy treatment. Univariable analysis was conducted with presence compared to absence (reference for each category) in chemotherapy v no chemotherapy. Multivariable analyses adjusted for patient age and sex. Sensitivity analyses were performed within individuals with haematological or solid tumour malignancy diagnoses.

DISCUSSION

At the start of the COVID-19 pandemic in 2020, the impact of SARS-CoV-2 in patients with cancer was poorly characterised. Any understanding was based on small-scale retrospective cohort studies making it challenging to infer conclusive findings. Up to the end of April 2020, the largest published study comprised only 105 patients with cancer which reported high mortality rates following a SARS-CoV-2 infection in patients with a haematological or lung cancer diagnosis as well as individuals with metastatic disease ²⁸⁵. However, the emergence of COVID-19 and publication of these studies had a profound effect on the delivery of cancer care across multiple health systems, including in the United Kingdom (UK), with patients considered to be one homogenous clinically vulnerable population group. This included alterations to cancer treatment plans such as delays and cancellations to planned treatment and operations in part resulting from hesitancy to expose patients to COVID-19 or prescribe otherwise effective anticancer treatment which may worsen COVID-19 disease course ^{284,296,298}. As it is critical to balance the potential benefits and risks of anticancer treatment, it is also pertinent to accurately evaluate the risk of COVID-19 disease on an individual level considering patient characteristics (e.g., age), cancer diagnosis and cancer prognosis. It was on this basis that the UK Coronavirus Cancer Monitoring Project (UKCCMP), later evolving into the UK Coronavirus Cancer Programme, was initiated in March 2020 to monitor the impact of COVID-19 on patients with cancer including the assessment of individualised risk following a SARS-CoV-2 infection ⁴⁶³.

In keeping with the aims of the UKCCMP, my UK cohort-based analysis of the impact of COVID-19 in patients with cancer successfully identified subgroups at higher risk of severe clinical outcomes during the first pandemic wave in 2020. This incorporated clinician reported UKCCMP data comprising 1044 patients with cancer and publicly available cancer data from the Office for National Statistics (ONS) from individuals who contracted and did not contract COVID-19, respectively, with both sub-grouped by patient demographics and cancer subtype. I identified that individuals with haematological malignancies were overrepresented in terms of susceptibility to SARS-CoV-2 infection and of presenting with a more severe COVID-19 disease phenotype. This reflects the observations in COVID-19 related hospital admissions of individuals with cancer during the early stages of the pandemic ^{285,291}, though I was the first to demonstrate this using real-world data through my comparisons with ONS cancer registration statistics. Furthermore, the age profiles of individuals with cancer who contracted COVID-19 and subsequently died were also considerably different when compared to ONS statistics, with deaths of patients aged 70 years and older overrepresented in the UKCCMP cancer cohort. This in combination with an overall case-fatality rate of 30.6% (0.306) within patients with cancer who contracted COVID-19 which is in contrast to reported case-fatality rates of between 1% and 15% within non-cancer populations upon contracting COVID-19 ^{280,281,286,424}.

Taken together, this highlighted the necessity to monitor and safeguard patients with cancer from contracting COVID-19 during the initial stages of the pandemic.

Absolute and relative case-fatality (death) rates were higher in those with a leukaemia diagnosis in comparison to other primary tumour subtypes following a SARS-CoV-2 infection. Before and after adjusting for patient age and sex, this finding is in keeping with findings from other cancer cohorts in the UK and China published between April and June 2020 ^{430,469,470}. This provided the basis to directly compare COVID-19 disease phenotype and outcomes of patients with haematological versus solid tumour cancers where I observed higher risk of a severe or critical disease presentation and death associated with a haematological malignancy diagnosis. Higher risk of severe outcomes associated with haematological cancers, especially leukaemia, may be attributable to the profound immune dysfunction observed with these types of cancers thereby increasing susceptibility to SARS-CoV-2 infection and likelihood of severe immune-related events such as cytokine storm. The increased risk of death in patients with haematological cancers also appears to be, when adjusted for age and sex, further elevated for those who had received recent chemotherapy treatment in keeping with a study published in China ⁴⁷¹. However, this was not found to be the case overall within all cancer subtypes and indeed the opposite for solid tumour malignancies. Taking all cancer subtypes together, this corresponds with other earlier studies describing COVID-19 cancer cohorts which reported no significant excess risk of death following recent chemotherapy while identifying patient demographics and other non-cancer comorbidities as significant risk factors ^{284,456}. The observed higher case-fatality rate in individuals with haematological cancers who received chemotherapy may therefore reflect cancer severity rather than being driven by chemotherapy treatment. It is also important to acknowledge that the effect of modifying or withholding anticancer treatment could profoundly affect cancer morbidity and mortality, perhaps more so than the risks associated with contracting COVID-19.

Contrary to my findings for haematological malignancies, diagnoses of lung, prostate and breast cancers were underrepresented in individuals who had contracted COVID-19 compared to the ONS cancer statistics. Individuals with a lung cancer diagnosis also didn't have a markedly higher case-fatality rate in the context of the UKCCMP cancer cohort, perhaps as a result of effective shielding measures implemented during the early stages of the pandemic following initial studies reporting increased clinical vulnerability ^{289,291,292,424,472}. Individuals with a prostate cancer diagnosis initially appeared to have an increased risk of death proceeding COVID-19 but, when accounting for age and sex by multivariable adjustment, shows that this risk is no greater than the rest of the UKCCMP population. On the other hand, patients diagnosed with a breast or gynaecological malignancy appeared to be at lower risk of death, though multivariable analyses demonstrated that this protection is likely to be the result of individuals being female with males typically observed to be at higher risk of adverse events following a SARS-CoV-2 infection ^{473,474}. In combination, my analyses suggest that the COVID-19 disease phenotype in patients with

cancer is likely to be determined by several factors including demographics (e.g., age, sex), primary tumour subtype and perhaps anticancer treatment.

There are some potential limitations of my analyses. Firstly, UK national COVID-19 testing policy at the time consisted of RT-PCR testing of individuals presenting with COVID-19 symptoms within secondary care settings (e.g., hospitals). The UKCCMP cohort therefore comprised patients with cancer presenting in cancer centres with COVID-19 symptoms which in turn may not be fully representative of all patients with cancer with SARS-CoV-2 infections during this period. I also analysed UKCCMP data with a specific data cut-off rather than determining cohort size through sample size calculations and a priori power analyses, with the pragmatism required for timely dissemination of results outweighing the benefits of delaying data cut-off.

Furthermore, there are also notable differences to acknowledge between the UKCCMP cohort and ONS cancer statistics, which I used as a historical control, whereby the UKCCMP only records individuals with 'active cancer' and ONS consists of all cancer diagnoses. This reflects the differences in data collection between the two datasets with the UKCCMP cohort capturing individuals presenting to cancer centres at the start of COVID-19 pandemic at a time where clinical guidelines advised against all-but-essential visits to secondary care centres^{298,467}. As this included in-person consultations and pauses or cessation of anticancer treatments for some patients^{298,467}, individuals with SARS-CoV-2 infections not presenting to cancer centres during this time period were not incorporated in the UKCCMP registry. In turn, these differences may be confounders when assessing outcomes of patients with cancer following a SARS-CoV-2 infection including mortality. It is also important to note that some clinical practices for disease management and available anticancer treatments have changed since 2017, and before the COVID-19 pandemic, when data collection for the ONS cancer statistics used in my analyses ended. However, my comparison of UKCCP patients who contracted COVID-19 with ONS statistics, a cancer dataset with a similar age distribution, has successfully contextualised the risk of contracting COVID-19 and death in the period proceeding infection. This is in contrast to other studies which have predominantly assessed COVID-19 related outcomes in individuals with cancer compared to the general population^{285,289,291}.

To summarise, I identified during the initial stages of the COVID-19 pandemic in 2020 that the risk of infection, severe COVID-19 phenotype and adverse clinical outcomes is heterogeneous within individuals living with cancer in the UK. Patient age, sex and primary tumour subtype were all associated with susceptibility to SARS-CoV-2 infection whereby older patients (70 years and above), males and those with a haematological malignancy representing higher proportions of SARS-CoV-2 infections compared to a representative historic ONS national cancer dataset. This in turn corresponded with increased risk of severe outcomes after contracting COVID-19 with a proportion of the initial variation between primary tumour subtype attributable to patient age and sex. Individuals with a haematological malignancy who contract COVID-19 also appear to be at higher risk of death following recent chemotherapy, however, it is crucial to consider potential cancer morbidity and mortality as well as COVID-19 when

determining individualised anticancer treatment plans. Taken together, my analyses and findings demonstrate the utility of monitoring the impact of COVID-19 in patients with cancer during the initial stages of the pandemic by facilitating a more individualised risk stratification of people living with cancer following a *SARS-CoV-2* infection.

**CHAPTER 6: AN EVALUATION OF HOW THE
IMPACT OF CORONAVIRUS DISEASE 2019
(COVID-19) HAS CHANGED IN PATIENTS WITH
CANCER ACROSS THE PANDEMIC**

INTRODUCTION AND AIMS

My evaluation of the impact of COVID-19 in patients with cancer during the initial stages of the COVID-19 pandemic identified subgroups who are at higher risk of severe COVID-19 disease, most notably individuals with a haematological malignancy. It is likely that the risk faced by patients with cancer has changed following the development and deployment of COVID-19 vaccines which have been reported to reduce susceptibility to infection and likelihood of severe disease in the general population. The UK Coronavirus Cancer Programme (UKCCP) was therefore tasked to address these questions, including through the linkage and analysis of population-level electronic healthcare records (EHRs), to provide real-world evidence of this impact and inform clinical decision-making to protect and safeguard people living with cancer ⁴⁶².

Specifically in the context of COVID-19 vaccines, the original clinical trials evaluating vaccine efficacy did not incorporate patients with cancer and/or individuals receiving anticancer therapy ^{442,443}, with indications that vaccine uptake within this population may be variable given this paucity of evidence ⁴⁷⁵. Vaccine effectiveness has been successfully assessed using diagnostic testing data and applying a test-negative case control (TNCC) methodology within non-cancer populations ^{451,476}. Given the potential of applying a TNCC design to SARS-CoV-2 diagnostic testing data, I sought to assess the impact of SARS-CoV-2 vaccines in patients with cancer compared to the non-cancer population and within cancer subgroups. This coincides with the launch of the primary and booster SARS-CoV-2 vaccination programmes in the UK in December 2020 and September 2021, respectively ^{477,478}. This includes the assessment of vaccine effectiveness following 2-dose primary and 3-dose booster vaccine regimens for providing protection against SARS-CoV-2 infection and severe clinical outcomes including hospital admission, intensive care admission and mortality in patients with cancer and the non-cancer population. Taken together, achieving these aims would show whether COVID-19 vaccine protect patients with cancer, if or how long this protection is sustained and if individuals could derive benefit from regular booster vaccine doses.

Following this, I aimed to better understand the biological basis of any differences in vaccine effectiveness in cancer patient subgroups by evaluating the association of immune responses following COVID-19 vaccination with cancer features and clinical outcomes through the interrogation of SARS-CoV-2 antibody test results obtained through multiple studies including the National COVID-19 Cancer Antibody Survey (NCCAS). In each case, I retrospectively analysed national public health datasets on a population-scale within England, comprising data including SARS-CoV-2 testing data, COVID-19 vaccination records, national cancer registrations and hospital records detailing clinical outcomes.

In addition to therapeutic changes such as COVID-19 vaccination, notable other changes have been associated with changes in disease infectiousness, symptom presentation and clinical severity at the population-level. Most notably, the *SARS-CoV-2* virus has continually evolved with novel variants and subvariants emerging as the primary driver of new COVID-19 infections with Omicron becoming the predominant circulating variant in the UK in December 2021 ³⁵⁷. In addition, new therapeutic strategies have been implemented to mitigate the risk of COVID-19 infections and severe clinical outcomes including prophylactic antibody therapies and utilisation of antiviral agents ^{301,399,414,417}. As the implications of these changes for patients with cancer in the context of COVID-19 disease presentation was poorly understood, my final aim was to temporally assess the changes in COVID-19 disease phenotype across the course of the pandemic and to understand if there was any excess residual risk for patients with cancer. This would incorporate analyses of *SARS-CoV-2* testing data and corresponding population-scale public health datasets prior to the start of the COVID-19 vaccination programme and up to the later stages of the UK-wide national COVID-19 testing programme when Omicron had become the predominant circulating *SARS-CoV-2* variant. These population-scale evaluations aimed to better understand and characterise COVID-19 in patients with cancer and identify patient subgroups at higher risk of severe clinical outcomes to inform clinical decision-making to continue to deliver high-quality cancer treatment and care while also mitigating the risks posed by COVID-19.

METHODS

UK Coronavirus Cancer Programme (UKCCP) datasets

UK Coronavirus Cancer Programme (UKCCP) overview

The mission of the UK Coronavirus Cancer Programme (UKCCP) has been to safeguard, evaluate and protect patients with cancer in the United Kingdom during the COVID-19 pandemic (www.ukcovidcancerprogramme.org). As a key component of this, the UKCCP sought to harness the potential of electronic healthcare records (EHRs) in order to provide real-world evidence of the ongoing impact of COVID-19 in patients with cancer and support rapid clinical decision-making ⁴⁶². The UKCCP datasets comprise EHRs from NHS England which was accessed, linked and processed within a secure research environment at the UK Health Security Agency (UKHSA), previously Public Health England ⁴⁶². This process was performed in accordance with the UK Policy Framework for Health and Social Care Research and studies within the programme supported by the United Kingdom Department of Health and Social Care (DHSC) with ethical approval from the Health Research Authority under reference 20/WA/0181.

UKCCP dataset curation and linkage

Data was linked and processed at the SARS-CoV-2 test level, with polymerase chain-reaction (PCR) and lateral flow test (LFT) testing data from community and hospital settings obtained from the NHS second-generation surveillance system (SGSS). SARS-CoV-2 test level data was then linked to clinically relevant datasets including National Immunisation Management Service (NIMS) (later termed National Immunisation and Vaccination System), Secondary Use Statistics (SUS) and Contact Tracing and Advice Service (CTAS) comprising information about COVID-19 vaccination records, clinical outcomes in secondary care settings (e.g., hospitals) and contact tracing as part of the NHS Test and Trace programme, respectively (Figure 32) ^{302,479–481}. Individuals with a cancer diagnosis were subsequently identified using the Public Health England (PHE) rapid registration national cancer dataset ⁴⁸², with non-cancer population control cohorts consisting of SARS-CoV-2 tests from individuals who were not contained within this national cancer dataset (Figure 32). Antibody testing data obtained from individuals who had received at least two COVID-19 vaccine doses was also collated from the pillar 3 antibody dataset and linked to SARS-CoV-2 testing data and secondary care outcomes (Figure 32) ⁴⁸³. All participants, including those in non-cancer population control groups, were 18 years or older.

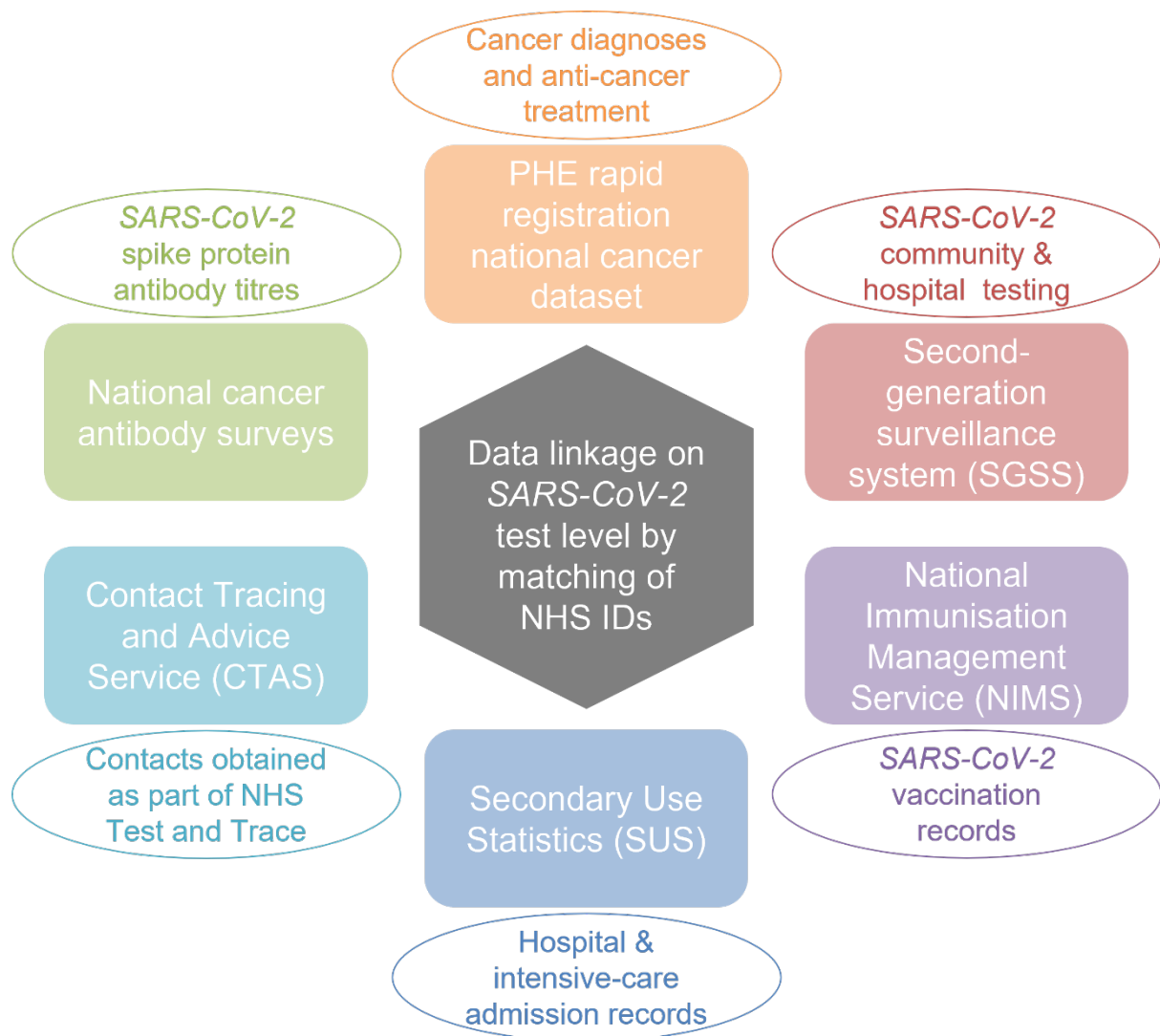


Figure 32: NHS England electronic healthcare datasets (EHRs) comprising the UK Coronavirus Cancer Programme datasets. The rapid registration national cancer dataset contains cancer diagnoses recorded between 1st January 2018 and 30th April 2021. UKCCP data was linked at the SARS-CoV-2 test level through exact matching of 10-digit unique NHS IDs within a secure research environment at the UK Health Security Agency (UKHSA).

Each UKCCP dataset comprised information about patient demographics including age, sex, ethnicity, and deprivation, with geographical location being used to determine the index of multiple deprivation (IMD), a national indicator of deprivation ⁴⁸⁴. Ethnicity was defined in line with UK government guidelines ⁴⁸⁵. I incorporated these demographics into my analyses as a result of evidence of demographic-specific differences in clinical outcomes following a cancer diagnosis and/or SARS-CoV-2 infection ^{279,433}. Within the cancer cohorts, detailed information relating to cancer features and anticancer treatments such as cancer stage and subtype, receipt of radiotherapy and systemic anticancer treatments (SACT) were included. Cancer subtype was defined by ICD-10

subtype ⁴⁶⁶ and SACT encompasses several classes of cancer treatment including cytotoxic agents (chemotherapy), targeted, immune or hormonal therapies ⁴⁸⁶.

COVID-19 infection and clinical outcome definitions

The UKCCP dataset was linked and processed at the SARS-CoV-2 test level rather than individual or patient level. This study design was implemented as individuals can have multiple SARS-CoV-2 or COVID-19 infections over time because of SARS-CoV-2 virus evolution and/or waning of immunity acquired through infection or vaccination. SARS-CoV-2 or COVID-19 infections and associated clinical outcomes were defined in line with definitions used by NHS England, NHS Test and Trace, UK Health Security Agency and Office for National Statistics (ONS) ²⁹⁷. I used the following definitions:

- **SARS-CoV-2 (COVID-19) infection or breakthrough infection** → A positive SARS-CoV-2 test indicating infection.
- **SARS-CoV-2 (COVID-19) symptomatic infection** → A positive SARS-CoV-2 test from an individual who reported not having an asymptomatic infection (i.e., reporting no COVID-19 symptoms). Infections or tests with no information about symptom presentation were classified as symptomatic infections.
- **SARS-CoV-2 (COVID-19) hospital assessment/admission or hospitalisation** → A positive SARS-CoV-2 test associated with a hospital episode or secondary care encounter between -1 and +14 days following the test. This includes inpatient and outpatient hospital assessments or admissions.
- **SARS-CoV-2 (COVID-19) inpatient hospitalisation** → A COVID-19 hospital assessment or admission which lasted multiple days (>1 day). This was determined by assessing the hospital admission and discharge date, whereby an inpatient hospitalisation was confirmed by a discharge date proceeding the admission date.
- **SARS-CoV-2 (COVID-19) intensive care admission** → An intensive care admission following a COVID-19 hospital assessment or admission.
- **SARS-CoV-2 (COVID-19) mortality (death)** → Any death up to 28 days following a positive SARS-CoV-2 test.

In terms of determining COVID-19 infections, NHS England and NHS Test and Trace initially used PCR testing for those with COVID-19 symptoms, and LFT testing for identification of asymptomatic cases ^{373,374,487}. This included testing in the community as well as inpatients and outpatients in secondary care settings during admission and/or prior to many procedures or treatments. However, guidelines for COVID-19 testing in the community were updated on 1st April 2022 which ended universal symptomatic and asymptomatic testing for the general public in

the community but retained testing provision for clinically vulnerable individuals and in some secondary care settings ³⁰⁷.

Data processing

The UKCCP datasets were curated through linkage of multiple data sources within a secure research environment at the UK Health Security Agency (UKHSA). I utilised R and R Studio (R version 4.0.3) and the broom (v0.7.2), dplyr (v1.0.2), lubridate (v1.7.9.2), robustbase (v0.93-6) and tidyr (v1.1.2) R packages for data cleaning and processing. I generated data for tables using the filter(), group_by() and summarize(n()) R functions and subsequently combined using the cbind() and rbind() functions of the R base package. Additional columns for processed information, such as individual age, ethnicity, and deprivation groups, were created using the mutate() and case_when() functions of the dplyr R package. During data processing, data rows where information SARS-CoV-2-related test status was missing or given as 'void' were excluded from all analyses, and rows missing information for stratified patient subgroups (demographics and cancer features) were excluded from subgroup analyses but still included in overall analyses. Missing data for these clinically relevant covariates such as demographics were converted to 'NA' prior to statistical testing and modelling in the latter case. Erroneous data entries such as a recorded age of above 120 years were processed in the same fashion as missing data. Following analyses, I imported data into Microsoft Excel for summary table generation.

COVID-19 evaluation study designs and outcome measures

Primary 2-dose SARS-CoV-2 vaccination

I implemented a case-control study design to evaluate the effectiveness of a 2-dose SARS-CoV-2 vaccine regimen in patients with cancer and non-cancer population. The study period was from the start of the COVID-19 immunisation programme in England on 8th December up to 15th October 2021, coinciding with the SARS-CoV-2 Alpha and Delta variant waves in the UK ^{477,488}. I sought to assess vaccine effectiveness against breakthrough SARS-CoV-2 infections following 2-dose vaccination across the study period (overall vaccine effectiveness) and effectiveness 3-6 months following a second vaccine dose.

To assess vaccine effectiveness, I applied a test-negative case-control (TNCC) methodology which is used as a standardised measure for vaccine surveillance studies (phase IV) and is known to have high concordance with results from randomised clinical trials (RCTs) ^{476,489}. Furthermore, this method addresses challenges that appear in observational studies relating to health interventions, including differences in healthcare-

seeking behaviours of individuals or access to infectious disease testing as previously reported^{476,489}. I calculated COVID-19 vaccine effectiveness (VE) against breakthrough infections by assessing the ratios of positive and negative SARS-CoV-2 tests in vaccinated and unvaccinated individuals at the SARS-CoV-2 test level. This methodology can also be applied to calculate vaccine effectiveness against related clinical outcomes such as hospital admission (hospital) and mortality (death) which I also investigated. The equations for calculating vaccine effectiveness using the TNCC method are outlined below:

$$VE (infection) = 1 - \left(\frac{Vaccinated, COVID - 19 positive}{Unvaccinated, COVID - 19 positive} \right) \div \left(\frac{Vaccinated, COVID - 19 negative}{Unvaccinated, COVID - 19 negative} \right)$$

$$VE (hospital) = 1 - \left(\frac{Vaccinated, COVID - 19 hospital}{Unvaccinated, COVID - 19 hospital} \right) \div \left(\frac{Vaccinated, COVID - 19 negative}{Unvaccinated, COVID - 19 negative} \right)$$

$$VE (mortality) = 1 - \left(\frac{Vaccinated, COVID - 19 death}{Unvaccinated, COVID - 19 death} \right) \div \left(\frac{Vaccinated, COVID - 19 negative}{Unvaccinated, COVID - 19 negative} \right)$$

$$Vaccine Effectiveness \% = 100 * VE$$

For this study, each data point corresponded to a single SARS-CoV-2 PCR test, with negative tests acting as an internal control which incorporates individuals who present with COVID-19 symptoms without being infected. In addition to comparing vaccine effectiveness in patients with cancer and the non-cancer population, I calculated vaccine effectiveness within cancer patient subgroups, to identify characteristics associated with lower overall vaccine effectiveness and effectiveness 3-6 months following a second dose. This was to determine whether specific cancer features are associated with lower vaccine effectiveness and/or more profound waning of vaccine protection after 3-6 months.

Booster 3-dose SARS-CoV-2 vaccination

For evaluating booster 3-dose SARS-CoV-2 vaccination, I applied a similar test-negative case-control (TNCC) methodology to my analysis of primary 2-dose vaccine effectiveness. I investigated vaccine effectiveness

against COVID-19 breakthrough infections, symptomatic infections, hospitalisation, and death following the receipt of either two or three doses of a COVID-19 vaccine within the study period from 8th December 2020 to 7th December 2021. Post-third dose effectiveness was calculated in individuals at any timepoint following a third dose booster vaccine and post-second dose effectiveness in individuals at least 20 weeks (140 days) after administration of second dose who had not received a booster. I also performed stratified analyses within cancer patient groups to determine the effect of cancer features on vaccine effectiveness against breakthrough and symptomatic infections. These subgroups included cancer subtype, date of cancer diagnosis, receipt of systemic anticancer therapy and/or radiotherapy, and COVID-19 vaccine manufacturer/combination.

Post-vaccine COVID-19 antibody responses

My analysis of COVID-19 antibody responses as a cross-sectional population-based study, incorporating antibody test data between 1st September 2021 and 4th March 2022. The study period coincided with that of the National COVID Cancer Antibody Survey (<https://covidcancersurvey.uk>) for which ethical approval was given by the Public Health England Research Ethics and Governance of Public Health Practice group (PHE REGG NR0278). The study included individuals with SARS-CoV-2 antibody test results following two or more COVID-19 vaccine doses, with the cancer cohort comprising individuals contained within the rapid registration national cancer dataset. Antibody testing was performed using the Roche Elecsys® test which provides quantification of antibodies against the SARS-CoV-2 spike protein, used as the immunogen for development of messenger RNA (BNT162b2, mRNA-1273) and adenoviral vector (AZD1222/ChAdOx1) COVID-19 vaccines, as previously described^{380,387–389}. The test has a saturation value of 25,000 units per millilitre (U/ml) with antibody responses detected with a titre greater or equal to 0.8 U/ml whereby undetectable, or negative, antibody test results indicated with a titre of below 0.8 U/ml³⁸⁰.

In this analysis, I sought to assess antibody responses (detected v undetected) and quantified titres (U/ml) in individuals with cancer in comparison with the non-cancer population and within cancer patient subgroups, including ICD-10 subtype. I then aimed to assess the relationship between antibody responses following vaccination and rates of COVID-19 breakthrough infections, defined by a positive SARS-CoV-2 PCR test, and hospital admissions.

Temporal evaluation of COVID-19

My temporal evaluation of COVID-19 was designed as a longitudinal case-control study of COVID-19 outcomes following a positive SARS-CoV-2 PCR and/or LFT test. I assessed individuals with SARS-CoV-2 infections between 1st November 2020 and 31st August 2022 to identify changes in COVID-19 disease phenotype in people

living with cancer and non-cancer population over the course of the *SARS-CoV-2* pandemic. This study period incorporated significant milestones such as the start of the UK-wide COVID-19 vaccination programme (December 2020), COVID-19 booster vaccination (September 2021), *SARS-CoV-2* Delta variant wave (April 2021-December 2021) and the first *SARS-CoV-2* Omicron variant wave (December 2021-April 2022) ^{477,478,488}.

For this study, I assessed COVID-19 case-outcome rates in patients with cancer across the COVID-19 pandemic and assessed the ongoing risk in cancer patient subgroups (between January and August 2022) relative to the non-cancer population. This included evaluation of COVID-19 hospital assessment, inpatient hospitalisation, intensive care admission and mortality to identify individuals who remain at higher risk following a *SARS-CoV-2* infection. As part of the analysis of cancer patient subgroups, I classified systemic anticancer treatments (SACT) into different classes, namely cytotoxic, B or T cell antibody, targeted, immunotherapy, and hormonal therapy. I also performed stratified analyses for SACT treatment classes A, B and C with all SACT classifications outlined in Supplementary Table 16.

Statistical analyses and data visualisation

I used R version 4.0.3 for performing statistical analyses and data visualisation. I generated summary statistics using the `summarise()` function. T-tests and Wilcoxon Rank-Sum tests required the `t.test()` and `wilcox.test` function of the R base package. Fisher/Chi tests were performed using the `fisher.test()` and `chi()` functions, respectively. I performed binary univariable and multivariable logistic regressions using the `glm(family = binomial(link = "logit"))` function, with odds ratios and 95% confidence intervals extracted using the `epiDisplay` (v3.5.0.1) R package. Multivariable regressions were adjusted by patient demographic features including age, sex (male or female), ethnicity (Asian/Asian British, Black/Black British, White/White British, and Mixed ethnic group/Other) and level of deprivation (Index of Multiple Deprivation deciles 1-10). Prior to statistical analyses, I processed the data as described in the 'data processing' section of the methods. In brief, new columns were created for subgroups of interest and missing data fields converted to 'NA' for clinically relevant covariates such as patient demographics. Outcome variables such as *SARS-CoV-2* infection, hospital admission or mortality were also converted to a numeric binary format (1 for event occurrence or 0 for no occurrence) prior to logistic regression analyses. Cancer status was also transformed into a numeric binary format, with '1' indicating a cancer diagnosis and '0' indicating no cancer diagnosis, respectively, with a similar format also used for vaccination status when assessing vaccine effectiveness (1= vaccinated, 0= unvaccinated). Generation of timeline plots, forest plots and heatmaps were performed utilising the `ggplot2` (v3.3.2, `geom_line()`, `geom_bar()`, `geom_errorbar()` and `geom_point()` functions), `forestplot` (v1.10.1) and `pheatmap` (v1.0.12) R packages, respectively.

RESULTS

Primary 2-dose SARS-CoV-2 vaccination

Cohort and data characteristics

The UKCCP cohort evaluating primary 2-dose vaccine regimens comprised 77,399,018 SARS-CoV-2 positive or negative polymerase chain reaction (PCR) tests from 28,010,955 individuals between 8th December 2020 and 15th October 2021. PCR tests were excluded due to a void SARS-CoV-2 test result (n=491,007) or could not be precisely linked to corresponding datasets through NHS identifiers (n=4,084,667). Of the linked SARS-CoV-2 PCR tests in my evaluation, 1,712,728 tests were performed by 377,194 individuals with a cancer diagnosis identified from the Public Health England Rapid National Cancer Registry which comprised the cancer cohort. 56,102 of these tests were SARS-CoV-2 positive and corresponded to 42,882 breakthrough infections. The baseline demographics of the cancer cohort and non-cancer population (study control), including age, sex, ethnicity, and index of deprivation (IMD), are shown in Table 9.

UKCCP data consisted of data from across the UK population, linking multiple datasets including NHS second-generation surveillance system, National Immunisation Management Service, Secondary Use Statistics and Contact Tracing and Advice Service, comprising one of the most comprehensive datasets within the UK Health Security Agency at the time of my analyses. Dataset linkage, processing, analyses, and storage required a high-level of computing power and server access, with some of these processes requiring a run time of several days.

SARS-CoV-2 vaccine effectiveness in patients with cancer and the non-cancer population

I implemented a test-negative case-control (TNCC) design to evaluate SARS-CoV-2 vaccine effectiveness. This method, first reported in the context of vaccines during the early stages of the pandemic ⁴⁵¹, has the advantage of utilising community diagnostic testing data and includes testing data from unvaccinated individuals as an internal control thereby not requiring an additional placebo group ^{476,489}.

In my analyses, I used two measures of assessing SARS-CoV-2 vaccine effectiveness, overall vaccine effectiveness within individuals at any time point following a second vaccine dose and effectiveness 3-6 months following a second dose against infection. In the non-cancer population, overall vaccine effectiveness was 69.8% (95% CI: 69.8-69.9) whereas this was lower in the cancer cohort at 65.5% (95% CI: 65.1-65.9). I subsequently observed that vaccine effectiveness 3-6 months post-second dose was reduced in the cancer cohort, with effectiveness against SARS-CoV-2 breakthrough infection determined to be 47.0% (95% CI: 46.3-47.6). This

waning of vaccine effectiveness in patients with cancer is notable over time with a prominent decrease between 0 and 24 weeks following second dose vaccination (Figure 33, Supplementary Table 17). However, effectiveness at 3-6 months post-second dose for protecting against infection remained higher in the non-cancer population at 61.4% (95% CI: 61.4-61.5), with less waning observed.

I subsequently compared the overall 2-dose effectiveness of the BNT162b2 (Pfizer) and ChAdOx1/AZD1222 (AstraZeneca) vaccines in the cancer cohort, identifying that vaccine effectiveness was higher in individuals (n=123,060) who had received BNT162b2 (72.1%, 95% CI: 71.6-72.7) compared to those (n=157,138) who received ChAdOx1 (59.0%, 95% CI: 58.5-59.6).

	<i>Patients with cancer (cancer cohort)</i>			<i>Non-cancer population (control cohort)</i>		
	All	PCR+	PCR-	All	PCR+	PCR-
Total	1712728	56102	1656626	75686290	5808432	69877858
Age median (IQR)	69 (58-78)	68 (56-77)	69 (58-78)	45 (29-61)	34 (20-51)	46 (30-62)
Sex						
Male	850559 (49.66)	28836 (51.40)	821723 (49.60)	29637195 (39.16)	2775160 (47.78)	26862035 (38.44)
Female	862169 (50.34)	27266 (48.60)	834903 (50.40)	45991583 (60.77)	3033061 (52.22)	42958522 (61.48)
Ethnicity						
White/White British	1533034 (89.51)	47856 (85.30)	1485178 (89.65)	55551500 (73.40)	2869777 (49.41)	52681723 (75.39)
Asian/Asian British	70859 (4.14)	3245 (5.78)	67614 (4.08)	5022431 (6.64)	359812 (6.19)	4662619 (6.67)
Black/Black British	50063 (2.92)	2051 (3.66)	48012 (2.90)	2611003 (3.45)	102911 (1.77)	2508092 (3.59)
Mixed/Other Ethnic Group	15885 (0.93)	617 (1.10)	15268 (0.92)	1267826 (1.68)	55454 (0.95)	1212372 (1.73)
Deprivation Index (IMD)						
IMD=1	129287 (7.55)	4280 (7.63)	125007 (7.55)	5735964 (7.58)	364776 (6.28)	5371188 (7.69)
IMD=2	134427 (7.85)	4390 (7.83)	130037 (7.85)	6073257 (8.02)	388336 (6.69)	5684921 (8.14)
IMD=3	143823 (8.40)	4715 (8.40)	139108 (8.40)	6252170 (8.26)	396746 (6.83)	5855424 (8.38)
IMD=4	151891 (8.87)	4339 (7.73)	147552 (8.91)	6351129 (8.39)	391737 (6.74)	5959392 (8.53)
IMD=5	157359 (9.19)	4106 (7.32)	153253 (9.25)	6296906 (8.32)	382963 (6.59)	5913943 (8.46)
IMD=6	163835 (9.57)	4371 (7.79)	159464 (9.63)	6319149 (8.35)	380069 (6.54)	5939080 (8.50)
IMD=7	168024 (9.81)	4450 (7.93)	163574 (9.87)	6103357 (8.06)	369624 (6.36)	5733733 (8.21)
IMD=8	166879 (9.74)	4178 (7.45)	162701 (9.82)	6102705 (8.06)	377014 (6.49)	5725691 (8.19)
IMD=9	168813 (9.86)	4178 (7.45)	164635 (9.94)	5958016 (7.87)	368232 (6.34)	5589784 (8.00)
IMD=10	160864 (9.39)	3913 (6.97)	156951 (9.47)	5731492 (7.57)	350993 (6.04)	5380499 (7.70)

Table 9: Table showing the baseline demographics of the UKCCP primary vaccine evaluation cohorts. Demographics including age, sex, ethnicity, and index of social deprivation (IMD). The numbers shown correspond the total number of PCR tests, positive PCR tests (PCR+) and negative PCR tests (PCR-) for each study cohort on the SARS-CoV-2 PCR test level. Brackets indicate percentages of the total number of PCR tests shown in the first numerical table row.

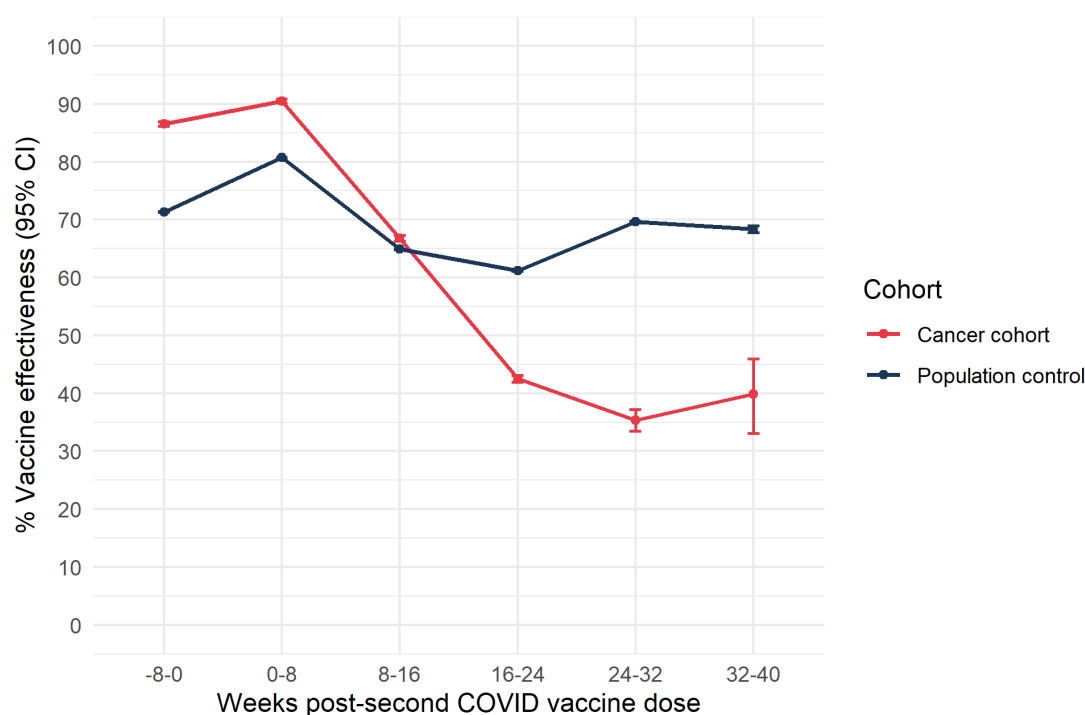


Figure 33: Temporal plot showing changes in 2-dose vaccine effectiveness against breakthrough SARS-CoV-2 infections over time in patients with cancer (cancer cohort) and non-cancer population (population control). Vaccine effectiveness was calculated in 8-week intervals before or after administration of the second vaccine dose.

Further to evaluating effectiveness against breakthrough SARS-CoV-2 infections, I sought to assess the level of protection provided by 2-dose vaccination against COVID-19 related hospital admission and mortality (death). I observed higher vaccine effectiveness for protecting against COVID-19 related hospital admission (83.3%, 95% CI 82.3-84.2) and mortality (93.4%, 95% CI 92.9-93.9) in the cancer cohort compared to breakthrough infections. Considerable protection was also provided by 2-dose vaccination 3-6 months following the second dose in patients with cancer, though vaccine effectiveness against hospital admission and mortality also waned as observed for breakthrough infections (Supplementary Table 18).

Vaccine effectiveness in cancer patient subgroups

To ascertain whether specific cancer features or characteristics were associated with differences in 2-dose vaccine effectiveness against breakthrough SARS-CoV-2 infections, I performed exploratory analyses within subgroups of patients within the cancer cohort (Tables 10 and 11).

	Overall vaccine effectiveness					Vaccine effectiveness at 3-6m				
	Exposed		Not exposed		Vaccine effectiveness (%)	Exposed		Not exposed		Vaccine effectiveness (%)
	(PCR-positive)		(PCR-negative)			(PCR-positive)		(PCR-negative)		
	2 doses (n)	Unvaccinated (N)	2 doses (n)	Unvaccinated (N)		2 doses (n)	Unvaccinated (N)	2 doses (n)	Unvaccinated (N)	
All cancer patients	18292	31649	780054	465982	65.47	12513	31649	347414	465982	46.97
Malignancy type										
Solid tumour malignancy	15070	26203	685675	390844	67.22	10245	26203	304288	390844	49.78
Haematological malignancy	3222	5446	94379	75138	52.9	2268	5446	43126	75138	27.44
Solid tumour malignancy										
Lip, oral cavity & pharynx (C00-C14)	441	684	16718	13798	46.79	297	684	7353	13798	18.52
Non-colorectal gastrointestinal (C15-C17, C22-C26)	921	2698	61577	45563	74.74	596	2698	25495	45563	60.52
Colorectal gastrointestinal (C18-C21)	2031	3740	114874	63005	70.22	1399	3740	49974	63005	52.84
Lung (C34)	1228	3344	70528	49068	74.45	820	3344	31250	49068	61.5
Respiratory & intrathoracic organs (C30-C39 excl. C34)	161	359	7376	5840	64.49	123	359	3304	5840	39.44
Bone, mesothelial and soft tissue (C40-C41, C45-C49)	283	637	14976	13091	61.16	185	637	6203	13091	38.71
Breast (C50)	3774	4877	147465	70606	62.95	2568	4877	66651	70606	44.22
Female gynaecological (C51-C58)	1095	2067	52094	33122	66.32	709	2067	23001	33122	50.61
Male urological (C60, C62, C63)	234	428	5328	4759	51.17	133	428	2294	4759	35.53
Prostate (C61)	3093	3867	108522	39592	70.82	2178	3867	50373	39592	55.73
Urinary tract (C64-C68)	1372	2223	70547	34539	69.78	968	2223	31654	34539	52.49
Central nervous system (C69-C72)	186	789	8127	11991	65.22	117	789	3506	11991	49.28
Endocrine glands (C73-C75)	251	490	7543	5870	60.14	152	490	3230	5870	43.63
Haematological malignancy										
Lymphoma (C81-C85)	1806	2427	37107	27855	44.14	1277	2427	16811	27855	12.82
Myeloma (C90)	472	918	29545	12921	77.51	345	918	13458	12921	63.92
Leukaemia (C91-C95)	809	1954	24555	32581	45.07	554	1954	11333	32581	18.49
Other	135	147	3172	1781	N/A	92	147	1524	1781	N/A

Table 10: Two-dose vaccine effectiveness against breakthrough SARS-CoV-2 infections by primary cancer subtype. Corresponding numbers of SARS-CoV-2 PCR positive and negative test results used for vaccine effectiveness calculations are provided. Primary tumour subtypes are classified by ICD-10 groupings.

When comparing primary cancer subtypes, I identified that 2-dose vaccine effectiveness was lowest in individuals with haematological malignancies, (52.9% overall vaccine effectiveness and 27.4% at 3-6 months, respectively), and that this was principally driven by those with a diagnosis of lymphoma or leukaemia (Table 10,

Figure 34, Supplementary Table 19, Supplementary Figure 4). In contrast, I observed that individuals with myeloma retained higher levels of protection (77.5% overall vaccine effectiveness and 63.9% at 3-6 months, respectively). Among individuals with solid tumour malignancies, vaccine effectiveness was lowest in those with head and neck cancers (46.8% overall vaccine effectiveness and 18.5% at 3-6 months, respectively) (Table 10, Figure 34, Supplementary Table 19, Supplementary Figure 4). Furthermore, there are notable differences in overall vaccine effectiveness and effectiveness 3-6 months post-second dose by primary tumour subtype when stratified by patient age and sex within the cancer cohort (Figure 35, Supplementary Figure 5). Perhaps most interestingly, I observed higher levels of protection in older individuals (age 80+ years) across solid tumour and haematological malignancies.

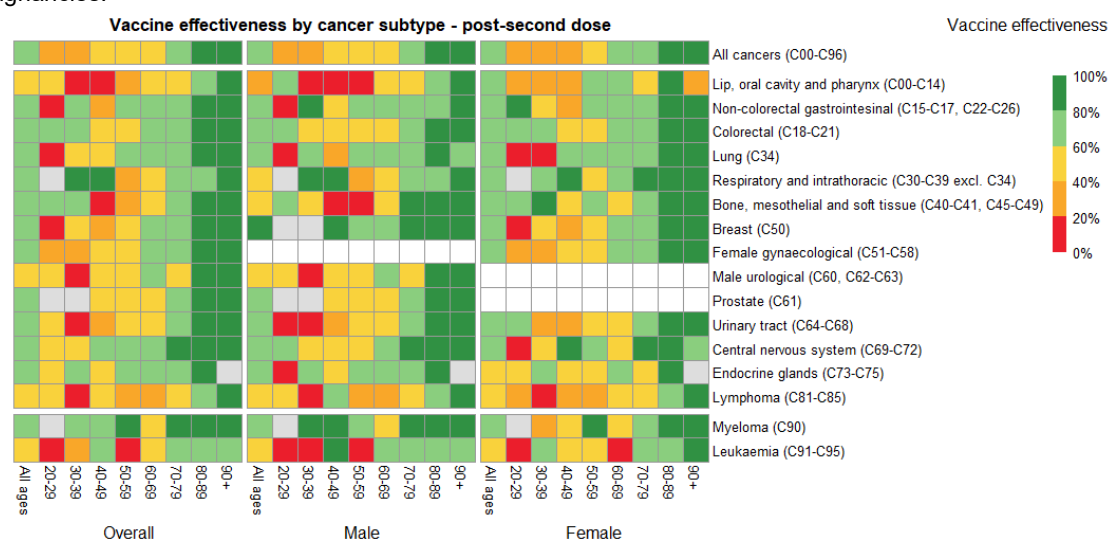


Figure 34: Heatmap showing 2-dose overall vaccine effectiveness and the association of primary tumour subtype with patient age and sex. Grey boxes denote categories with insufficient data for vaccine effectiveness calculations (<10 PCR tests). White boxes denote inapplicable sections.

Among other cancer features aside from primary subtype, I identified that prior or concurrent systemic anticancer therapy (SACT) and/or radiotherapy were associated with lower overall vaccine effectiveness and 3-6 months post-second dose in comparison to those who had not received these types of treatment (Table 11). Overall vaccine effectiveness for individuals in receipt of SACT was 61.0% (95% CI: 60.1-61.9), and 36.6% at 3-6 months post-second dose (95% CI: 35.1-38.0), compared to 67.5% (95% CI: 67.1-67.9) and 51.1% (95% CI: 50.4-51.8) for those who did not receive any SACT. Similarly for individuals receiving radiotherapy, overall vaccine effectiveness was 59.8% (95% CI: 58.6-60.9) and 36.6% at 3-6 months (95% CI: 34.7-38.5), compared to 66.5% (95% CI: 66.1-66.9) and 48.8% (95% CI: 48.1-49.4) for those who were not treated with radiotherapy. Furthermore, I observed lower vaccine effectiveness at 3-6 months for individuals who had been treated with SACT and/or radiotherapy within the last 12 months in comparison to those who had been treated over 12 months ago.

	Overall vaccine effectiveness					Vaccine effectiveness at 3-6m				
	Exposed		Not exposed		Vaccine effective-ness (%)	Exposed		Not exposed		Vaccine effective-ness (%)
	(PCR-positive)		(PCR-negative)			(PCR-positive)		(PCR-negative)		
	2 doses (n)	Unvacci-nated (N)	2 doses (n)	Unvacci-nated (N)		2 doses (n)	Unvacci-nated (N)	2 doses (n)	Unvacci-nated (N)	
All cancer patients	18292	31649	780054	465982	65.47	12513	31649	347414	465982	46.97
Cancer Stage										
Stage 1	3748	4678	139476	60749	65.1	2551	4678	64551	60749	48.68
Stage 2	2532	3387	104254	50455	63.82	1755	3387	46566	50455	43.86
Stage 3	2203	3649	109286	58389	67.74	1569	3649	48642	58389	48.39
Stage 4	966	3115	69574	47760	78.71	674	3115	30209	47760	65.79
Vaccine Name/Manufacturer										
2x BNT162b2 (Pfizer)	7050	31649	372674	465982	72.15	4667	31649	167336	465982	58.84
2x ChAdOx1 (AstraZeneca)	11192	31649	402308	465982	59.04	7828	31649	177512	465982	35.51
Mixed/Other	50	0	5072	0	N/A	18	0	2566	0	N/A
Cancer Diagnosis and Treatment										
Cancer diagnosis <12 months of data cut-off	2807	8286	162082	164729	65.57	1778	8286	63335	164729	44.19
Cancer diagnosis >12 months of data cut-off	15485	23363	617972	301253	67.69	10735	23363	284079	301253	51.27
Systemic Anti-Cancer Therapy	4633	9024	208369	158293	61	3328	9024	92068	158293	36.59
No Systemic Anti-Cancer Therapy	13659	22625	571685	307689	67.51	9185	22625	255346	307689	51.08
SACT <12 months of data cut-off	3061	6509	144513	121632	60.42	2152	6509	62253	121632	35.4
SACT >12 months of data cut-off	1572	2515	63856	36661	64.11	1176	2515	29815	36661	42.5
Radiotherapy	2576	4591	114754	82298	59.76	1823	4591	51564	82298	36.62
No Radiotherapy	15716	27058	665300	383684	66.5	10690	27058	295850	383684	48.76
Radiotherapy <12 months cut-off	911	2230	49023	50364	58.03	657	2230	21194	50364	29.99
Radiotherapy >12 months cut-off	1665	2361	65731	31934	65.74	1166	2361	30370	31934	48.07

Table 11: Two-dose vaccine effectiveness against breakthrough SARS-CoV-2 infections by patient subgroups in the cancer cohort. Corresponding numbers of SARS-CoV-2 PCR positive and negative test results used for vaccine effectiveness calculations are provided.

Vaccine effectiveness in the context of clinically relevant patient demographics

Having identified differences in vaccine effectiveness when stratifying primary cancer subtype by patient demographics, I fitted multivariable logistic regression models to adjust for clinically relevant patient demographics for the overall cohort and by patient subgroups. These demographics, which might be contributing to observed differences in vaccine effectiveness within the cancer cohort, comprised patient age, sex, ethnicity and deprivation

index (IMD). Most notably, I observed that vaccine effectiveness at 3-6 months post-second dose remains lower for individuals having received SACT within 12 months and/or a new cancer diagnosis within 12 months compared to those who had not ($p < 0.0001$ for both, Figure 35, Supplementary Table 19, Supplementary Figure 5).

In addition to older patients (80+ years), I found that individuals with stage 4 cancers had higher vaccine effectiveness which corresponded with fewer breakthrough SARS-CoV-2 infections. I therefore sought to investigate whether this could be attributed to differences in patient behaviour through linking the cancer cohort to data from the Contact Tracing and Advice Service (CTAS) which records information about the number of interpersonal contacts prior to SARS-CoV-2 infection or following exposure. From this analysis, I observed that older cancer patients and/or those with stage 4 cancers had fewer total contacts compared to those with stage 1 cancers (1.32 vs 2.04 contacts). Moreover, I was able to identify evidence of an inverse relationship between patient age and number of contacts (Supplementary Table 20, Supplementary Figure 6).

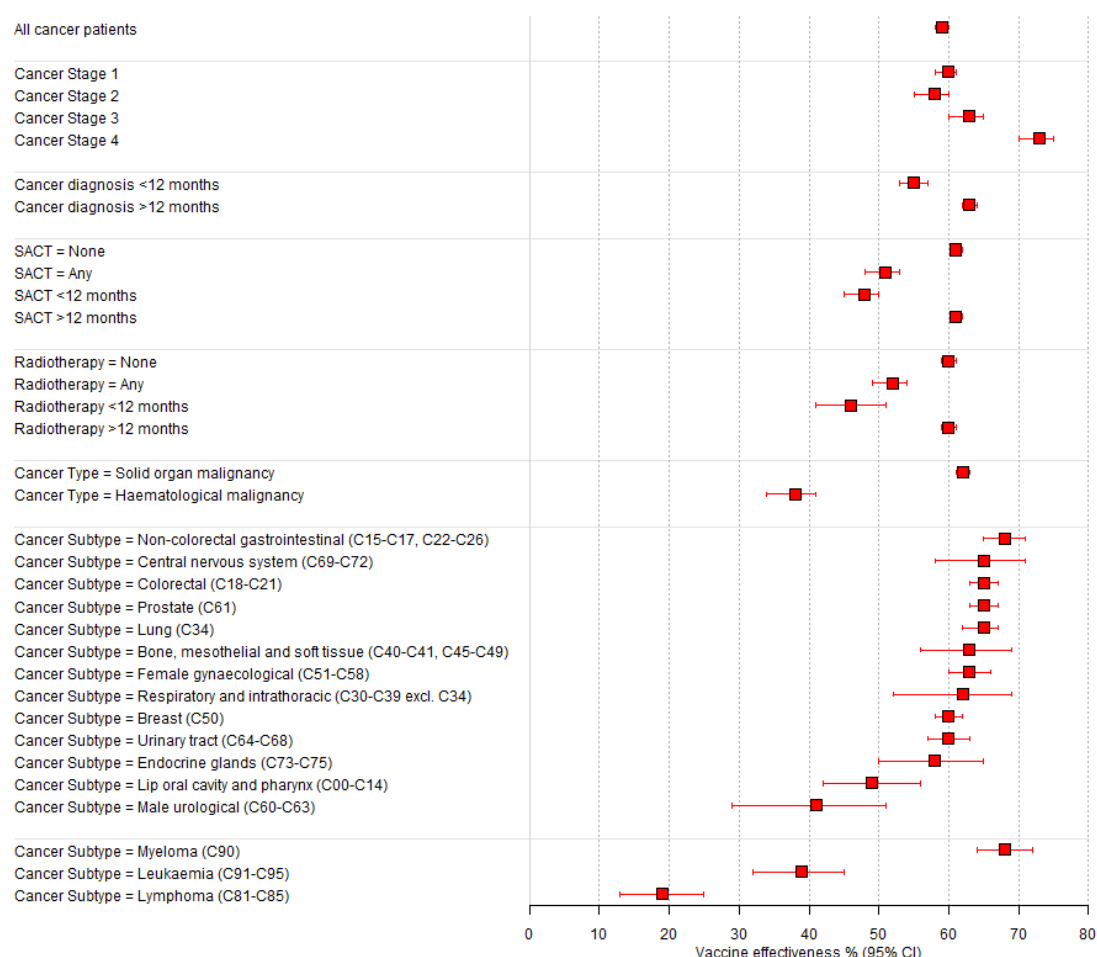


Figure 35: Forest plot showing overall vaccine effectiveness by cancer subgroups within the cancer cohort with multivariable adjustment for patient demographics. Multivariable regression models were fitted with adjustment for clinically relevant covariates of age, sex, ethnicity, and deprivation index.

Vaccine effectiveness waning by cancer subgroup

Noting differences in 2-dose vaccine effectiveness between cancer subgroups as well as waning within the overall cancer cohort 3-6 months after second dose administration, I performed further exploratory analyses to determine cancer subgroups that were more likely to experience waning vaccine effectiveness at 3-6 months. following multivariable correction. I defined waning vaccine effectiveness against breakthrough SARS-CoV-2 infection as 2-dose vaccine effectiveness over the whole study period (overall) subtracted from vaccine effectiveness at 3-6 months. Accounting for clinically relevant confounders, I observed that the waning of vaccine effectiveness was highest in individuals with a diagnosis of lymphoma or leukaemia, those with a recent diagnosis and those receiving SACT or radiotherapy following multivariable adjustment (Figure 36, Supplementary Figure 7).

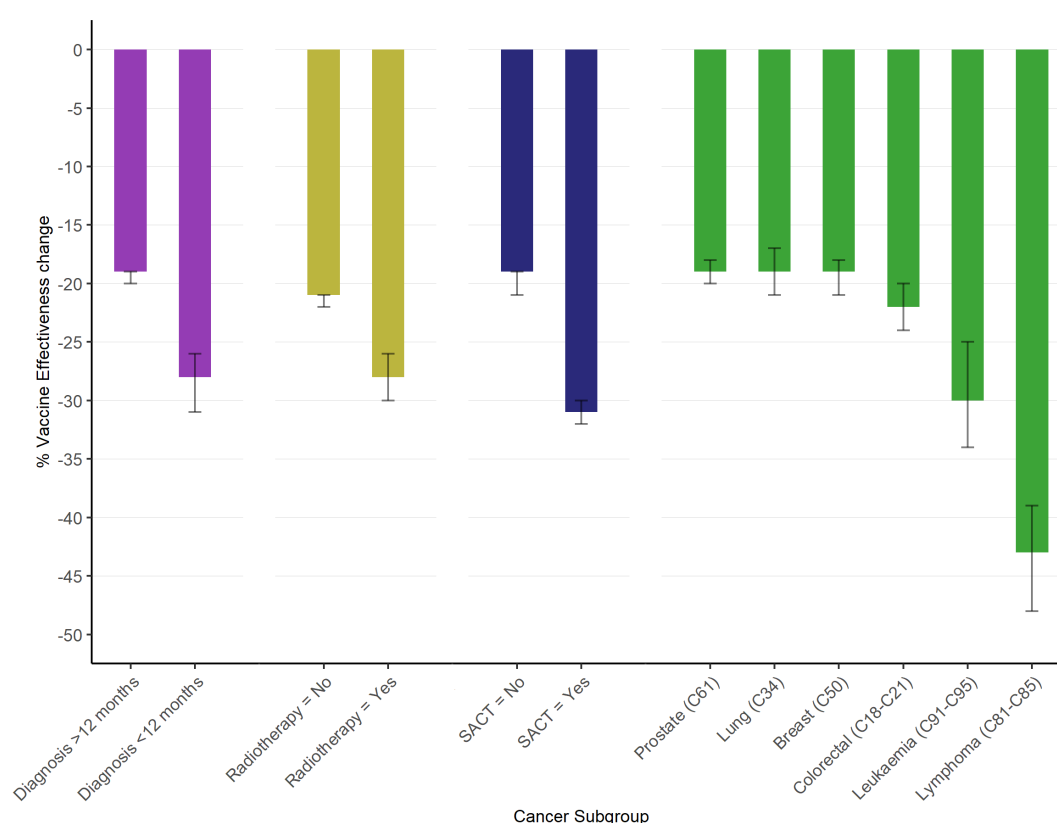


Figure 36: Waterfall plot showing waning vaccine effectiveness at 3-6 months post-second dose by key cancer subgroups. Multivariable adjustment of vaccine effectiveness was performed to account for clinically relevant confounders.

Booster 3-dose SARS-CoV-2 vaccination

Cohort characteristics

The UKCCP cohort for evaluating 3-dose booster vaccine regimens comprised 87,039,743 SARS-CoV-2 PCR tests from 29,929,073 individuals between 8th December 2020 and 7th December 2021. Of these, 2,258,553 of PCR tests were identified to have been taken by 361,098 individuals with a cancer diagnosis identified from the national cancer registry which comprised the cancer cohort. 8,371,139 individuals had received a third dose booster vaccination with 230,666 of these included in the cancer cohort. Of these individuals in the cancer cohort, 97.8% received a BNT162b2 (Pfizer) third dose and 1.5% received a mRNA-1273 (Moderna) booster. The baseline characteristics of the cancer cohort and non-cancer population (control cohort) are shown in Table 12.

SARS-CoV-2 second-dose primary and third-dose booster vaccine effectiveness

Firstly, I evaluated third-dose booster vaccine effectiveness (VE) against SARS-CoV-2 breakthrough infection, symptomatic infection, COVID-19 (coronavirus) hospitalisation and COVID-19 death in the cancer cohort and non-cancer population compared to VE of individuals who had received a primary 2-dose regimen. I observed that third dose boosters increased vaccine effectiveness against all primary outcome measures in both cohorts when compared to those who had received two doses during the study period (Figure 37, Supplementary Table 21).

In terms of primary and booster vaccine types received by individuals within the cancer cohort, I found that vaccine effectiveness was broadly similar among individuals who had received a BNT162b2 (Pfizer, messenger RNA) third dose booster following two primary vaccination doses with ChAdOx1 (AstraZeneca, adenoviral vector), compared to those who had received three BNT162b2 doses (Supplementary Table 22). However, the number of individuals who had received the mRNA-1273 (Moderna) vaccine as a third dose booster was too low to accurately assess vaccine effectiveness.

	Cancer cohort			Non-cancer population		
	Total	PCR Negative	PCR Positive	Total	PCR Negative	PCR Positive
Overall	2258553	1906816	351737	84781190	77135957	7645233
Age Group						
18-19	6704 (0.3%)	3275 (0.2%)	3429 (1.0%)	1542445 (1.8%)	1312966 (1.7%)	229479 (3.0%)
20-29	32592 (1.4%)	22606 (1.2%)	9986 (2.8%)	12268104 (14.5%)	10915482 (14.1%)	1352622 (17.7%)
30-39	79174 (3.5%)	59537 (3.1%)	19637 (5.6%)	15032642 (17.7%)	13640231 (17.7%)	1392411 (18.2%)
40-49	157804 (7.0%)	125757 (6.6%)	32047 (9.1%)	14212766 (16.8%)	12899883 (16.7%)	1312883 (17.2%)
50-59	371635 (16.5%)	305755 (16.0%)	65880 (18.7%)	15638289 (18.4%)	14397690 (18.7%)	1240599 (16.2%)
60-69	531067 (23.5%)	449631 (23.6%)	81436 (23.2%)	10557704 (12.5%)	9710645 (12.6%)	847059 (11.1%)
70-79	638202 (28.3%)	545798 (28.6%)	92404 (26.3%)	7058259 (8.3%)	6396119 (8.3%)	662140 (8.7%)
80-89	359827 (15.9%)	319012 (16.7%)	40815 (11.6%)	5729098 (6.8%)	5277821 (6.8%)	451277 (5.9%)
90+	81548 (3.6%)	75445 (4.0%)	6103 (1.7%)	2741883 (3.2%)	2585120 (3.4%)	156763 (2.0%)
Sex						
Male	1128166 (50.0%)	944990 (49.6%)	183176 (52.1%)	32588547 (38.5%)	28961653 (37.6%)	3626894 (47.4%)
Female	1130385 (50.0%)	961824 (50.4%)	168561 (47.9%)	52155575 (61.5%)	48137490 (62.4%)	4018085 (52.6%)
Ethnicity						
White/White British	2008740 (88.9%)	1716193 (90.0%)	292547 (83.2%)	64909441 (76.6%)	60035859 (77.8%)	4873582 (63.7%)
Asian/Asian British	99253 (4.4%)	76218 (4.0%)	23035 (6.5%)	5783735 (6.8%)	5194104 (6.7%)	589631 (7.7%)
Black/Black British	72380 (3.2%)	55433 (2.9%)	16947 (4.8%)	3051051 (3.6%)	2842800 (3.7%)	208251 (2.7%)
Mixed/Other Ethnic Group	22143 (1.0%)	16853 (0.9%)	5290 (1.5%)	1235403 (1.4%)	1146947 (1.5%)	88456 (1.2%)
Unknown	56037 (2.5%)	42119 (2.2%)	13918 (4.0%)	9801560 (11.6%)	7916247 (10.3%)	1885313 (24.7%)
Deprivation Index (IMD)						
IMD=Low (1-3)	631193 (28.0%)	517019 (27.1%)	114174 (32.5%)	25984682 (30.6%)	23377311 (30.3%)	2607371 (34.1%)
IMD=Medium (4-7)	932921 (41.3%)	787675 (41.3%)	145246 (41.3%)	34658310 (40.9%)	31625313 (41.0%)	3032997 (39.7%)
IMD=High (8-10)	694385 (30.7%)	602071 (31.6%)	92314 (26.2%)	24120895 (28.5%)	22118571 (28.7%)	2002324 (26.2%)

Table 12: Baseline characteristics of the cancer cohort and the non-cancer population (control). The total number of SARS-CoV-2 PCR positive and PCR negative tests are given for each baseline characteristic.

Having previously identified notable differences between cancer subgroups in primary 2-dose vaccine effectiveness against breakthrough *SARS-CoV-2* infections, I assessed third dose booster vaccine effectiveness against break-through and symptomatic infections. The frequencies of COVID-19 related hospitalisations and deaths were insufficient for accurately determining vaccine effectiveness against these clinical outcomes on a cancer subgroup level in most cases. In terms of vaccine effectiveness against breakthrough and symptomatic *SARS-CoV-2* infections, effectiveness increased for most primary cancer subtypes following a third dose booster (Figure 38, Supplementary Table 23). Moreover, in keeping with my previous analyses, third dose booster vaccine effectiveness against breakthrough and symptomatic infections was higher in individuals with solid tumour malignancies (66.0%, 95% CI: 65.5-66.4 and 69.6%, 95% CI 69.2-70.1, respectively), compared to those with a haematological malignancy (53.2%, 95% CI: 52.8-53.6 and 56.0%, 95% CI: 55.5-56.4, respectively) (Figure 38).

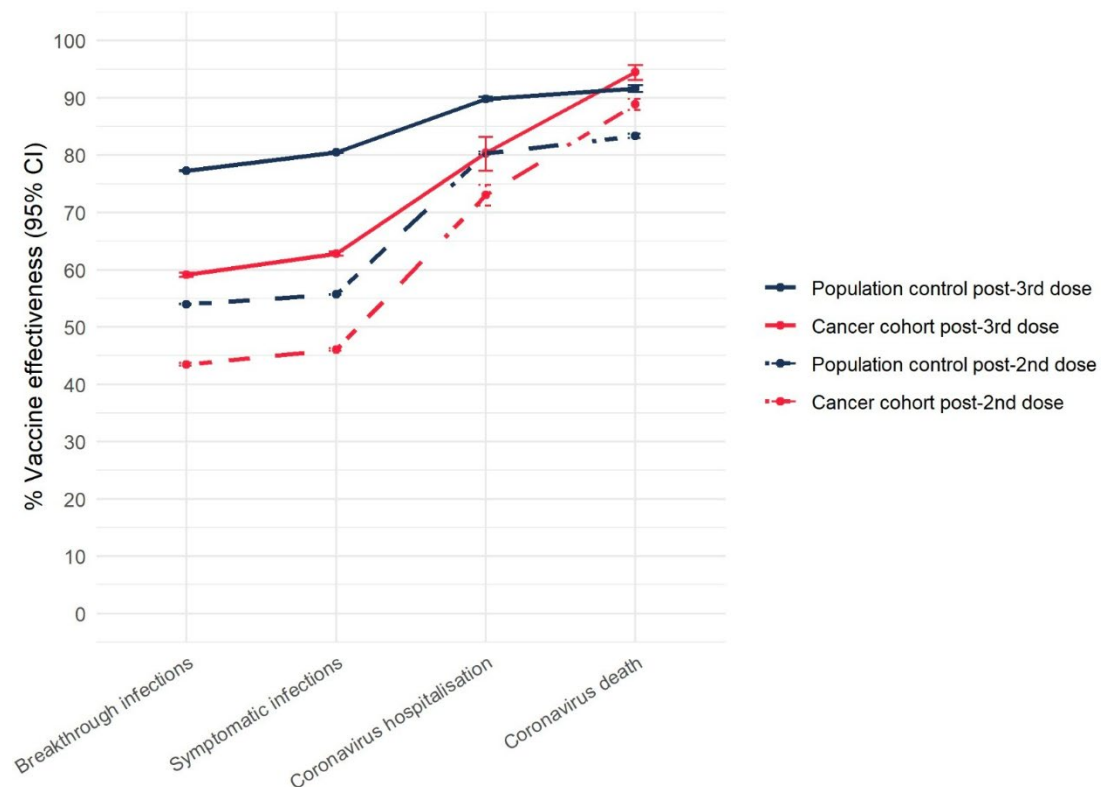


Figure 37: Vaccine effectiveness against SARS-CoV-2 breakthrough and symptomatic infections, COVID-19 related (coronavirus) hospitalisation and COVID-19 related death in the cancer cohort and non-cancer population control. Vaccine effectiveness was calculated following primary second dose and booster third dose SARS-CoV-2 vaccination. Error bars represent 95% confidence intervals for vaccine effectiveness calculations.

Specifically in individuals with a lymphoma diagnosis, I observed lower levels of vaccine effectiveness against breakthrough and symptomatic SARS-CoV-2 infections (10.5%, 95% CI: 9.9-11.1 and 13.6% 95% CI: 12.9-14.3, respectively), with no improvement after a third dose booster vaccination. However, third dose booster vaccination conferred higher levels of protection against COVID-19 related hospitalisation and death at 23.2% (95% CI: -2.4-43.5) and 80.1% (95% CI: 69.7-87.0), respectively in keeping with the overall cancer cohort. I also observed that vaccine effectiveness against breakthrough and symptomatic SARS-CoV-2 infections were lower in individuals with a recent cancer diagnosis, in recent receipt of SACT, and/or radiotherapy within 12 months following a third dose booster when compared to individuals not within these subgroups (Supplementary Table 23, Supplementary Figure 8).

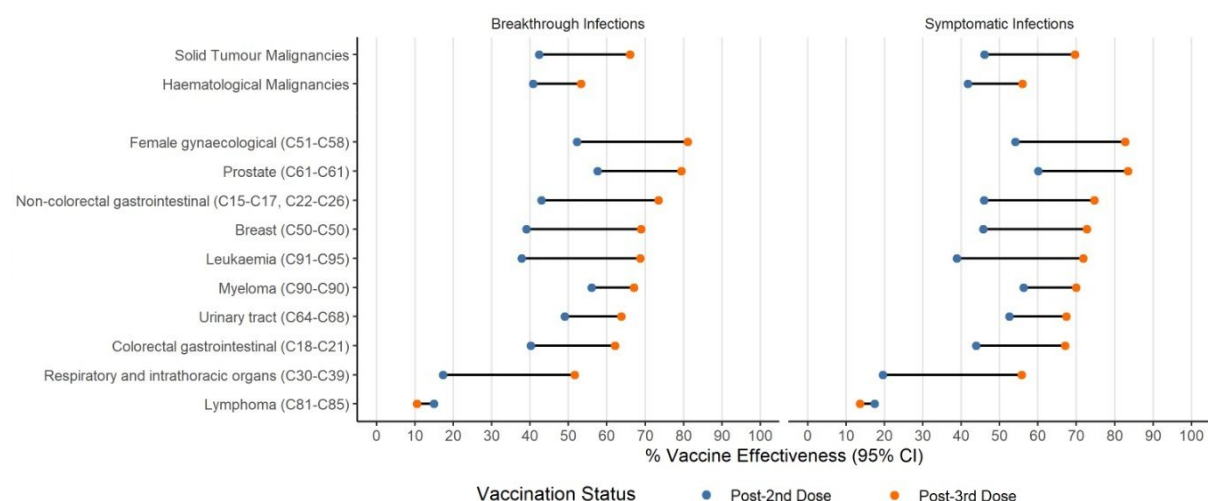


Figure 38: Vaccine effectiveness against SARS-CoV-2 breakthrough and symptomatic infections by primary tumour subtype within the cancer cohort. Vaccine effectiveness was calculated following primary second dose (blue dots) SARS-CoV-2 vaccination and booster third dose (orange dots) SARS-CoV-2 vaccination.

COVID-19 third-dose booster vaccination protection and age

My evaluation of 2-dose primary vaccine effectiveness identified differences in levels of protection by age within the cancer cohort. As the age distribution between the cancer cohort and the non-cancer population control differed, I sought to explore whether this corresponded with any differences in relative risk of SARS-CoV-2 infection and related outcomes by age group between patients with cancer and the non-cancer population. I fitted multivariable logistic regression models on PCR test data from individuals who had received three vaccination doses to assess this, with baseline characteristics outlined in Supplementary Table 24. Each model was adjusted

to account for the confounding effects of age, sex, ethnicity, deprivation index, primary dose type/manufacturer and booster dose type/manufacturer.

Across all age groups, I found that individuals within the cancer cohort were not at increased risk of breakthrough infections (OR: 0.99, 95% CI: 0.96-1.03, $p=0.770$), and were actually at reduced risk of symptomatic infections (OR: 0.73, 95% CI: 0.70-0.76, $p<0.00001$) compared to the non-cancer population. However, following a positive SARS-CoV-2 PCR test, patients with cancer were at increased risk of COVID-19 related (coronavirus) hospitalisation (OR: 3.38, 95% CI: 3.03-3.77, $p<0.0001$) and death (OR 3.01, 95% CI 2.48-3.65, $p<0.0001$) compared to individuals in the non-cancer population control cohort (Table 13, Figure 39). This trend was observable within all age categories (18-39, 40-59, 60-79, 80+ years) for COVID-19 related (coronavirus) hospitalisation and death (Table 13, Figure 39).

Age	Breakthrough Infections		Symptomatic Infections		Hospitalisation		Death (Mortality)	
	OR (95% CI)	p value	OR (95% CI)	p value	OR (95% CI)	p value	OR (95% CI)	p value
All	0.99 (0.96-1.03)	0.770	0.73 (0.70-0.76)	<0.00001	3.38 (3.03-3.77)	<0.00001	3.01 (2.48-3.65)	<0.00001
18-39	1.37 (1.09-1.71)	0.0064	0.73 (0.55-0.98)	0.033	11.50 (5.67-23.34)	<0.00001	67.15 (5.05-893.14)	0.0014
40-59	1.06 (0.98-1.15)	0.150	0.66 (0.60-0.73)	<0.00001	5.71 (4.36-7.48)	<0.00001	10.77 (5.66-20.51)	<0.00001
60-79	0.94 (0.89-0.99)	0.013	0.76 (0.72-0.81)	<0.00001	3.17 (2.75-3.66)	<0.00001	3.07 (2.33-4.05)	<0.00001
80+	1.10 (1.00-1.21)	0.043	0.93 (0.83-1.03)	0.160	1.49 (1.19-1.87)	0.00057	1.53 (1.12-2.09)	0.0073

Table 13: Multivariable logistic regression models showing the relative risk of SARS-CoV-2 breakthrough and symptomatic infections, and COVID-19 related hospitalisation and death in patients with cancer relative to the non-cancer population. All individuals in the cancer and non-cancer cohorts had received third dose booster SARS-CoV-2 vaccines. Odds Ratios and 95% confidence intervals were adjusted for clinically relevant covariates.

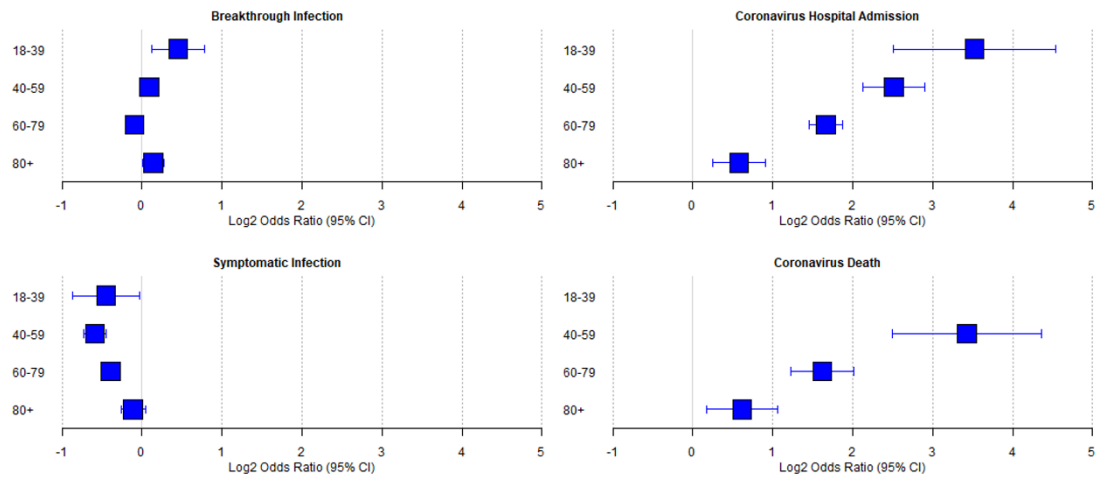


Figure 39: Forest Plots showing the relative risk of SARS-CoV-2 breakthrough and symptomatic infections, and COVID-19 related hospitalisation and death in patients with cancer relative to the non-cancer population. All individuals in the cancer and non-cancer cohorts had received third dose booster SARS-CoV-2 vaccines. A log2 transformation has been applied for displaying Odds Ratios and 95% confidence intervals which were initially calculated by fitting multivariable logistic regression models accounting for clinically relevant covariates.

Post-vaccine COVID-19 antibody responses

Cohort characteristics

The UKCCP cohort for post-vaccine SARS-CoV-2 antibody responses was formed of 4,249 antibody tests from patients with cancer between 1st September 2021 and 4th March 2022. Of these, 2,313 tests were performed following two SARS-CoV-2 vaccine doses, and 1,936 tests were taken following three SARS-CoV-2 vaccine doses. A corresponding non-cancer population control cohort was also assembled, comprising 294,230 SARS-CoV-2 antibody tests of which 230,417 were performed following two vaccine doses and 63,813 following a third dose booster vaccine. The baseline characteristics of the cancer and non-cancer cohorts are shown in Table 15.

Post-vaccination COVID-19 antibody responses in patients with cancer and non-cancer population

I first compared post-vaccination SARS-CoV-2 antibody responses in patients with cancer with the non-cancer population, assessing detection and quantitation of anti-COVID-19 spike protein (COVS) antibodies in each cohort following either two or three SARS-CoV-2 vaccine doses. From this comparison, I identified that 4.7% (199/4249) tests from the cancer cohort reported undetectable antibody responses (COVS negative) whereas only 0.1% (376/294230) of recorded tests showed undetectable antibody levels in the non-cancer population ($p<0.0001$, Table 14). I also observed that within both the cancer cohort and non-cancer population, individuals who had received a third dose booster vaccine had significantly higher COVS antibody titres than those who had only received two vaccination doses at the point of antibody testing ($p<0.0001$ for both, Table 14).

COVID-19 antibody response	Cancer cohort		Non-cancer population (control)	
Overall	4249 (100.0)		294230 (100.0)	
COVS Negative	199 (4.7)		376 (0.1)	
COVS Positive	4050 (95.3)		293854 (99.9)	
COVS Median (all)	10176.0 (1388.0-25000.0)		14518.0 (4057.0-25000.0)	

COVID-19 vaccine doses	Cancer cohort		Non-cancer population (control)	
	Total	COVS Median Titre (IQR)	Total	COVS Median Titre (IQR)
Dose 2	2313	8765.0 (1093.0-25000.0)	230417	12126.0 (2766.0-25000.0)
Dose 3	1936	11146.5 (2278.8-25000.0)	63813	23667.0 (12875.5-25000.0)
Overall (Dose 2 & 3)	4249	10176.0 (1388.0-25000.0)	294230	14518.0 (4057.0-25000.0)

Table 14: COVID-19 (COVS) antibody titres in the cancer cohort and non-cancer population. Top: Overall antibody responses following either two or three vaccinations. Bottom: COVS titres following two vaccine doses (dose2), three vaccine doses (dose 3), and overall.

	Cancer cohort			Non-cancer population		
	Dose 2	Dose 3	Overall	Dose 2	Dose 3	Overall
Overall	2313	1936	4249	230417	63813	294230
Age						
18-19	0 (0.00)	0 (0.00)	0 (0.00)	849 (0.37)	153 (0.24)	1002 (0.34)
20-29	29 (1.25)	27 (1.39)	56 (1.32)	12184 (5.29)	3941 (6.18)	16125 (5.48)
30-39	99 (4.28)	105 (5.42)	204 (4.80)	35649 (15.47)	9703 (15.21)	45352 (15.41)
40-49	349 (15.09)	273 (14.10)	622 (14.64)	75411 (32.73)	15252 (23.90)	90663 (30.81)
50-59	704 (30.44)	493 (25.46)	1197 (28.17)	61095 (26.51)	16465 (25.80)	77560 (26.36)
60-69	720 (31.13)	586 (30.27)	1306 (30.74)	32976 (14.31)	11703 (18.34)	44679 (15.19)
70-79	378 (16.34)	387 (19.99)	765 (18.00)	11308 (4.91)	5796 (9.08)	17104 (5.81)
80-89	34 (1.47)	65 (3.36)	99 (2.33)	901 (0.39)	772 (1.21)	1673 (0.57)
90+	0 (0.00)	0 (0.00)	0 (0.00)	44 (0.02)	28 (0.04)	72 (0.02)
Sex						
Male	879 (38.00)	740 (38.22)	1619 (38.10)	91068 (39.52)	23050 (36.12)	114118 (38.79)
Female	1434 (62.00)	1196 (61.78)	2630 (61.90)	139348 (60.48)	40763 (63.88)	180111 (61.21)
Ethnicity						
White/ White British	2225 (96.20)	1866 (96.38)	4091 (96.28)	216343 (93.89)	59287 (92.91)	275630 (93.68)
Asian/ Asian British	45 (1.95)	35 (1.81)	80 (1.88)	8140 (3.53)	2642 (4.14)	10782 (3.66)
Black/ Black British	21 (0.91)	9 (0.46)	30 (0.71)	1617 (0.70)	596 (0.93)	2213 (0.75)
Mixed/ Other Ethnic Group	21 (0.91)	25 (1.29)	46 (1.08)	4039 (1.75)	1198 (1.88)	5237 (1.78)
Deprivation index (IMD)						
IMD Low (1-3)	462 (19.97)	275 (14.20)	737 (17.35)	40501 (17.58)	12165 (19.06)	52666 (17.90)
IMD Medium (4-7)	904 (39.08)	799 (41.27)	1703 (40.08)	94355 (40.95)	26414 (41.39)	120769 (41.05)
IMD High (8-10)	947 (40.94)	862 (44.52)	1809 (42.57)	95505 (41.45)	25222 (39.52)	120727 (41.03)

Table 15: Baseline characteristics table of cancer cohort and non-cancer population (control). Numbers correspond to the number of antibody tests attributed to each baseline characteristic with percentages of the total number of antibody tests included in brackets (outlined in the first table row).

Post-vaccination COVID-19 antibody responses by cancer patient subgroup

Having identified differences in antibody responses between the cancer cohort and non-cancer population, I assessed antibody levels across clinically relevant subgroups within the cancer cohort including cancer stage, treatment regimens and subtype. I observed that the proportions of undetectable antibody responses were higher in individuals with a haematological malignancy (19.2%, 105/546) compared to those with a solid tumour malignancy (4.2%, 118/2791). This corresponded with a considerable difference in COVS antibody titres whereby patients with a haematological cancer diagnosis had significantly lower antibody titres than those with a solid cancer diagnosis (1872.5 U/ml vs 16,165.0 U/ml, $p<0.0001$). Furthermore, individuals with a diagnosis of leukaemia or lymphoma had the highest proportions of undetectable antibody responses and the lowest antibody titres (Figure 40, Supplementary Table 25).

In addition to primary tumour subtype, I identified that individuals who were recorded as having received systemic anticancer therapy (SACT) had lower median post-vaccination COVS antibody titres than those not receiving SACT (8,131.0 U/ml vs 15,443.0 U/ml, $p<0.0001$) (Supplementary Figure 9). However, I observed no discernible difference in antibody titres between those who were recorded as having received radiotherapy compared to those who had not (Supplementary Figure 9).

COVID-19 antibody status, breakthrough SARS-CoV-2 infection and hospitalisation

Further to assessing detectable and quantifiable antibody responses, I sought to understand if these are associated with COVID-19 related outcomes including breakthrough SARS-CoV-2 infection and hospitalisation. I observed that 259 patients with cancer went on to have a breakthrough SARS-CoV-2 infection, of which 55 patients were admitted to hospital (hospitalisation) following their SARS-CoV-2 antibody test though no deaths were recorded. Cancer patients with a breakthrough SARS-CoV-2 infection had significantly lower median antibody titres than those who were not infected during the study period (2,699.0 U/ml, IQR: 346.9-12,552.0 vs 10,961.0 U/ml, IQR: 1,611.0-25,000.0, $p<0.0001$) (Figure 41). Moreover, patients with cancer hospitalised following a SARS-CoV-2 infection had significantly lower median antibody titres than individuals who were not infected with SARS-CoV-2 or admitted to hospital (147.0 U/ml, IQR: 6.6-2104.0 vs 10,961.0 U/ml, IQR: 1,611.0-25,000.0, $p<0.0001$) (Figure 41).

In the context of risk of SARS-CoV-2 related outcomes, individuals with an undetectable antibody response upon testing were at significantly higher risk of contracting a SARS-CoV-2 breakthrough infection (OR 2.56, 95% CI: 1.69-4.00, 13.57% (27/199) v 5.73% (232/4050), $p<0.0001$) and being hospitalised following SARS-CoV-2 infection (OR 5.88, 3.13-11.11, 6.03% (12/199) v 1.06% (43/4050), $p<0.0001$) than individuals with a detectable antibody response (COVS-positive). This higher risk was still observed for SARS-CoV-2 breakthrough

infection and hospitalisation when I assessed relative risk by fitting multivariable adjusted models (accounting for patient age, sex, ethnicity and levels of deprivation) with Odds Ratios of 3.05 (95% CI: 1.96-4.72) and 6.48 (95% CI: 3.31-12.67), respectively (both $p < 0.0001$), suggesting that detectable antibody responses remained an independent risk factor.

Noting differences observed in antibody responses by the number of SARS-CoV-2 vaccines received and cancer subtype, I sought to evaluate whether this corresponded with differences in patient outcomes. I performed a sensitivity analysis to assess this, using univariable and multivariable logistic regression models on study subpopulations based the number of vaccine doses (2 or 3) received prior to the point of antibody testing and a diagnosis of either a haematological or solid tumour malignancy. From this I observed that higher risk is maintained irrespective of the number of vaccine doses (2 or 3) at the point of antibody testing and diagnosis of a haematological or solid tumour malignancy (Supplementary Table 26).

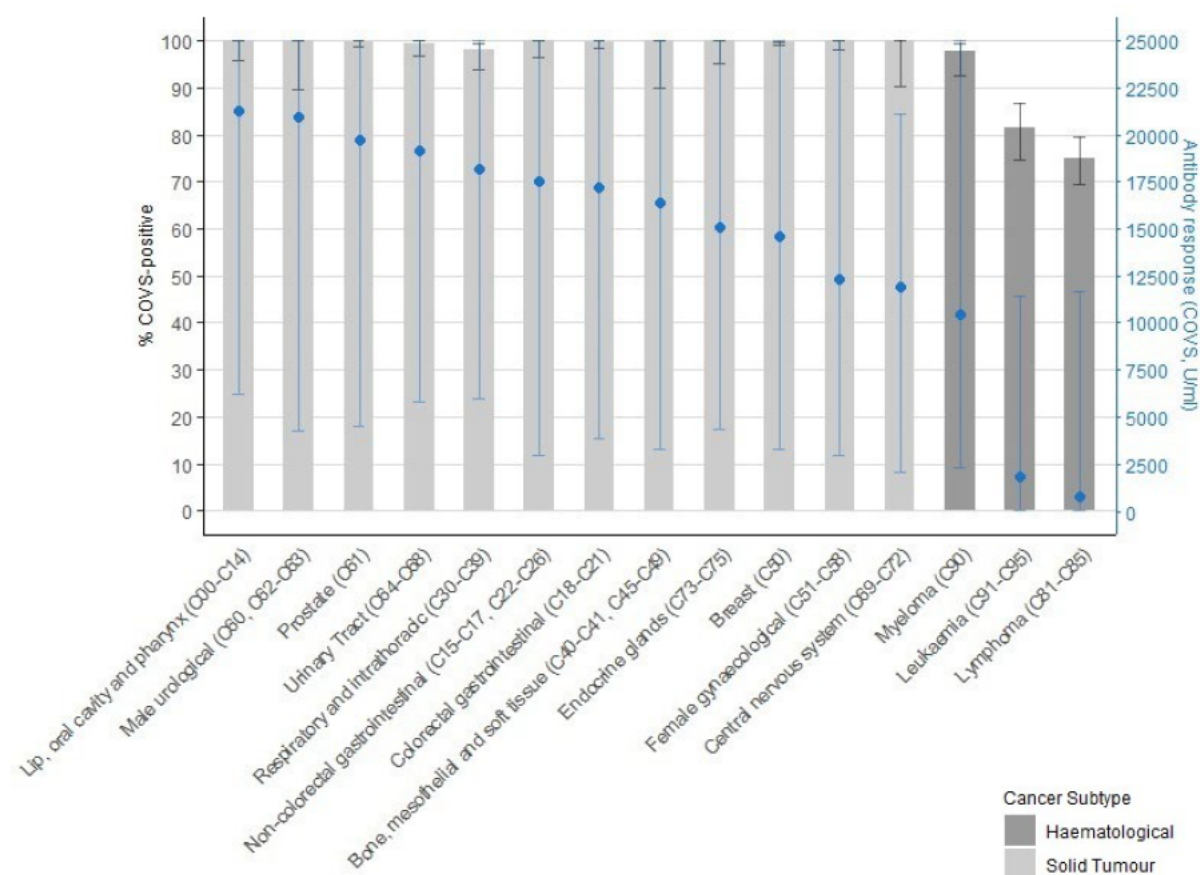


Figure 40: Median SARS-CoV-2 COVS antibody titres and detectable responses in patients with cancer by primary cancer subtype. Blue dots represent median COVS antibody titres (U/ml) with blue error bars denoting interquartile ranges. Black error bars denote 95% confidence intervals for proportions of detectable antibody responses (COVS-positivity). ICD-10 codes are included in brackets for each primary cancer subtype.

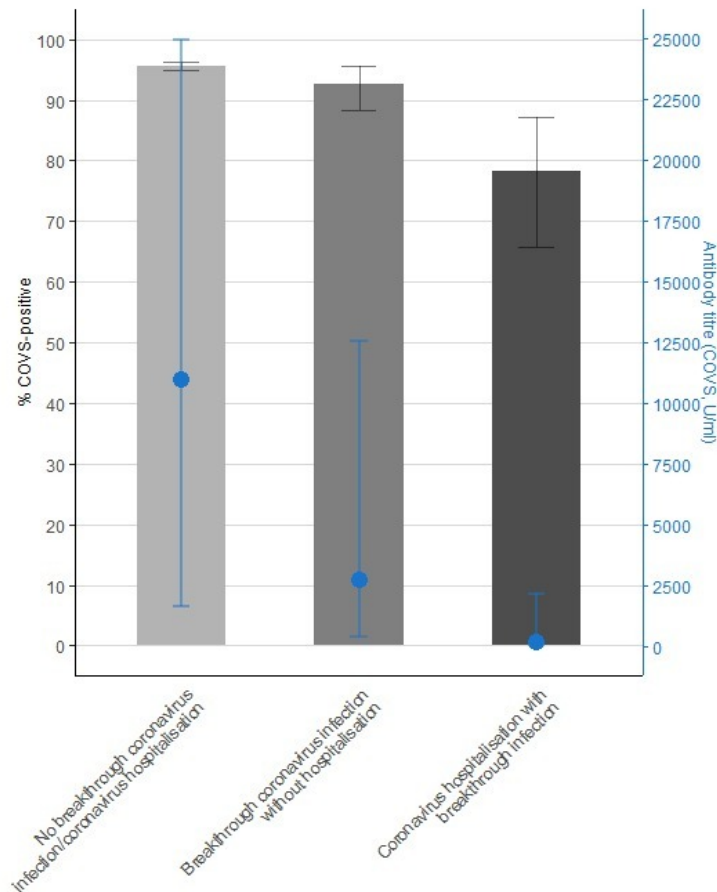


Figure 41: Median SARS-CoV-2 COVS antibody titres and detectable responses in patients with cancer by SARS-CoV-2 (coronavirus) breakthrough infection and hospitalisation status. Bars represent the proportion of detectable antibody responses (% COVS-positive responses) with black error bars denoting 95% confidence intervals. Blue dots represent median antibody titres and blue error bars denote interquartile ranges.

Quantitative SARS-CoV-2 antibody responses, breakthrough infections and hospitalisation

Having identified association of detectable antibody responses post-vaccination with patient outcomes, I fitted multivariable logistic regression models to ascertain the relationship between SARS-CoV-2 antibody titres and the risk of breakthrough infection or hospitalisation. I observed that the risk of SARS-CoV-2 breakthrough infection and hospitalisation increases for individuals with antibody titres below 5000 U/ml (Figure 42), which I confirmed upon comparing the relative risk of SARS-CoV-2 breakthrough infection and hospitalisation between individuals with titres of greater or less than 5000 U/ml. The relative risks of SARS-CoV-2 breakthrough infection and hospitalisation for individuals with titres below 5000 U/ml were 3.05 (95% CI: 2.33-4.01, 10.08% (167/1656) v 3.55% (92/2594)) and 7.22 (95% CI: 3.57-16.10, 2.72% (45/1656) v 0.39% (10/2594)), respectively (both $p < 0.0001$).

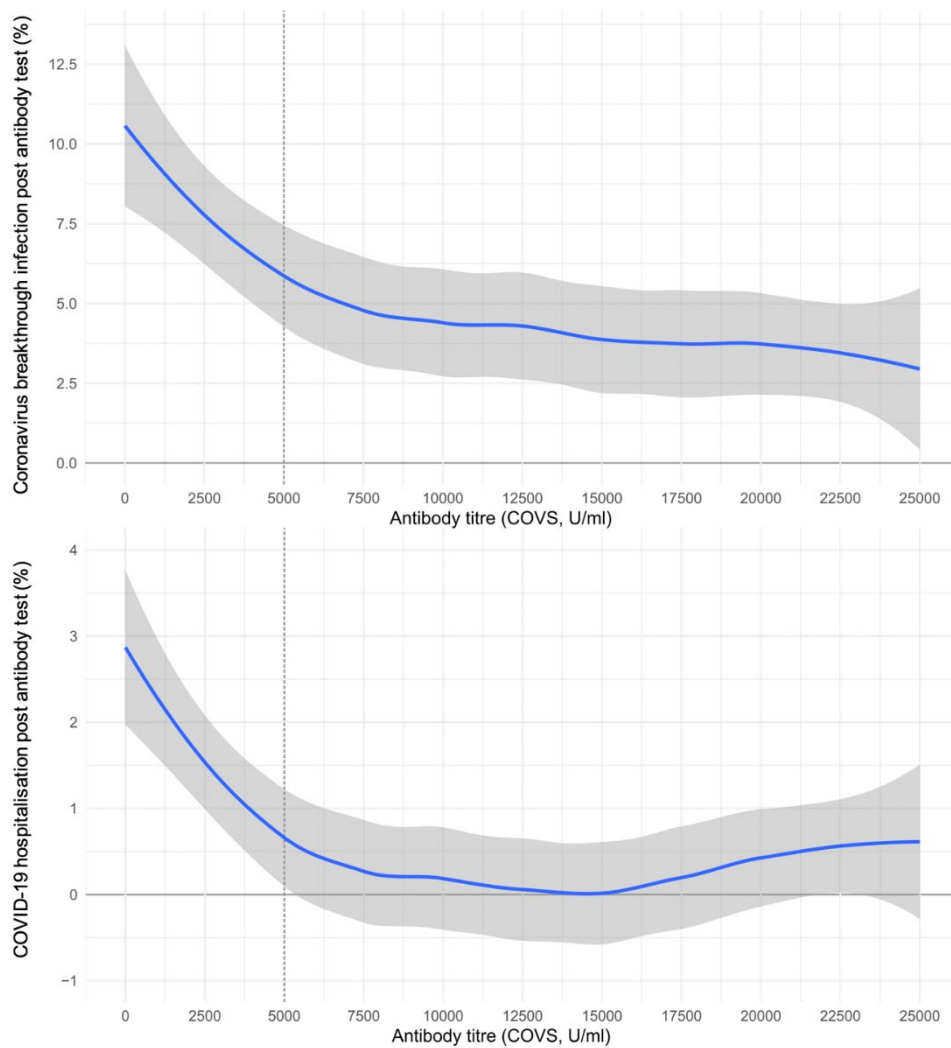


Figure 42: Multivariable logistic regression curves showing the relationship between COVS antibody titre and risk of SARS-CoV-2 breakthrough infection (top) or SARS-CoV-2 hospitalisation (bottom). The grey area represents 95% confidence intervals, and the dotted line indicates the cut-off at 5,000 U/ml where relative risk comparisons by COVS titres were performed.

I subsequently assessed the relationship between median antibody titres and SARS-CoV-2 related outcomes by primary cancer subtype. Lower median antibody titres were associated with higher case-outcome rates (%) for SARS-CoV-2 breakthrough infections and hospitalisations (Figure 43). Moreover, a higher proportion of individuals with a diagnosis of leukaemia or lymphoma were admitted to hospital following a SARS-CoV-2 infection (3.8% and 3.6%, respectively) than those with a diagnosis of myeloma or any solid tumour malignancy.

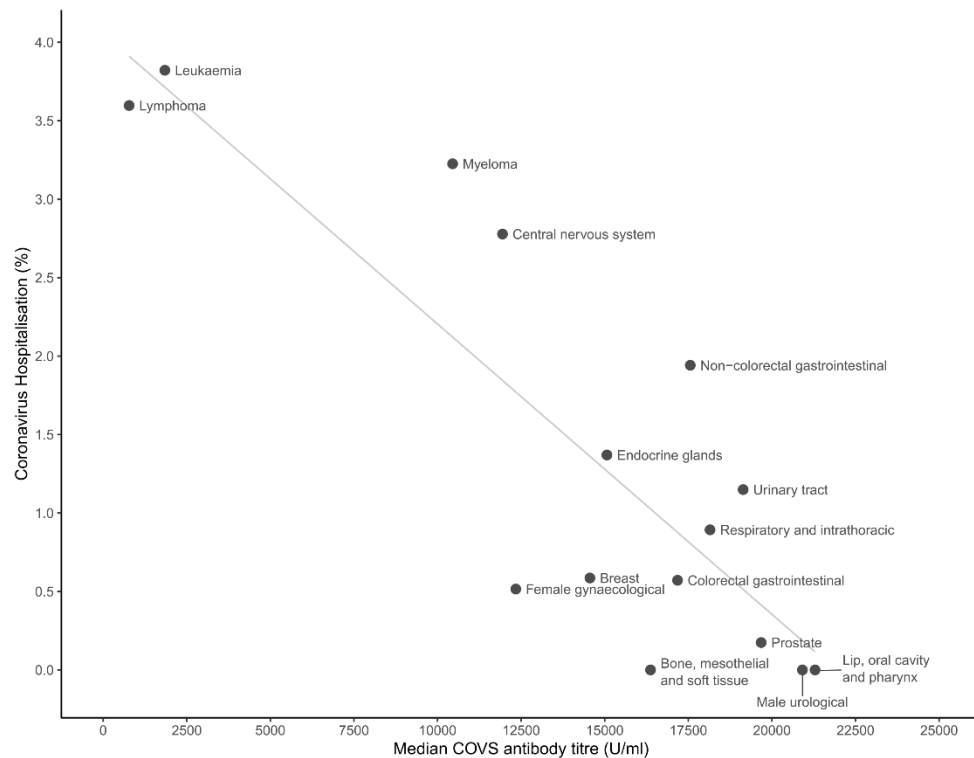
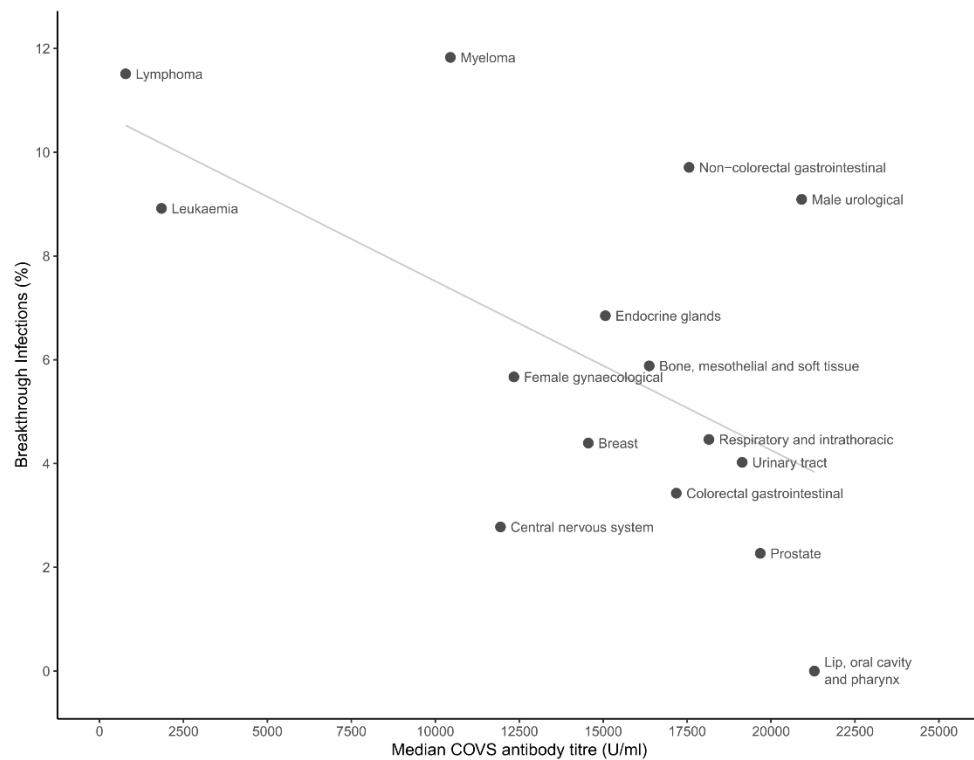


Figure 43: Scatter plot showing the relationship between median COVS antibody titre (U/ml) and SARS-CoV-2 breakthrough infections (top) and hospitalisation (bottom) by primary cancer subtype. Primary cancer subtypes were classified by ICD-10 codes.

Temporal evaluation of COVID-19 disease phenotype

Cohort characteristics

My temporal evaluation of COVID-19 disease phenotype comprised UKCCP data linked to SARS-CoV-2 positive PCR and lateral flow tests reported between 1st November 2020 and 31st August 2022. During the study period 198,819 positive SARS-CoV-2 tests, corresponding to 127,322 individual infections, were identified from individuals identified in the national cancer registry which comprised my cancer cohort. SARS-CoV-2 positive tests not linked to individuals within the national cancer registry comprised my non-cancer population cohort which contained 18,188,573 positive tests from 15,801,004 individual SARS-CoV-2 infections. Within the cancer cohort, 39,033 SARS-CoV-2 positive tests were associated with a hospital assessment, 28,061 with inpatient hospitalisation, 2,168 with intensive care admission and 15,278 with COVID-19 mortality (death) (Table 16). The baseline characteristics of the cancer and non-cancer cohorts comprising patient demographics and vaccination status are outlined in Supplementary Table 27.

Clinical Outcomes	Patients with cancer		Non-cancer population	
	Tests	Patients	Tests	Patients
SARS-CoV-2 infection (overall cohort)	198,819	127,322	18,188,573	15,801,004
Hospital assessment	39,033	23,264	837,081	536,566
Inpatient hospitalisation	28,061	16,261	671,144	412,087
Intensive care admission	2,168	1,136	68,088	36,162
COVID-19 mortality	15,278	9,580	184,498	113,043

Table 16: Clinical outcomes of the UKCCP cancer cohort and non-cancer population control at the SARS-CoV-2 positive test level and individual participant/patient level. SARS-CoV-2 positive tests included results from PCR and lateral flow testing.

Temporal changes in COVID-19 outcomes in cancer and non-cancer populations

During the study period from November 2020 to August 2022, I observed that the disease phenotype of the SARS-CoV-2 virus became less severe in both patients with cancer and the non-cancer population. COVID-19 hospital assessment and hospitalisation rates in cancer patients decreased from 35.73% and 30.58% in early 2021 to 13.26% and 7.45% by spring 2022, respectively, with similar decreases also noted in the non-cancer population (Figure 44, Supplementary Table 28). Furthermore, within the cancer and non-cancer cohorts, I observed COVID-19 case-intensive care admission rates of less than 0.5% in Spring 2022 as opposed to 2.54% and 1.52% in early 2021, respectively (Figure 44). The COVID-19 case-mortality rate in patients with cancer also decreased in this

time period from 20.62% in late 2020/early 2021 to 3.25% by spring 2022 (Figure 44), with a similar decrease from 4.16% to 0.40% in case-mortality rate within the non-cancer population was noted. This corresponded with increased uptake of one or more COVID-19 vaccine dose(s) over the course of the study period, with over 90% vaccine uptake within both the cancer and non-cancer study populations by Spring 2022 (Figure 45, Supplementary Table 29).

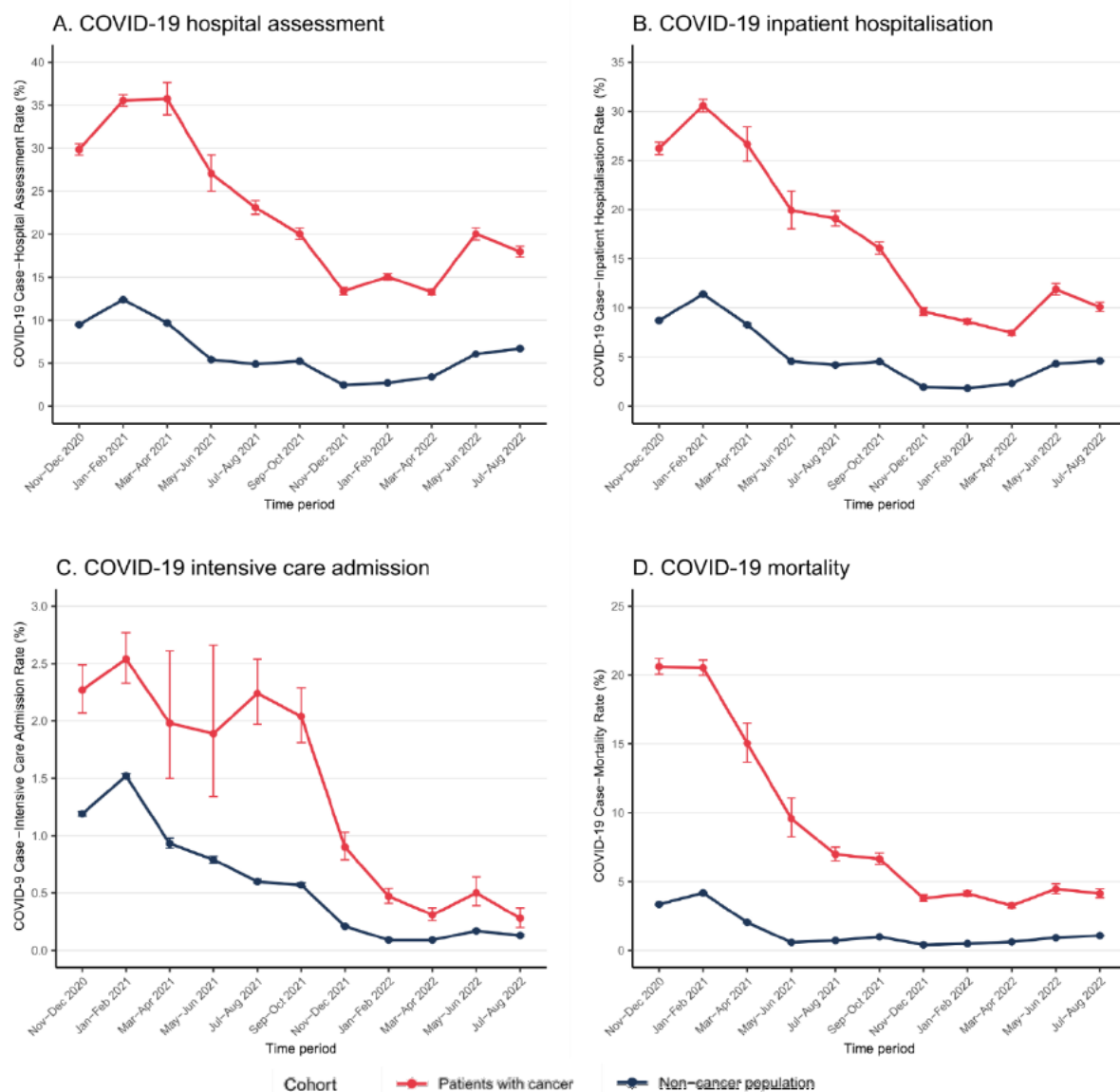


Figure 44: Temporal trends of case-outcome rates between patients with cancer and the non-cancer population between 1st November 2020 and 31st August 2022. Case-outcome rates (%) with 95% confidence intervals are shown at 2-month intervals for COVID-19 hospital assessment (A), inpatient hospitalisation (B), intensive care admission (C) and mortality (D).

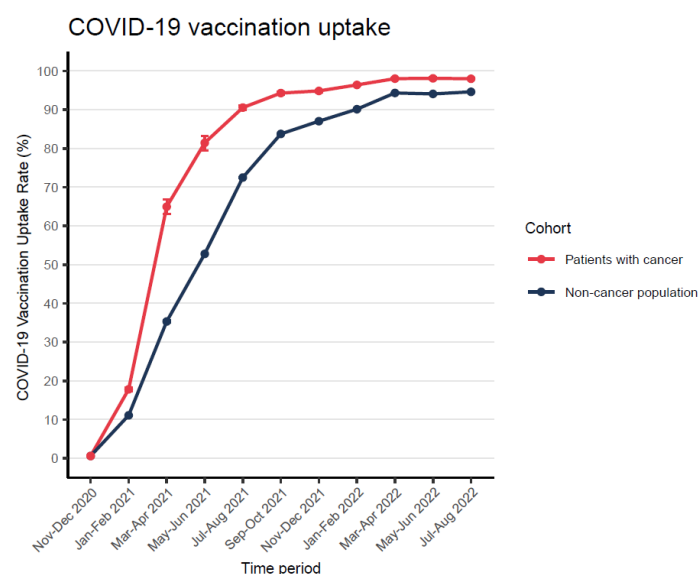


Figure 45: Temporal plot showing the proportion of SARS-CoV-2 positive tests from vaccinated individuals within the cancer and non-cancer study populations. The proportions of positive tests from vaccinated individuals were calculated at each timepoint between 1st November 2020 and 31st August 2022. Individuals were vaccinated if they have received at least one SARS-CoV-2 vaccination dose.

COVID-19 disease phenotype in 2022

Having identified temporal trends in COVID-19 disease phenotype throughout the pandemic, I sought to understand the relative risk of severe COVID-19 related outcomes in patients with cancer compared to the non-cancer population in 2022. This coincided with Omicron (BA.1-BA.5) becoming the predominant SARS-CoV-2 variant in circulation within the UK population. Through performing logistic regression models, I observed that the rate of COVID-19 hospital assessment was 5.34x (95% CI: 5.23-5.45) higher for patients with cancer which corresponded with a 4.22x (95% CI: 4.11-4.33) higher rate of inpatient hospitalisation (Supplementary Figure 10). When adjusting for clinically relevant demographics including age, sex, ethnicity, and deprivation index, I identified that the relative risks of COVID-19 hospital assessment and inpatient hospitalisation in patients with cancer were 3.02x (95% CI: 2.95-3.08) and 2.10x (95% CI: 2.04-2.16), respectively (Figure 46). I found that this also corresponded with increased risk of COVID-19 intensive care and mortality for patients with cancer, with risk 2.53x (95% CI: 2.24-2.86) and 2.54x (95% CI: 2.44-2.65) greater compared to the non-cancer population, respectively, when adjusted for age, sex, ethnicity and deprivation (all $p < 0.00001$) (Figure 46, Supplementary Figure 10, Supplementary Table 30).

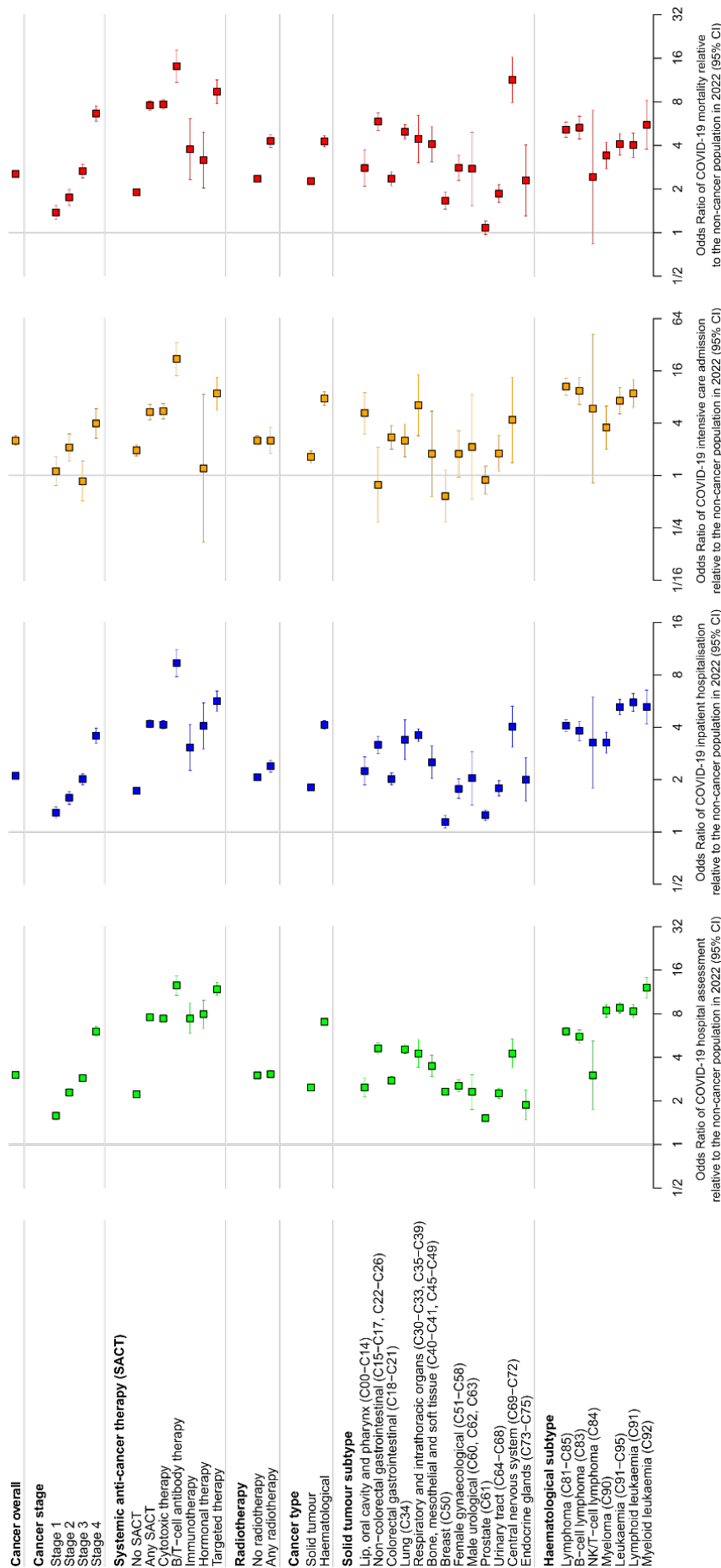


Figure 46. Relative risk of severe clinical outcomes in subgroups of patients with cancer compared to the non-cancer population following a SARS-CoV-2

infection. Multivariable logistic regression models adjusting for age, sex, ethnicity and deprivation show the relative risk of COVID-19 hospital assessment (green), inpatient hospitalisation (blue), intensive care admission (orange) and mortality (red). Odds ratios are used to approximate relative risk with 95% confidence intervals. Corresponding ICD-10 codes for specific primary tumour subtypes are listed in brackets.

I then assessed the relative risk of each COVID-19 related outcome (hospital assessment, inpatient hospitalisation, intensive care admission, mortality) within subgroups of patients with cancer compared to the non-cancer population, including assessment of cancer stage, cancer treatment and primary tumour subtype. From this, I identified that specific subgroups were associated with significantly higher risk including individuals with a haematological malignancy (particularly leukaemia), those receipt of certain types of SACT (cytotoxic therapy, B/T cell antibody therapy, targeted therapy) and individuals with a stage 4 cancer (Figure 46, Supplementary Figure 10).

COVID-19 disease phenotype by primary cancer subtype

Having identified differences in COVID-19 related outcomes between different subgroups in 2022, I performed a more granular assessment of case-outcome rates across the COVID-19 pandemic by primary cancer subtype. Firstly, comparing individuals with a haematological or solid tumour malignancy, I identified similar temporal changes in COVID-19 hospital assessment, inpatient hospitalisation, intensive care admission and mortality rates for each cancer type between November 2020 and August 2022 (Supplementary Figure 11, Supplementary Table 31). However, rates for each COVID-19 related outcome remained higher in patients with haematological malignancies at the end of the study period.

I subsequently built upon my stratified analyses of COVID-19 related outcomes by cancer subtype (ICD-10 classified) in 2020 (chapter 5) and calculated case-outcome rates for individual cancer subtypes across different age groups and sex in late 2020, 2021 and 2022. From this, I observed that case-outcome rates across most cancer subtypes were lower in 2022 compared to both 2020 and 2021 (Supplementary Figure 12). This corresponded with SARS-CoV-2 infection waves driven by Omicron variants (e.g., BA.1, BA.2) and when most patients with cancer had received one or more COVID-19 vaccine doses. In keeping with my observations of temporal trends, I identified that Individuals with haematological malignancies remained at greater risk than those with solid tumour malignancies, with elevated risks also observed to a lesser extent in those with lung cancers across the study period (Figure 47, Supplementary Figure 12). However, for COVID-19 mortality in 2022, patient age (>80 years) was more strongly associated with higher case-mortality rates than primary tumour subtype (Figure 47). Within the cancer cohort in this time period, the case-mortality rate for patients with cancer aged 80+ was 10.32% (95% CI: 9.87-10.78) in comparison with 2.83% (95% CI: 2.73-2.94) for those aged 18-79. This was higher than the case-mortality rate for individuals with haematological malignancies in 2022 (5.68%, 95% CI: 5.33-6.04).

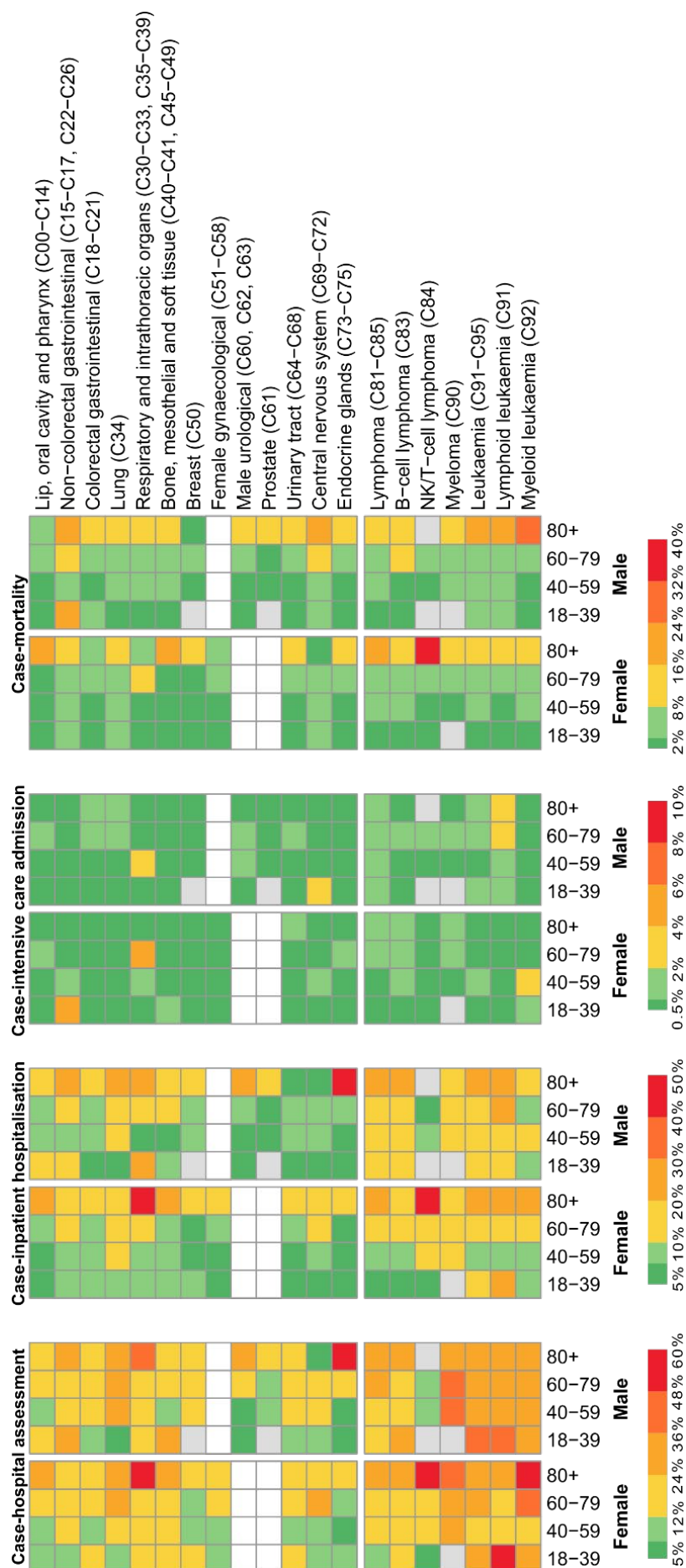


Figure 46: Case-outcome rates of solid and haematological cancer subtypes by age group and sex (female or male) following a SARS-CoV-2 infection in 2022.

Case-outcome rates for COVID-19 hospital assessment, inpatient hospitalisation, intensive care admission and mortality are shown for each cancer subtype and are colour coded according to the scales below each heatmap. Grey boxes denote <10 positive SARS-CoV-2 tests within a subgroup. White boxes denote <10 positive SARS-CoV-2 tests within sex-specific cancer subtypes (female gynaecological, male urological, prostate cancers). Corresponding ICD-10 codes for specific primary tumour subtypes are listed in brackets.

COVID-19 disease phenotype by systemic anticancer therapy (SACT) types and classes

In addition to primary cancer subtype, my univariable and multivariable regression models identified that patients in receipt of systemic anticancer therapy (SACT) were at higher risk of severe outcomes following a SARS-CoV-2 infection than other subgroups of patients with cancer relative to the non-cancer population. This also corresponded with my evaluations of 2-dose primary and 3-dose booster vaccine effectiveness which indicated lower levels of protection provided by vaccination for individuals receiving SACT. Considering this, I performed a more granular analysis of case-outcome rates across the COVID-19 pandemic by SACT types and treatment classes (A/B/C). My initial temporal comparison within the cancer cohort of individuals in receipt of SACT with those not receiving SACT showed similar decreases in case-outcome rates of COVID-19 hospital assessment, inpatient hospitalisation, intensive care admission and mortality between November 2020 and August 2022 (Supplementary Figure 13, Supplementary Table 32).

Subsequently, I calculated case-outcome rates for individual SACT types and classes in late 2020, 2021 and 2022 to assess temporal trends in more detail. I observed that the case-outcome rates for all outcomes decreased across all SACT types and classes in 2022 compared to 2020 and 2021, most notably with significant drops in COVID-19 case-mortality rates (Figure 48). However, case-mortality rates in 2022 remained higher in individuals recorded as receiving hormonal therapy (8.1%) and/or targeted therapy (6.9%) compared to those receiving no SACT or other SACT types (all <6.0%). 2022 COVID-19 case-outcome rates for hospital assessment and inpatient hospitalisation were also highest for individuals in receipt of targeted treatment (36.5% and 16.2%, respectively) with those receiving a SACT class C treatment regimen at higher risk of intensive care admission (1.8%).

Noting higher case-outcome rates within some SACT classes, I sought to assess the relative risk of patients with cancer in receipt with SACT compared to those not receiving treatment in 2022, coinciding with high vaccine uptake and Omicron becoming the predominant circulating SARS-CoV-2 variant. I observed that the relative risk of all COVID-19 related outcomes for individuals in receipt of various SACT regimens was greater than 1.0 in nearly every instance when compared to individuals not receiving SACT (Figure 49, Supplementary Table 33). For individuals receiving any SACT, the relative risk of hospital assessment, inpatient hospitalisation, intensive care admission and mortality were 2.29 (95% CI: 2.21-2.37), 1.58 (95% CI: 1.50-1.65), 1.79 (95% CI: 1.45-2.21) and 1.83 (95% CI: 1.71-1.96), respectively (all $p<0.0001$). Individuals receiving targeted therapy were at the highest risk of hospital assessment 3.95 (95% CI: 3.63-4.31) and inpatient hospitalisation 2.26 (95% CI: 2.01-2.52) relative to those not receiving treatment (both $p<0.0001$). Furthermore, receipt of SACT class C and hormonal therapy were most indicative of highest risk of COVID-19 intensive care admission (SACT class C: 5.67, 95% CI: 2.41-11.44, $p=0.0001$) and mortality (hormonal: 2.60, 95% CI: 1.88-3.53, $p<0.0001$), respectively (Figure 49, Supplementary Table 33).

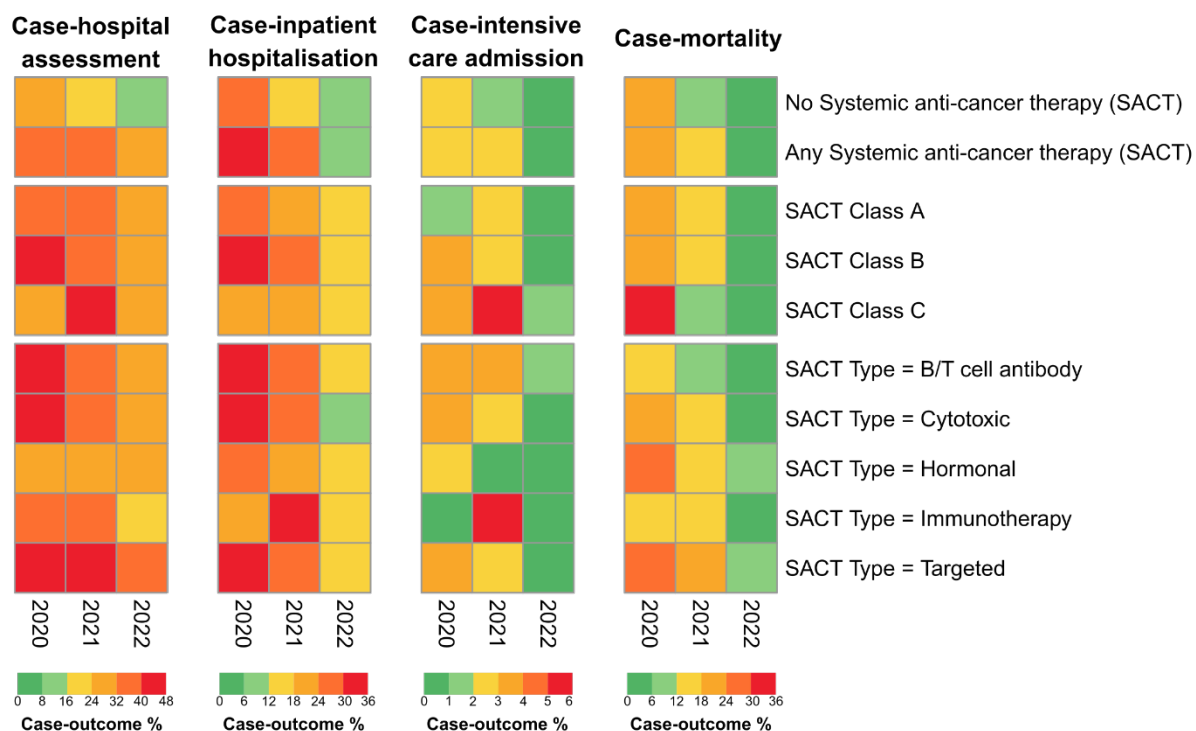


Figure 48: Case-outcome rates of SACT treatments and treatment classes following a COVID-19 infection in 2020, 2021 and 2022. Case-outcome rates for COVID-19 hospital assessment, inpatient hospitalisation, intensive care admission and mortality are shown for each cancer subtype and are colour coded according to the scales below each heatmap.

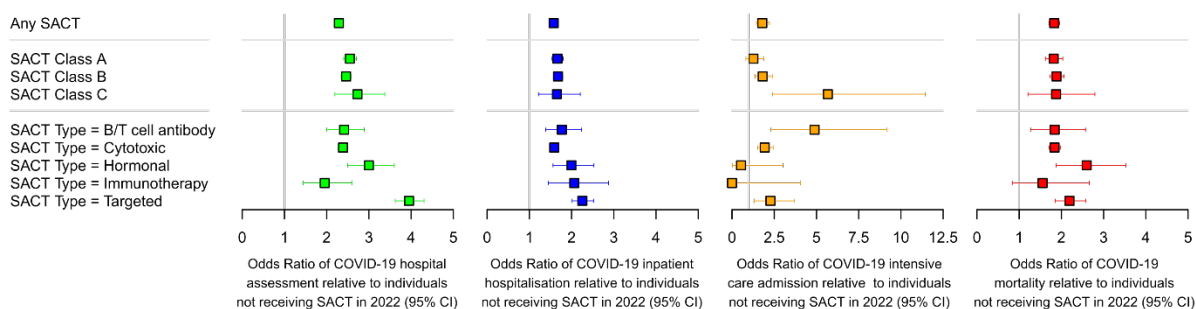


Figure 49: Relative risk of severe clinical outcomes in patients with cancer receiving SACT compared to those not receiving SACT following a COVID-19 infection in 2022. Univariable Fisher's Exact testing shows the relative risk of COVID-19 hospital assessment (green), inpatient hospitalisation (blue), intensive care admission (orange) and mortality (red). Odds ratios are used to approximate relative risk with 95% confidence intervals.

DISCUSSION

My population-scale evaluations of the impact of Coronavirus Disease 2019 (COVID-19) in patients with cancer have successfully identified subgroups at higher risk of severe COVID-19 related clinical events, assessed the effectiveness of *SARS-CoV-2* vaccines for preventing infection and associated outcomes and how the COVID-19 disease phenotype has evolved over the course of the pandemic in the United Kingdom. These analyses, within the scope of the United Kingdom Coronavirus Programme (UKCCP), were the first to be performed using fully integrated national population level health data whereas previous studies typically comprised small patient cohorts and/or combined data reported from a few cancer centres^{289,291,292}. This has grown over the course of the pandemic to incorporate COVID-19 related data beyond that of equivalent registries including OnCovid, Isaric and US CCC-19^{490–492}, with my *SARS-CoV-2* booster vaccine effectiveness dataset comprised 87 million PCR tests, with 2.25 million of those linked to a cancer diagnosis, in the 12-month period proceeding mass vaccination in the UK alone (December 2020–December 2021). Perhaps more importantly, this national-level single registry has enabled me to perform analyses with high granularity and statistical power as well as dissemination in a timely manner to inform oncologists and medical professionals of the ongoing risks of COVID-19 in patients with cancer.

More specifically, my evaluations identified that the primary and secondary effects of COVID-19 have changed markedly for patients with cancer and non-cancer population from 2020. This is reflected in trend globally, with a less severe disease course experienced by the majority of individuals, coinciding with the evolution of *SARS-CoV-2* variants, development and deployment of *SARS-CoV-2* vaccines, availability of additional treatment options such as antivirals and prophylactic antibodies, and changes in clinical practices to promote better infection control and disease management. However, patients with cancer remain at higher risk of severe clinical outcomes following a *SARS-CoV-2* infection compared to the non-cancer population, with specific patient subgroups disproportionately so. Thus, continued COVID-19 disease monitoring and deployment of therapeutic strategies remain important to mitigate the risks presented by COVID-19.

I firstly assessed the effectiveness of *SARS-CoV-2* vaccines within individuals with a cancer diagnosis and the wider non-cancer population. Implementing a test-negative case-control (TNCC) study design, my evaluation represents the first use of COVID-19 diagnostic (PCR) testing data to determine vaccine effectiveness in patients with cancer which has been replicated in other vaccine studies assessing clinically vulnerable population groups^{493–495}. I identified that a two-dose (primary) vaccination regimen provided protection for patients with cancer and the non-cancer population from *SARS-CoV-2* infection within the study period, though this was 4.3% lower in individuals with a cancer diagnosis (65.5% versus 69.8%). At three to six months post-vaccination, effectiveness against *SARS-CoV-2* infection was 14.4% lower in patients with cancer than the non-cancer population,

representing a reduction of nearly a third in vaccine effectiveness in patients with cancer as well as a more rapid waning of vaccine protection. This rapid waning could be in part have been driven by behavioural changes, including shielding, thereby reducing the likelihood of viral transmission, and resulting in a higher calculated vaccine effectiveness. In support of this, I show that patients with cancer are typically older than the non-cancer population, with median ages of 69.0 versus 45.0 years respectively, and that older individuals within the cancer cohort have fewer interpersonal contacts prior to *SARS-CoV-2* infection or following exposure. I also observed lower vaccine effectiveness in individuals with a diagnosis of lymphoma or leukaemia, new diagnosis of any cancer subtype, and/or those who had received radiotherapy or SACT within the preceding 12 months. These findings are in keeping with clinical data from cancer cohorts in the United States (n=184,485) and France (n=2391) assessing two-dose vaccine effectiveness ^{496,497}. Waning vaccine effectiveness has also been reported in other clinically vulnerable non-cancer populations, with a UK Health Security Agency (UKHSA) study identifying effectiveness waning of 25% at week 20 post-vaccination within a clinically extremely vulnerable subgroup incorporating multiple different comorbidities including those with severe respiratory diseases ^{498,499}.

Although vaccine effectiveness waned more quickly in patients with cancer, I also found that vaccine effectiveness against breakthrough *SARS-CoV-2* infection, hospitalisation and mortality increased for the majority following a third dose booster. This is reassuring and indicates the continuing importance of *SARS-CoV-2* vaccination booster provision for people living with cancer and other immunocompromised population subgroups as COVID-19 becomes an endemic disease. However, my analyses also show that patients with cancer retain lower vaccine effectiveness than the non-cancer population following a third-dose booster and that the benefits of booster vaccination are heterogeneous within the cancer population. Lower booster vaccine effectiveness is observable in individuals with a with haematological malignancy (especially lymphoma), those with a recent cancer diagnosis and/or having received anticancer treatments in the last year in turn corresponding with higher risk of adverse outcomes following *SARS-CoV-2* infection. This is contrary to a previous UK study of 2515 patients with cancer which had identified that recent SACT was not associated with poorer post COVID-19 outcomes ²⁸³, though this was before the rollout of vaccines with my analyses subsequently showing reduced vaccine effectiveness in individuals receiving SACT. Moreover, patients with cancer remain at higher risk of hospitalisation and death following a *SARS-CoV-2* infection despite receiving a third dose booster, even when adjusting for salient factors such as age. Thus, while behavioural changes and reducing social contacts may provide additional protection from infection, it is unable to compensate for the higher risk of adverse outcomes after an infection has occurred.

Lower primary and booster vaccine effectiveness as well as more rapid waning of protection may be linked to differences B-cell and T-cell immune responses following vaccination. For context, *SARS-CoV-2* vaccination in a healthy individual enables the formation of durable B cell-mediated immunity characterised by stable memory B cell populations which respond rapidly upon antigen exposure via natural infection or booster vaccination ^{500–502}.

These responses are characterised by production and release of neutralising antibodies specific to *SARS-CoV-2* upon antigen exposure which in turn compensates for waning antibody levels in the period following initial antigen exposure ^{500–502}. A protective clinical effect is also seen in T cell responses in a short time frame following two-dose vaccination with a level of protection comparable to natural infection ⁵⁰³. This includes the observation of strong CD8+ T cell responses observed after initial dosing and subsequent second dose (or booster) vaccination has been shown to induce *SARS-CoV-2* spike protein-specific CD4+ T cell responses which last several months ⁵⁰³. Within the cancer population however, my findings correspond to those of previous serology and T-cell response studies showing that these B and T cell responses to *SARS-CoV-2* vaccination are considerably reduced within immunocompromised patient cohorts ^{437,439,504–508}.

The underlying reason for poorer B and T cell responses in these immunocompromised individuals, in particular those with a leukaemia or lymphoma diagnosis, may have a limited intrinsic capacity to maintain immunity acquired post-vaccination (i.e., immunological memory) leading to waning effectiveness but that receipt of a third dose booster can trigger and re-activate antibody responses in some of these patients ^{437,439,504–508}. Large and rapid increases in antibody responses (titres) following *SARS-CoV-2* vaccination have also been observed in individuals, including clinically vulnerable subgroups, who had already received at least one prior vaccine dose, indicating good memory B cell responses following prior exposure ^{509–511}. However, it is important to note that immunocompromised subgroups are still reported to have lower post-vaccine antibody titres than the general population, with this corresponding to higher risk of adverse outcomes following an infection ^{504–508}. This in combination highlights the potential utility of immunological B-cell antibody and/or T-cell testing assays for evaluating the level of protection provided by *SARS-CoV-2* vaccination and mitigating risk within immunocompromised population subgroups including those with a cancer diagnosis.

Through assessing post-vaccine *SARS-CoV-2* antibody responses from the National COVID Cancer Antibody Survey, I show that anti-COVID spike (COVS) antibody testing can be effective for identifying patients with cancer who, following vaccination, have lower levels of immunological protection against infection. Weaker antibody responses and lower post-vaccine antibody titres correspond with a higher risk of *SARS-CoV-2* infection and hospitalisation, particularly for individuals with titres lower than 5000 U/ml, and that the benefit of vaccination on an individual level is not homogenous. Individuals with a diagnosis of leukaemia or lymphoma also typically have lower *SARS-CoV-2* antibody titres than those with other tumour types following vaccination which is in keeping with the lower levels of protection observed in these individuals following two or three vaccination doses. Taken together, these findings suggests that increased access to antibody testing for patients with cancer should be considered. This could be beneficial for supporting decision-making for national clinical guidelines, including future vaccination booster programmes, and as a diagnostic tool to better identify individuals who would derive the

greatest benefit from receiving early treatment following SARS-CoV-2 exposure and/or pre-exposure prophylactic monoclonal antibodies such as tixagevimab/cilgavimab and sotrovimab.

Lastly, I evaluated the COVID-19 disease phenotype over the course of the SARS-CoV-2 pandemic in patients with cancer within the UK, observing reduced frequency of severe outcomes (hospital assessment, inpatient hospitalisation, intensive care admission and mortality) following a SARS-CoV-2 infection. This change in disease phenotype is also reflected in the wider population and has, in part, likely been driven by the successful rollout of SARS-CoV-2 vaccines by UK public health bodies and NHS to clinically vulnerable individuals including those with a cancer diagnosis. The development and deployment of new therapeutic agents such as antiviral and monoclonal antibody therapies have also contributed to ameliorating patient outcomes ^{301,512–514}. However, I also show that the risk of severe COVID-19 related outcomes remains significantly higher in patients with cancer compared to the non-cancer population and thus maintaining access to protective measures for mitigating risk remains salient. In particular, individuals receiving SACT are at higher risk of adverse clinical outcomes with patients in receipt of B and T cell-depleting therapies appearing to be at greatest risk when comparing SACT treatment groups. This contradicts earlier studies performed in 2020 assessing COVID-19 outcomes ^{284,456} but is in keeping with my findings of reduced primary and booster vaccine effectiveness within this subgroup of patients with cancer. Taken together, the clinical benefit of anticancer treatment needs to be weighed up against the possibility of lower post-vaccine protection and continuing higher risk of adverse COVID-19 outcomes.

There are some potential limitations of my analyses. Firstly, my population-scale evaluations of vaccine effectiveness and temporal changes in disease phenotype only included patients with a diagnosis of cancer recorded before 30th April 2021 with those with more recent diagnoses not included within my cancer cohorts. This resulted from delays of newer cancer diagnoses being incorporated into the Public Health England rapid registration national cancer dataset. My assessments of primary and third dose booster vaccine effectiveness for patients with cancer may therefore underestimate the difference in protection provided by vaccination, as individuals with a more recent diagnosis were shown to have lower levels of vaccine effectiveness and more likely to be in receipt of anticancer treatment (e.g., SACT). Furthermore, my implementation of a test-negative case-control (TNCC) design for vaccine effectiveness was based on SARS-CoV-2 PCR testing data, with the mean number of tests per individual higher within the cancer cohort compared to the non-cancer population (e.g., 4.54 v 2.74 tests per person in my 2-dose vaccine effectiveness dataset). This is likely driven by clinical guidelines whereby patients with cancer underwent more PCR testing in order to access primary and secondary care settings (e.g., for anticancer treatment ^{295,515}) which in turn may alter the denominator used for TNCC analyses (negative tests) and therefore calculated vaccine effectiveness. Subsequently, multiple studies comprising clinically vulnerable individuals have assessed vaccine effectiveness using the TNCC method. This includes a retrospective

cancer study in Ontario (Canada), temporal evaluation of efficacy during the Delta and Omicron variant waves in Portugal and assessment of effectiveness against Omicron subvariants in Quebec (Canada) ^{493–495}.

Specifically for my assessment of SARS-CoV-2 post-vaccination antibody responses comprising individuals with a cancer diagnosis, the dataset captured a range of different time intervals between last vaccination and antibody testing. This may be reflected in the antibody responses recorded, with the timing of SARS-CoV-2 vaccination relative to cancer progression and the immunosuppressive effects of some anticancer treatments likely to be an important determinant. The overall size of my cancer cohort in this evaluation is also considerably smaller than for my other analyses, making it more challenging to draw firm conclusions from assessing cancer subgroups including antibody responses by primary tumour subtype. However, the cross-sectional study design is easier to deliver within a primary or secondary care setting and, more pertinently, more representative of the heterogeneity of SARS-CoV-2 vaccination schedules that exists within patients with cancer and the wider population.

For all my analyses, some of the data fields within the combined UKCCP dataset were not always complete. This includes information relating to some patient demographics such as deprivation index and more granular information about any anticancer treatments received within my cancer cohorts. In spite of this, I included some of these data field within my analyses due to their known clinical relevance ^{516–519} and the size of national-level population-scale datasets which still ensured a high level of granularity and power when assessing patient outcomes. This balanced and pragmatic approach enabled more rapid dissemination of my findings to facilitate informed decision-making by clinicians and people living with cancer. Notably, lower rates of severe COVID-19 outcomes following SARS-CoV-2 vaccination, the association of SARS-CoV-2 infection and hospital admission with post-vaccine antibody responses and increased risk faced by individuals with haematological malignancies have subsequently been shown within other cancer cohorts across multiple healthcare systems ^{493,520–522}. This is in addition to recent studies implementing test-negative case-control methodology to assess vaccine effectiveness in clinically vulnerable population subgroups ^{493–495}. It is also important to note that I defined clinical outcomes in the context of time between confirmed SARS-CoV-2 infection and event occurrence rather than being directly attributable to COVID-19 (e.g., COVID-19 being listed as cause of death on a death certificate). Clinical guidelines for hospital and intensive care admissions were also adapted during the COVID-19 pandemic for managing healthcare capacity and infection control which may have impacted admissions of patients with cancer more profoundly than the non-cancer population. However, I implemented these definitions as they were the best possible surrogate measures for assessing clinical outcomes with the available data and, more importantly, were in line with definitions used by UK public health bodies as well as other COVID-19 studies within UK cohorts. Further to these data-associated limitations, additional checks could have been performed to confirm the suitability of univariable and multivariable regressions for my evaluating vaccine effectiveness and patient outcomes. This could have included ascertaining the normality of the distributions of residuals within regression models and

whether any multicollinearity existed between covariates for multivariable analyses. The first of these, residual distributions, could have been assessed through plotting Q-Q (quantile-quantile) plots for each logistic regression which would have enabled evaluation of the conformity between the empirical (actual) distribution and the given theoretical (normal) distribution ⁵²³. Multicollinearity on the other hand refers to a strong linear relationship between two or more independent (or predictor) variables in a regression model in which they are highly correlated with each other whereby this could have been identified by performing a principal component analysis (PCA) or checking a specific measure of multicollinearity such as the variance inflation factor (VIF) between covariates ^{524,525}. Implementing these checks would therefore demonstrate suitability and robustness of these regression models for explaining the relationship between predictor (independent) and outcome variables in my analyses.

In summary, I have successfully evaluated the impact of COVID-19 in patients with cancer in the United Kingdom using population-level data. I have shown that many individuals living with cancer derive considerable benefit from SARS-CoV-2 vaccination with elevated levels of protection against severe COVID-19 outcomes provided by primary and booster vaccine doses. The clinical effectiveness of vaccines effectiveness is typically lower than that observed within the non-cancer population and typically wanes more quickly, thereby highlighting the importance of access to regular vaccine boosters to these clinically vulnerable population groups. Vaccine effectiveness is also heterogeneous, with individuals with a haematological malignancy and/or those receiving systemic anticancer therapies (SACT) associated with the lowest levels of protection against SARS-CoV-2 infection and adverse outcomes. This approach to determine vaccine protection, using SARS-CoV-2 diagnostic testing data and implementing a test-negative case-control study design, has subsequently been utilised within other cohorts comprising patients with cancer ^{493–495}. However, my vaccine effectiveness evaluations cover the period preceding the emergence and circulation of Omicron variants in the wider population, thus future studies are important to determine the level of protection provided by vaccines in these clinically vulnerable individuals considering emerging SARS-CoV-2 variants.

My analyses of vaccine effectiveness correspond with those from my assessment of antibody responses in patients with cancer following two or three vaccine doses, demonstrating that anti-spike SARS-CoV-2 (COVS) antibody testing can be implemented to identify individuals with the lowest level of antibody-derived protection and immunity from SARS-CoV-2 infection and severe outcomes. Expansion of COVS antibody testing could therefore be implemented for identifying individuals with a cancer diagnosis who could derive the most benefit from pre-emptive treatment with prophylactic monoclonal antibodies and/or earlier intervention upon contracting a SARS-CoV-2 infection. Lastly, through evaluating the impact of COVID-19 in patients with cancer over time, I show that the risk of severe COVID-19 events is at its lowest since the start of the pandemic in 2020. The COVID-19 disease phenotype is now markedly less severe for most patients with cancer and is likely the result of infection control measures, therapeutic interventions (e.g., vaccines, antivirals, monoclonal antibodies). However, patients with

cancer remain at elevated risk of severe COVID-19 outcomes compared to the general (non-cancer) population, whereby individuals with a haematological cancer and/or receiving SACT correspond with the highest risk. Continued monitoring and evaluation of the impact of COVID-19 is therefore required to accurately individualise risk for as COVID-19 has become an endemic disease and new SARS-CoV-2 variants (e.g., Pirola/BA.2.86⁵²⁶) emerge and circulate within the wider population in order to safeguard people living with cancer.

ACKNOWLEDGEMENTS AND CONTRIBUTION DISCLOSURES

Chapter 1: Overview of methods for characterising anti-tumour immune responses in cancer

There are no acknowledgements or contributions to disclose for this chapter.

Chapter 2: Immunogenomic analysis of colorectal cancer primary tumours and association with tumour features

I would like to acknowledge Dr Roland Arnold and Dr Anshita Goel (both University of Birmingham) for bioinformatic support and advice which enabled me to develop coding scripts to process and analyse transcriptomic (RNA sequencing) data as part of this chapter. This included advice for developing scripts that enabled me to perform data processing and analysis within the Birmingham Environment for Academic Research (BEAR) computer cluster as well as creating my analysis stream for quantifying the stages of the Cancer Immunity Cycle. Dr Takashi Akiyoshi (Department of Gastroenterological Surgery, Cancer Institute Hospital, Tokyo) and Dr Seichi Mori (Cancer Precision Medicine Centre, Tokyo) kindly provided approval for me to utilise data originally published by Akiyoshi *et al.* in 2019 (reference ²⁰⁹) for bulk transcriptomic analyses of rectal cancer biopsies prior to neoadjuvant chemoradiotherapy (CRT) cohort. I also received feedback for improving my neoadjuvant CRT analyses from Dr João D. Barros-Silva (Francis Crick Institute, formerly University of Birmingham).

Chapter 3: Immunogenomic analysis of local and distant metastases of colorectal cancer

I would like to acknowledge Professor Gary Middleton, Dr Neeraj Lal, and Dr Ghaleb Goussous (all University of Birmingham) for support and advice provided while I developed my transcriptomic analyses of secondary metastases of colorectal cancer (liver and lung). Dr Lennard Lee (University of Oxford and University of Birmingham) prepared samples for the METALLIC cohort (local lymph node metastases) prior to 3' RNA sequencing. Dr Ghaleb Goussous and Dr Lennard Lee provided FFPE specimens for the distant metastasis cohort, comprising colorectal primary tumours and secondary liver/lung metastases, from which I extracted RNA and prepared cDNA libraries for 3' RNA sequencing. Members of the Genomics Birmingham sequencing facility, namely Dr Celina Whalley, Ms Charlie Poxon (both Nonacus Limited, formerly University of Birmingham) and Mr Kostas Kagiopoulos (University of Manchester, formerly University of Birmingham) were involved in performing 3' RNA sequencing within the Genomics Birmingham sequencing facility as well as providing expertise for RNA extraction and cDNA library preparation. Dr Roland Arnold and Dr Anshita Goel (both University of Birmingham) provided bioinformatic support and advice which enabled me to develop coding scripts to process transcriptomic (RNA sequencing) data as part of this chapter. Data from the METALLIC cohort has previously been published by

Demetriou *et al.* in 2020 (reference ⁵²⁷) for which I provided transcriptomic analysis as part of a wider publication but for which only the relevant overlapping methods sections are included within the body of work of my thesis.

Chapter 4: Overview of the characteristics of SARS-CoV-2 and impact of COVID-19 at the population level

There are no acknowledgements or contributions to disclose for this chapter.

Chapter 5: Monitoring the impact of Coronavirus Disease 2019 (COVID-19) in patients with cancer at the start of the pandemic.

I would like to acknowledge the United Kingdom Coronavirus Cancer Monitoring Project (UKCCMP) team for their support initiating and delivering the monitoring project. This support enabled me to perform granular analyses of COVID-19 disease phenotype and assess clinical outcomes of patients with cancer during the initial stages of the COVID-19 pandemic in 2020. In particular, I acknowledge the contributions of Dr Lennard Lee (University of Oxford and University of Birmingham), Professor Jean-Baptiste Cazier (University of Birmingham), Professor Rachel Kerr (University of Oxford) and Professor Gary Middleton (University of Birmingham) who co-led the UKCCMP programme. Data acquisition via RedCAP prior to my analyses was led by Dr Claire Palles, Dr Naomi Campton (both University of Birmingham) and Dr Helen Curley (Queen Elizabeth Hospitals Birmingham) with data processing coordinated by the Centre for Computational Biology (CCB), Dr Roland Arnold, Dr Anshita Goel and Dr Csilla Varnai (all University of Birmingham). I received experimental design feedback from the UKCCMP programme co-leads as well as Dr Chris Turnbull, Dr Sarah Briggs (both University of Oxford) and Dr Adam Briggs (University of Warwick). Institution-level approval for initiating the UKCCMP provided by the University of Birmingham College of Medical and Dental Sciences (MDS), then chaired by Professor David Adams, enabled the provision of granular data for my analyses. Parts of this chapter, namely the analyses I performed detailed in the methods and results sections, have been incorporated as part of the publication listed below which I coauthored and has been included in the reference list (⁵²⁸).

- Lee LYW, Cazier JB, Starkey T, *et al.* COVID-19 prevalence and mortality in patients with cancer and the effect of primary tumour subtype and patient demographics: a prospective cohort study. *Lancet Oncol.* **21**(10):1309–25 (2020).

Chapter 6: An evaluation of how the impact of coronavirus disease 2019 (covid-19) has changed in patients with cancer across the pandemic.

I would like to acknowledge the UKCCP team (<https://ukcovidcancerprogramme.org/>) for their support initiating the programme which allowed me to evaluate the impact of COVID-19 disease in patients with cancer on a population-scale across the COVID-19 pandemic. In particular, I acknowledge the contributions of Dr Lennard Lee (University of Oxford and University of Birmingham) and Professor Peter Johnson (University of Southampton and national cancer lead) who co-initiated the UKCCP. Data acquisition was performed within the UK Health Security Agency (UKHSA) secure research environment and was linked by Ms Maria Ionescu, with data cleaning and processing of linked data being performed jointly by Ms Ionescu and myself. This process was facilitated by the UKHSA and in particular Dr Tom Fowler (UKHSA). I received feedback from the project co-leads as well as Maria Ionescu, Dr Michael Tilby (University Hospitals Birmingham) and Dr Matthew Fittall (University College London), with Dr Fittall specifically providing support for classifying systemic anticancer therapies (SACT) which I assessed in my temporal analyses. Patient recruitment for the National COVID-19 Cancer antibody survey, which formed a key part of the dataset I used for assessing antibody response to COVID-19 vaccine evaluation, was led by Dr Tilby and funding was received from Blood Cancer UK to enable study completion. Parts of this chapter, namely the analyses I performed detailed in the methods and results sections, have been published within multiple peer-reviewed publications which I coauthored (references ^{529–532}). The list of publications is as below:

- Lee LYW, Starkey T, Ionescu MC, *et al.* Vaccine effectiveness against COVID-19 breakthrough infections in patients with cancer (UKCCEP): a population-based test-negative case-control study. *Lancet Oncol.* **23**(6):748–57 (2022).
- Lee LYW, Ionescu MC, Starkey T, *et al.* COVID-19: Third dose booster vaccine effectiveness against breakthrough coronavirus infection, hospitalisations and death in patients with cancer: A population-based study. *Eur J Cancer.* **175**:1–10 (2022).
- Lee LYW, Tilby M, Starkey T, *et al.* Association of SARS-CoV-2 Spike Protein Antibody Vaccine Response With Infection Severity in Patients With Cancer: A National COVID Cancer Cross-sectional Evaluation. *JAMA Oncol.* **9**(2):188-196 (2023).
- Starkey T, Ionescu MC, Tilby M, *et al.* A population-scale temporal case–control evaluation of COVID-19 disease phenotype and related outcome rates in patients with cancer in England (UKCCP). *Sci Rep.* **13**(11327) (2023).

REFERENCES

1. Bowel cancer statistics | Cancer Research UK. Available at: <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer#heading-Two>. (Accessed: 24th September 2020)
2. Duchalais, E. *et al.* Colorectal Cancer Cells Adhere to and Migrate Along the Neurons of the Enteric Nervous System. *CMGH* **5**, 31–49 (2018).
3. Pang, R. *et al.* Cell Stem Cell Article A Subpopulation of CD26 + Cancer Stem Cells with Metastatic Capacity in Human Colorectal Cancer. (2010). doi:10.1016/j.stem.2010.04.001
4. Angelova, M. *et al.* Evolution of Metastases in Space and Time under Immune Selection. *Cell* **175**, 751–765.e16 (2018).
5. De Visser, K. E., Eichten, A. & Coussens, L. M. Paradoxical roles of the immune system during cancer development. *Nature Reviews Cancer* **6**, 24–37 (2006).
6. Dranoff, G. Cytokines in cancer pathogenesis and cancer therapy. *Nature Reviews Cancer* **4**, 11–22 (2004).
7. Farhood, B., Najafi, M. & Mortezaee, K. CD8 + cytotoxic T lymphocytes in cancer immunotherapy: A review. *J. Cell. Physiol.* **234**, 8509–8521 (2019).
8. Nicholson, S. E., Keating, N. & Belz, G. T. Natural killer cells and anti-tumor immunity. *Mol. Immunol.* (2017). doi:10.1016/J.MOLIMM.2017.12.002
9. Nair, S. & Dhodapkar, M. V. Natural Killer T Cells in Cancer Immunotherapy. *Front. Immunol.* **8**, 1178 (2017).
10. Zhao, Y., Niu, C. & Cui, J. Gamma-delta ($\gamma\delta$) T cells: friend or foe in cancer development? *J. Transl. Med.* **16**, 3 (2018).
11. Groscurth, P. Cytotoxic effector cells of the immune system. *Anatomy and Embryology* **180**, 109–119 (1989).
12. Charles A Janeway, J., Travers, P., Walport, M. & Shlomchik, M. J. T cell-mediated cytotoxicity. (2001).
13. Finn, O. J. Immuno-oncology: understanding the function and dysfunction of the immune system in cancer. *Ann. Oncol.* **23**, viii6–viii9 (2012).
14. Wellenstein, M. D. & De Visser, K. E. Review Cancer-Cell-Intrinsic Mechanisms Shaping the Tumor Immune Landscape. *Immunity* **48**, 399–416 (2018).
15. Martelotto, L. G., Ng, C. K. Y., Piscuoglio, S., Weigelt, B. & Reis-Filho, J. S. Breast cancer intra-tumor heterogeneity. *Breast Cancer Res.* **16**, 1–11 (2014).
16. Liu, J., Dang, H. & Wang, X. W. The significance of intertumor and intratumor heterogeneity in liver cancer. *Exp. Mol. Med.* **2018 501** **50**, e416–e416 (2018).
17. McGranahan, N. & Swanton, C. Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. *Cell* **168**, 613–628 (2017).
18. Marusyk, A., Janiszewska, M. & Polyak, K. Intratumor heterogeneity: the Rosetta stone of therapy resistance. *Cancer Cell* **37**, 471 (2020).
19. Ramón y Cajal, S. *et al.* Clinical implications of intratumor heterogeneity: challenges and opportunities. *J. Mol. Med. (Berl)*. **98**, 161 (2020).
20. Mroz, E. A. & Rocco, J. W. MATH, a novel measure of intratumor genetic heterogeneity, is high in poor-outcome classes of head and neck squamous cell carcinoma. *Oral Oncol.* **49**, 211–215 (2013).
21. Zhang, J. *et al.* Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing. *Science (80-.)*. **346**, 256–259 (2014).
22. Patel, A. P. *et al.* Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science (80-.)*. **344**, 1396–1401 (2014).
23. Zou, Y. *et al.* The Single-Cell Landscape of Intratumoral Heterogeneity and The Immunosuppressive Microenvironment in Liver and Brain Metastases of Breast Cancer. *Adv. Sci.* **10**, 2203699 (2023).

24. Ling, A. *et al.* The infiltration, and prognostic importance, of Th1 lymphocytes vary in molecular subgroups of colorectal cancer. *J. Pathol. Clin. Res.* **2**, 21–31 (2016).
25. Glaire, M. A. *et al.* Tumour-infiltrating CD8+ lymphocytes and colorectal cancer recurrence by tumour and nodal stage. *Br. J. Cancer* **121**, 474–482 (2019).
26. Coppola, A. *et al.* NK cell inflammation in the clinical outcome of colorectal carcinoma. *Frontiers in Medicine* **2**, (2015).
27. Ottaiano, A. *et al.* Natural killer cells activity in a metastatic colorectal cancer patient with complete and long lasting response to therapy. *World J. Clin. Cases* **5**, 390 (2017).
28. Zhang, B. *et al.* Circulating and Tumor-Infiltrating Myeloid-Derived Suppressor Cells in Patients with Colorectal Carcinoma. *PLoS One* **8**, 57114 (2013).
29. Ai, L. *et al.* Prognostic role of myeloid-derived suppressor cells in cancers: A systematic review and meta-analysis 11 Medical and Health Sciences 1112 Oncology and Carcinogenesis. *BMC Cancer* **18**, 1220 (2018).
30. Yahaya, M. A. F., Lila, M. A. M., Ismail, S., Zainol, M. & Afizan, N. A. R. N. M. Tumour-Associated Macrophages (TAMs) in colon cancer and how to reeducate them. *Journal of Immunology Research* **2019**, (2019).
31. Pardoll, D. M. The blockade of immune checkpoints in cancer immunotherapy. *Nat. Rev. Cancer* **12**, 252–64 (2012).
32. Passardi, A., Canale, M., Valgiusti, M. & Ulivi, P. Immune Checkpoints as a Target for Colorectal Cancer Treatment. *Int. J. Mol. Sci.* **18**, (2017).
33. Yaghoubi, N., Soltani, A., Ghazvini, K., Hassanian, S. M. & Hashemy, S. I. PD-1/ PD-L1 blockade as a novel treatment for colorectal cancer. *Biomedicine and Pharmacotherapy* **110**, 312–318 (2019).
34. Andre, T. *et al.* Pembrolizumab versus chemotherapy for microsatellite instability-high/mismatch repair deficient metastatic colorectal cancer: The phase 3 KEYNOTE-177 Study. *J. Clin. Oncol.* **38**, LBA4-LBA4 (2020).
35. Kalyan, A., Kircher, S., Shah, H., Mulcahy, M. & Benson, A. Updates on immunotherapy for colorectal cancer. *J. Gastrointest. Oncol.* **9**, 160–169 (2018).
36. Le, D. T. *et al.* PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N. Engl. J. Med.* **372**, 2509–2520 (2015).
37. Lenz, H.-J. *et al.* Nivolumab (NIVO) + low-dose ipilimumab (IPI) as first-line (1L) therapy in microsatellite instability-high/mismatch repair-deficient (MSI-H/dMMR) metastatic colorectal cancer (mCRC): Two-year clinical update. *J. Clin. Oncol.* **38**, 4040–4040 (2020).
38. Berntsson, J. *et al.* The clinical impact of tumour-infiltrating lymphocytes in colorectal cancer differs by anatomical subsite: A cohort study. doi:10.1002/ijc.30869
39. Edin, S. *et al.* The Prognostic Importance of CD20+ B lymphocytes in Colorectal Cancer and the Relation to Other Immune Cell subsets. *Sci. Rep.* **9**, 1–9 (2019).
40. Engstrand, J., Nilsson, H., Strömberg, C., Jonas, E. & Freedman, J. Colorectal cancer liver metastases - a population-based study on incidence, management and survival. *BMC Cancer* **18**, (2018).
41. Zhang, L. *et al.* Immune landscape of colorectal cancer tumor microenvironment from different primary tumor location. *Front. Immunol.* **9**, (2018).
42. Galon, J. & Bruni, D. Approaches to treat immune hot, altered and cold tumours with combination immunotherapies. *Nat. Rev. Drug Discov.* **18**, 197–218 (2019).
43. Ulivi, P. *et al.* Right- versus left-side metastatic colorectal cancer: Differences in tumor biology and bevacizumab efficacy. *Ann. Oncol.* **28**, iii8 (2017).
44. Baran, B. *et al.* Difference Between Left-Sided and Right-Sided Colorectal Cancer: A Focused Review of Literature. *Gastroenterol. Res.* **11**, 264 (2018).
45. Guinney, J. *et al.* The consensus molecular subtypes of colorectal cancer. *Nat. Med.* (2015). doi:10.1038/nm.3967
46. Isella, C. *et al.* Selective analysis of cancer-cell intrinsic transcriptional traits defines novel clinically relevant subtypes of colorectal cancer. *Nat. Commun.* **8**, 1–16 (2017).

47. Thorsson, V. *et al.* The Immune Landscape of Cancer. *Immunity* **48**, 812–830.e14 (2018).
48. Fukuoka, S. *et al.* Regorafenib Plus Nivolumab in Patients With Advanced Gastric or Colorectal Cancer: An Open-Label, Dose-Escalation, and Dose-Expansion Phase Ib Trial (REGONIVO, EPOC1603). *J. Clin. Oncol.* **38**, 2053–2061 (2020).
49. García-Mulero, S. *et al.* Lung metastases share common immune features regardless of primary tumor origin. *J. Immunother. cancer* **8**, 491 (2020).
50. Kim, S. W., Roh, J. & Park, C. S. Immunohistochemistry for pathologists: Protocols, pitfalls, and tips. *Journal of Pathology and Translational Medicine* **50**, 411–418 (2016).
51. Lindor, N. M. *et al.* Immunohistochemistry versus microsatellite instability testing in phenotyping colorectal tumors. *J. Clin. Oncol.* **20**, 1043–1048 (2002).
52. Shia, J. Immunohistochemistry versus microsatellite instability testing for screening colorectal cancer patients at risk for hereditary nonpolyposis colorectal cancer syndrome: Part I. The utility of immunohistochemistry. *J. Mol. Diagnostics* **10**, 293–300 (2008).
53. Yuan, L. *et al.* Immunohistochemistry and microsatellite instability analysis in molecular subtyping of colorectal carcinoma based on mismatch repair competency. *Int. J. Clin. Exp. Med.* **8**, 20988–21000 (2015).
54. Rosenbaum, M. W., Bledsoe, J. R., Morales-Oyarvide, V., Huynh, T. G. & Mino-Kenudson, M. PD-L1 expression in colorectal cancer is associated with microsatellite instability, BRAF mutation, medullary morphology and cytotoxic tumor-infiltrating lymphocytes. in *Modern Pathology* **29**, 1104–1112 (Nature Publishing Group, 2016).
55. Loh, Z., Williams, D. S., Salmon, L., Dow, E. & John, T. Impact of universal immunohistochemistry on Lynch syndrome diagnosis in an Australian colorectal cancer cohort. (2019). doi:10.1111/imj.14230
56. Dvorak, K., Higgins, A., Palting, J., Cohen, M. & Brunhoeber, P. Immunohistochemistry with Anti-BRAF V600E (VE1) Mouse Monoclonal Antibody is a Sensitive Method for Detection of the BRAF V600E Mutation in Colon Cancer: Evaluation of 120 Cases with and without KRAS Mutation and Literature Review. (2017). doi:10.1007/s12253-017-0344-x
57. Oyer, J. L., Gitto, S. B., Altomare, D. A. & Copik, A. J. PD-L1 blockade enhances anti-tumor efficacy of NK cells. *Oncoimmunology* **7**, e1509819 (2018).
58. Alsaab, H. O. *et al.* PD-1 and PD-L1 Checkpoint Signaling Inhibition for Cancer Immunotherapy: Mechanism, Combinations, and Clinical Outcome. *Front. Pharmacol.* **8**, 561 (2017).
59. Tiernan, J. P. *et al.* Carcinoembryonic antigen is the preferred biomarker for in vivo colorectal cancer targeting. *Br. J. Cancer* **108**, 662–667 (2013).
60. Saito, G. *et al.* Relation between Carcinoembryonic Antigen Levels in Colon Cancer Tissue and Serum Carcinoembryonic Antigen Levels at Initial Surgery and Recurrence. *Oncology* **91**, 85–89 (2016).
61. Gray-Owen, S. D. & Blumberg, R. S. CEACAM1: Contact-dependent control of immunity. *Nature Reviews Immunology* **6**, 433–446 (2006).
62. Galon, J. *et al.* Type, Density, and Location of Immune Cells Within Human Colorectal Tumors Predict Clinical Outcome. *Science (80-.)*. **313**, 1960–1964 (2006).
63. Pagès, F. *et al.* International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study. *Lancet (London, England)* **391**, 2128–2139 (2018).
64. Bindea, G., Mlecnik, B., Angell, H. K. & Galon, J. The immune landscape of human tumors: Implications for cancer immunotherapy. *Oncoimmunology* **3**, e27456 (2014).
65. Mlecnik, B. *et al.* Integrative Analyses of Colorectal Cancer Show Immunoscore Is a Stronger Predictor of Patient Survival Than Microsatellite Instability. *Immunity* **44**, 698–711 (2016).
66. Mlecnik, B. *et al.* Comprehensive Intrametastatic Immune Quantification and Major Impact of Immunoscore on Survival. (2017). doi:10.1093/jnci/djx123
67. Mlecnik, B. *et al.* The tumor microenvironment and Immunoscore are critical determinants of dissemination to distant metastasis. *Sci. Transl. Med.* **8**, 327ra26–327ra26 (2016).
68. Van den Eynde, M. *et al.* The Link between the Multiverse of Immune Microenvironments in Metastases and the Survival of Colorectal Cancer Patients. *Cancer Cell* **34**, 1012–1026.e3 (2018).

69. Angelova, M. *et al.* Evolution of Metastases in Space and Time under Immune Selection. *Cell* **175**, 751–765.e16 (2018).
70. Galon, J. *et al.* Towards the introduction of the 'Immunoscore' in the classification of malignant tumours. *J. Pathol.* **232**, 199 (2014).
71. McKinnon, K. M. Flow cytometry: An overview. *Curr. Protoc. Immunol.* **2018**, 5.1.1-5.1.11 (2018).
72. Spitzer, M. H. & Nolan, G. P. Mass Cytometry: Single Cells, Many Features. *Cell* **165**, 780–791 (2016).
73. Shimabukuro-Vornhagen, A. *et al.* Characterization of tumor-associated B-cell subsets in patients with colorectal cancer. *Oncotarget* **5**, 4651–4664 (2014).
74. Zhang, T. *et al.* Immunocyte Profiling Using Single-Cell Mass Cytometry Reveals EpCAM+ CD4+ T Cells Abnormal in Colon Cancer. *Front. Immunol.* **10**, 1571 (2019).
75. Shinko, D. *et al.* Mass Cytometry Reveals a Sustained Reduction in CD16+ Natural Killer Cells Following Chemotherapy in Colorectal Cancer Patients. *Front. Immunol.* **10**, 2584 (2019).
76. Pasero, C. *et al.* Highly effective NK cells are associated with good prognosis in patients with metastatic prostate cancer. *Oncotarget* **6**, 14360–73 (2015).
77. Michea, P. *et al.* Adjustment of dendritic cells to the breast-cancer microenvironment is subset specific. *Nat. Immunol.* **19**, (2018).
78. Dalerba, P. *et al.* Phenotypic characterization of human colorectal cancer stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 10158–10163 (2007).
79. De Vries, N. L. *et al.* High-dimensional cytometric analysis of colorectal cancer reveals novel mediators of antitumour immunity. *Gut* **69**, 691–703 (2020).
80. Xu, B. *et al.* Prognostic value of tumor infiltrating NK cells and macrophages in stage II+III esophageal cancer patients. *Oncotarget* **7**, 74904–74916 (2016).
81. Bustin, S. A. & Nolan, T. Pitfalls of quantitative real- time reverse-transcription polymerase chain reaction. *Journal of Biomolecular Techniques* **15**, 155–166 (2004).
82. Chang, D. H. *et al.* The Effect of Lung Cancer on Cytokine Expression in Peripheral Blood Mononuclear Cells. *PLoS One* **8**, e64456 (2013).
83. Huang, H. *et al.* Multiplexed immunoassay-based serum cytokine profiling for potential prognosis predictors in patients with metastatic breast cancer. *Transl. Cancer Res.* **7**, 1561–1566 (2018).
84. Gorelik, E. *et al.* Multiplexed immunobead-based cytokine profiling for early detection of ovarian cancer. *Cancer Epidemiol. Biomarkers Prev.* **14**, 981–987 (2005).
85. Mackay, S. *et al.* Single-cell cytokine profiling of tumor-infiltrating T cells to measure patient responses to anti-PD-1 therapy. *J. Clin. Oncol.* **35**, 49–49 (2017).
86. Koboldt, D. C., Steinberg, K. M., Larson, D. E., Wilson, R. K. & Mardis, E. R. XThe next-generation sequencing revolution and its impact on genomics. *Cell* **155**, 27 (2013).
87. Kulski, J. K. in *Next Generation Sequencing - Advances, Applications and Challenges* (InTech, 2016). doi:10.5772/61964
88. Tomczak, K., Czerwińska, P. & Wiznerowicz, M. The Cancer Genome Atlas (TCGA): An immeasurable source of knowledge. *Współczesna Onkologia* **1A**, A68–A77 (2015).
89. Thorisson, G. A., Smith, A. V., Krishnan, L. & Stein, L. D. The International HapMap Project Web site. *Genome Res.* **15**, 1592–1593 (2005).
90. Whewey, G. *et al.* Opportunities and challenges for molecular understanding of ciliopathies—the 100,000 genomes project. *Frontiers in Genetics* **10**, 127 (2019).
91. The 100,000 Genomes Project | Genomics England. Available at: <https://www.genomicsengland.co.uk/about-genomics-england/the-100000-genomes-project/>. (Accessed: 29th September 2020)
92. Secretary of State for Health and Social Care announces ambition to sequence 5 million genomes within five years | Genomics England. Available at: <https://www.genomicsengland.co.uk/matt-hancock-announces-5-million-genomes-within-five-years/>. (Accessed: 29th September 2020)

93. Illumina. *Technology Spotlight: Illumina® Sequencing*.
94. Wendl, M. C. & Wilson, R. K. Aspects of coverage in medical DNA sequencing. *BMC Bioinformatics* **9**, 239 (2008).
95. Maby, P. *et al.* Correlation between density of CD8 + T-cell infiltrate in microsatellite unstable colorectal cancers and frameshift mutations: A rationale for personalized immunotherapy. *Cancer Res.* **75**, 3446–3455 (2015).
96. Ilieva, K. M. *et al.* Effects of BRAF mutations and BRAF inhibition on immune responses to melanoma. *Molecular Cancer Therapeutics* **13**, 2769–2783 (2014).
97. Canon, J. *et al.* The clinical KRAS(G12C) inhibitor AMG 510 drives anti-tumour immunity. *Nature* **575**, 217–223 (2019).
98. Ioannidis, N. M. *et al.* Gene expression imputation identifies candidate genes and susceptibility loci associated with cutaneous squamous cell carcinoma. *Nat. Commun.* **9**, 4264 (2018).
99. Shahamatdar, S. *et al.* Germline Features Associated with Immune Infiltration in Solid Tumors. *Cell Rep.* **30**, 2900–2908.e4 (2020).
100. Lim, Y. W. *et al.* Germline genetic polymorphisms influence tumor gene expression and immune cell infiltration. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E11701–E11710 (2018).
101. Xin, J. *et al.* Risk assessment for colorectal cancer via polygenic risk score and lifestyle exposure: a large-scale association study of East Asian and European populations. *Genome Med.* **15**, 1–14 (2023).
102. Briggs, S. E. W. *et al.* Integrating genome-wide polygenic risk scores and non-genetic risk to predict colorectal cancer diagnosis using UK Biobank data: population based cohort study. *BMJ* **379**, (2022).
103. Sud, A. *et al.* Realistic expectations are key to realising the benefits of polygenic scores. *BMJ* **380**, (2023).
104. Our Future Health supports Genome UK Implementation Plan for England – Our Future Health. Available at: <https://ourfuturehealth.org.uk/news/our-future-health-supports-genome-uk-implementation-plan-for-england/>. (Accessed: 15th October 2023)
105. Roussel, L. *et al.* Loss of human ICO SL results in combined immunodeficiency. *J. Exp. Med.* **215**, 3151–3164 (2018).
106. Compagnone, M., Cifaldi, L. & Fruci, D. Regulation of ERAP1 and ERAP2 genes and their disfunction in human cancer. *Human Immunology* **80**, 318–324 (2019).
107. Huang, D. *et al.* Mutations of key driver genes in colorectal cancer progression and metastasis. *Cancer Metastasis Rev.* **37**, 173–187 (2018).
108. Kandoth, C. *et al.* ARTICLE Mutational landscape and significance across 12 major cancer types. (2013). doi:10.1038/nature12634
109. Caswell, D. R. & Swanton, C. The role of tumour heterogeneity and clonal cooperativity in metastasis, immune evasion and clinical outcome. doi:10.1186/s12916-017-0900-y
110. Ligresti, G. *et al.* PIK3CA mutations in human solid tumors Role in sensitivity to various therapeutic approaches. *Cell Cycle* **8**, 1352–1358 (2009).
111. Nusrat, M. *et al.* Association of PIK3CA mutations (mut) with immune engagement and clinical benefit from immunotherapy in microsatellite stable (MSS) colorectal cancer (CRC) patients (pts). *J. Clin. Oncol.* **37**, 3604–3604 (2019).
112. Yaeger, R. *et al.* Clinical Sequencing Defines the Genomic Landscape of Metastatic Colorectal Cancer. *Cancer Cell* **33**, 125–136.e3 (2018).
113. Marabelle, A. *et al.* Association of tumour mutational burden with outcomes in patients with advanced solid tumours treated with pembrolizumab: prospective biomarker analysis of the multicohort, open-label, phase 2 KEYNOTE-158 study. *Lancet Oncol.* **0**, (2020).
114. Chalmers, Z. R. *et al.* Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Med.* **9**, 34 (2017).
115. Damotte, D. *et al.* The tumor inflammation signature (TIS) is associated with anti-PD-1 treatment benefit in the CERTIM pan-cancer cohort. *J. Transl. Med.* **17**, 357 (2019).

116. Thapa, B. *et al.* The Immune Microenvironment, Genome-wide Copy Number Aberrations, and Survival in Mesothelioma. *J. Thorac. Oncol.* **12**, 850–859 (2017).
117. Gong, J. *et al.* A Pan-cancer Analysis of the Expression and Clinical Relevance of Small Nucleolar RNAs in Human Cancer. *Cell Rep.* **21**, 1968–1981 (2017).
118. Lu, Z. *et al.* Tumor copy-number alterations predict response to immune-checkpoint-blockade in gastrointestinal cancer. *J. Immunother. cancer* **8**, 374 (2020).
119. Cancer Genome Atlas Research Network. Electronic address: [REDACTED] T. C. G. A. R., Cancer Genome Atlas Research Network, D. A. & Roberts, L. R. Comprehensive and Integrative Genomic Characterization of Hepatocellular Carcinoma. *Cell* **169**, 1327–1341.e23 (2017).
120. Budinska, E. *et al.* Gene expression patterns unveil a new level of molecular heterogeneity in colorectal cancer. *J. Pathol.* **231**, 63–76 (2013).
121. Haan, J. C. *et al.* Genomic landscape of metastatic colorectal cancer. *Nat. Commun.* **5**, (2014).
122. Lee-Six, H. *et al.* The landscape of somatic mutation in normal colorectal epithelial cells. *Nature* **574**, 532–537 (2019).
123. Cai, H. *et al.* Mutational landscape of gastric cancer and clinical application of genomic profiling based on target next-generation sequencing. *J. Transl. Med.* **17**, 189 (2019).
124. Nakagawa, H. & Fujita, M. Whole genome sequencing analysis for cancer genomics and precision medicine. *Cancer Sci.* **109**, 513–522 (2018).
125. Bräunlein, E. & Krackhardt, A. M. Identification and Characterization of Neoantigens As Well As Respective Immune Responses in Cancer Patients. *Frontiers in Immunology* **8**, 1702 (2017).
126. De Mattos-Arruda, L. *et al.* Neoantigen prediction and computational perspectives towards clinical benefit: recommendations from the ESMO Precision Medicine Working Group. *Annals of Oncology* **31**, 978–990 (2020).
127. Ward, J. P., Gubin, M. M. & Schreiber, R. D. in *Advances in Immunology* **130**, 25–74 (Academic Press Inc., 2016).
128. Rooney, M. S., Shukla, S. A., Wu, C. J., Getz, G. & Hacohen, N. Molecular and Genetic Properties of Tumors Associated with Local Immune Cytolytic Activity. *Cell* **160**, 48–61 (2015).
129. Snyder, A. *et al.* Genetic Basis for Clinical Response to CTLA-4 Blockade in Melanoma. *N. Engl. J. Med.* **371**, 2189–2199 (2014).
130. Thomas, A. *et al.* Tumor mutational burden is a determinant of immune-mediated survival in breast cancer. *Oncoimmunology* **7**, (2018).
131. Zaravinos, A. *et al.* Cytolytic activity correlates with the mutational burden and deregulated expression of immune checkpoints in colorectal cancer. *J. Exp. Clin. Cancer Res.* **38**, 364 (2019).
132. Workman, H. C. *et al.* The membrane mucin MUC4 is elevated in breast tumor lymph node metastases relative to matched primary tumors and confers aggressive properties to breast cancer cells. *Breast Cancer Res.* **11**, (2009).
133. Bruyère, E., Jonckheere, N., Frénois, F., Mariette, C. & Van Seuningen, I. The MUC4 membrane-bound mucin regulates esophageal cancer cell proliferation and migration properties: Implication for S100A4 protein. *Biochem. Biophys. Res. Commun.* **413**, 325–329 (2011).
134. Schenck, R. O., Lakatos, E., Gatenbee, C., Graham, T. A. & Anderson, A. R. A. NeoPredPipe: High-throughput neoantigen prediction and recognition potential pipeline. *BMC Bioinformatics* **20**, 264 (2019).
135. Wu, J. *et al.* TSNAdb: A Database for Tumor-specific Neoantigens from Immunogenomics Data Analysis. *Genomics, Proteomics Bioinforma.* **16**, 276–282 (2018).
136. Mauger, D. M. *et al.* mRNA structure regulates protein expression through changes in functional half-life. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 24075–24083 (2019).
137. Kukurba, K. R. & Montgomery, S. B. RNA sequencing and analysis. *Cold Spring Harb. Protoc.* **2015**, 951–969 (2015).
138. Patterson, J. *et al.* Impact of sequencing depth and technology on de novo RNA-Seq assembly. *BMC Genomics* **20**, 604 (2019).

139. Martens-Uzunova, E. S. *et al.* C/D-box snoRNA-derived RNA production is associated with malignant transformation and metastatic progression in prostate cancer. *Oncotarget* **6**, 17430–17444 (2015).
140. Hong, B. S. *et al.* Tumor suppressor miRNA-204-5p regulates growth, metastasis, and immune microenvironment remodeling in breast cancer. *Cancer Res.* **79**, 1520–1534 (2019).
141. Srivastava, S. K. *et al.* MicroRNA-150 directly targets MUC4 and suppresses growth and malignant behavior of pancreatic cancer cells. *Carcinogenesis* **32**, 1832–1839 (2011).
142. Lin, Y. C. *et al.* Tumor suppressor miRNA-204-5p promotes apoptosis by targeting BCL2 in prostate cancer cells. *Asian J. Surg.* **40**, 396–406 (2017).
143. Liang, C. Y. *et al.* Downregulation of hsa-microRNA-204-5p and identification of its potential regulatory network in non-small cell lung cancer: RT-qPCR, bioinformatic- And meta-analyses. *Respir. Res.* **21**, 1–27 (2020).
144. Bindea, G. *et al.* Spatiotemporal Dynamics of Intratumoral Immune Cells Reveal the Immune Landscape in Human Cancer. *Immunity* **39**, 782–795 (2013).
145. Charoentong, P. *et al.* Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. *Cell Rep.* **18**, 248–262 (2017).
146. Angelova, M. *et al.* Characterization of the immunophenotypes and antigenomes of colorectal cancers reveals distinct tumor escape mechanisms and novel targets for immunotherapy. *Genome Biol.* **16**, 64 (2015).
147. Charoentong, P., Finotello, F., Angelova, M., Hackl, H. & Correspondence, Z. T. Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. (2017). doi:10.1016/j.celrep.2016.12.019
148. Leman, J. K. H., Sandford, S. K., Rhodes, J. L. & Kemp, R. A. Multiparametric analysis of colorectal cancer immune responses. *World Journal of Gastroenterology* **24**, 2995–3005 (2018).
149. Macosko, E. Z. *et al.* Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).
150. See, P., Lum, J., Chen, J. & Ginhoux, F. A Single-Cell Sequencing Guide for Immunologists. *Front. Immunol.* **9**, 2425 (2018).
151. Zhang, Y. *et al.* Deep single-cell RNA sequencing data of individual T cells from treatment-naïve colorectal cancer patients. *Sci. Data* **6**, (2019).
152. Li, H. *et al.* Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors. *Nat. Genet.* **49**, 708–718 (2017).
153. NHS England » National genomic test directory. Available at: <https://www.england.nhs.uk/publication/national-genomic-test-directories/>. (Accessed: 17th October 2023)
154. Zhang, L. *et al.* Lineage tracking reveals dynamic relationships of T cells in colorectal cancer. *Nature* **564**, 268–272 (2018).
155. Rosati, E. *et al.* Overview of methodologies for T-cell receptor repertoire analysis. *BMC Biotechnol.* **17**, (2017).
156. Jung, D. & Alt, F. W. Unraveling V(D)J Recombination: Insights into Gene Regulation. *Cell* **116**, 299–311 (2004).
157. Alt, F. W. *et al.* VDJ recombination. *Immunology Today* **13**, 306–314 (1992).
158. Chlewicki, L. K., Holler, P. D., Monti, B. C., Clutter, M. R. & Kranz, D. M. High-affinity, peptide-specific T cell receptors can be generated by mutations in CDR1, CDR2 or CDR3. *J. Mol. Biol.* **346**, 223–239 (2005).
159. Kjer-Nielsen, L. *et al.* A Structural Basis for the Selection of Dominant T Cell Receptors in Antiviral Immunity number of nucleotides at the VDJ junctional sites, called N-regions. The residual repertoire of unique TCRs following thymic selection is between 10⁷ and 10⁸ (Arstila *et al.* *Immunity* **18**, (2003).
160. Fozza, C. *et al.* Study of the T-cell receptor repertoire by CDR3 spectratyping. *Journal of Immunological Methods* **440**, 1–11 (2017).
161. Wang, C. Y. *et al.* Analysis of the CDR3 length repertoire and the diversity of T cell receptor α and β

- chains in swine CD4+ and CD8+ T lymphocytes. *Mol. Med. Rep.* **16**, 75–86 (2017).
162. Freeman, J. D., Warren, R. L., Webb, J. R., Nelson, B. H. & Holt, R. A. Profiling the T-cell receptor beta-chain repertoire by massively parallel sequencing. *Genome Res.* **19**, 1817–1824 (2009).
 163. Li, B. *et al.* Landscape of tumor-infiltrating T cell repertoire of human cancers. *Nat. Genet.* **48**, 725–732 (2016).
 164. Brown, S. D., Raeburn, L. A. & Holt, R. A. Profiling tissue-resident T cell repertoires by RNA sequencing. *Genome Med.* **7**, (2015).
 165. Chen, W. *et al.* UMI-count modeling and differential expression analysis for single-cell RNA sequencing. *Genome Biol.* **19**, 70 (2018).
 166. Fu, Y., Wu, P. H., Beane, T., Zamore, P. D. & Weng, Z. Elimination of PCR duplicates in RNA-seq and small RNA-seq using unique molecular identifiers. *BMC Genomics* **19**, 531 (2018).
 167. Adam, J. K., Odhav, B. & Bhoola, K. D. Immune responses in cancer. doi:10.1016/S0163-7258(03)00056-1
 168. Wittig, M. *et al.* Development of a high-resolution NGS-based HLA-typing and analysis pipeline. *Nucleic Acids Res.* **43**, (2015).
 169. Bentley, G. *et al.* High-resolution, high-throughput HLA genotyping by next-generation sequencing. *Tissue Antigens* **74**, 393–403 (2009).
 170. Shiina, T., Hosomichi, K., Inoko, H. & Kulski, J. K. The HLA genomic loci map: Expression, interaction, diversity and disease. *Journal of Human Genetics* **54**, 15–39 (2009).
 171. Campbell, K. S. & Purdy, A. K. Structure/function of human killer cell immunoglobulin-like receptors: Lessons from polymorphisms, evolution, crystal structures and mutations. *Immunology* **132**, 315–325 (2011).
 172. Kobayashi, K. S. & Van Den Elsen, P. J. NLRC5: A key regulator of MHC class I-dependent immune responses. *Nature Reviews Immunology* **12**, 813–820 (2012).
 173. Ozaki, Y. *et al.* Cost-efficient multiplex PCR for routine genotyping of up to nine classical HLA loci in a single analytical run of multiple samples by next generation sequencing. *BMC Genomics* **16**, (2015).
 174. Kishore, A. & Petrek, M. Next-Generation Sequencing Based HLA Typing: Deciphering Immunogenetic Aspects of Sarcoidosis. *Front. Genet.* **9**, (2018).
 175. Chowell, D. *et al.* Patient HLA class I genotype influences cancer response to checkpoint blockade immunotherapy. *Science (80-.)*. **359**, 582–587 (2018).
 176. Stockton, J. D. *et al.* Rapid, highly accurate and cost-effective open-source simultaneous complete HLA typing and phasing of class I and II alleles using nanopore sequencing. *HLA* **96**, 163–178 (2020).
 177. Tamaki, K. *et al.* Somatic Mutations and Loss of Heterozygosity of HLA Genes Are Frequently Occurred and Tightly Associated with Poor Prognosis in Adult T Cell Leukemia-Lymphoma. *Blood* **134**, 2785–2785 (2019).
 178. McGranahan, N. *et al.* Allele-Specific HLA Loss and Immune Escape in Lung Cancer Evolution. *Cell* **171**, 1259–1271.e11 (2017).
 179. Hasan Ali, O. *et al.* Human leukocyte antigen variation is associated with adverse events of checkpoint inhibitors. *Eur. J. Cancer* **107**, 8–14 (2019).
 180. Sidaway, P. Immunotherapy: HLA-1 genotype influences response to checkpoint inhibitors. *Nature Reviews Clinical Oncology* **15**, 66 (2018).
 181. Response to Immune Checkpoint Blockade is Influenced by HLA-I Genotype. *Cancer discovery* **8**, 139 (2018).
 182. Pahl, J. & Cerwenka, A. Tricking the balance: NK cells in anti-cancer immunity. *Immunobiology* **222**, 11–20 (2017).
 183. Beksac, K., Beksac, M., Dalva, K., Karaagaoglu, E. & Tirnaksiz, M. B. Impact of 'killer immunoglobulin-like receptor/ligand' genotypes on outcome following surgery among patients with colorectal cancer: Activating KIRs are associated with long-term disease free survival. *PLoS One* **10**, (2015).
 184. Morales-Estevez, C. *et al.* KIR genes and their ligands predict the response to anti-EGFR monoclonal

- antibodies in solid tumors. *Front. Immunol.* **7**, 561 (2016).
185. Pyke, R. M. *et al.* Computational KIR copy number discovery reveals interaction between inhibitory receptor burden and survival. in *Pacific Symposium on Biocomputing* **24**, 148–159 (World Scientific Publishing Co. Pte Ltd, 2019).
 186. Henry, N. L. & Hayes, D. F. Cancer biomarkers. *Mol. Oncol.* **6**, 140 (2012).
 187. Green, R. C. *et al.* ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genet. Med.* **15**, 565–574 (2013).
 188. Gallego Romero, I., Pai, A. A., Tung, J. & Gilad, Y. RNA-seq: impact of RNA degradation on transcript quantification. *BMC Biol.* **12**, 42 (2014).
 189. Early Detection and Diagnosis of Cancer Roadmap | Cancer Research UK. Available at: <https://www.cancerresearchuk.org/funding-for-researchers/research-opportunities-in-early-detection-and-diagnosis/early-detection-and-diagnosis-roadmap>. (Accessed: 26th March 2024)
 190. Hammond, M. E. H. *et al.* American society of clinical oncology/college of american pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. *J. Clin. Oncol.* **28**, 2784–2795 (2010).
 191. Wolff, A. C. *et al.* American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *J. Clin. Oncol.* **25**, 118–145 (2007).
 192. Rehm, H. L. Evolving health care through personal genomics. *Nat. Rev. Genet.* **18**, 259 (2017).
 193. Bedrosian, I. *et al.* Germline Testing in Patients with Breast Cancer: ASCO-Society of Surgical Oncology Guideline. *J. Clin. Oncol.* **42**, 584–604 (2024).
 194. Richards, S. *et al.* Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* **2015** **17**, 405–423 (2015).
 195. Angelova, M. *et al.* Characterization of the immunophenotypes and antigenomes of colorectal cancers reveals distinct tumor escape mechanisms and novel targets for immunotherapy. *Genome Biol.* **16**, 64 (2015).
 196. Chen, D. S. & Mellman, I. Oncology Meets Immunology: The Cancer-Immunity Cycle. *Immunity* **39**, 1–10 (2013).
 197. Chen, D. S. & Mellman, I. Elements of cancer immunity and the cancer-immune set point. *Nature* **541**, 321–330 (2017).
 198. Karasaki, T. *et al.* An Immunogram for the Cancer-Immunity Cycle: Towards Personalized Immunotherapy of Lung Cancer. *J. Thorac. Oncol.* **12**, 791–803 (2017).
 199. Pio, R., Ajona, D., Ortiz-Espinosa, S., Mantovani, A. & Lambris, J. D. Complementing the cancer-immunity cycle. *Frontiers in Immunology* **10**, 774 (2019).
 200. Lee, S.-H. How to Achieve a Higher Pathologic Complete Response in Patients With Locally Advanced Rectal Cancer Who Receive Preoperative Chemoradiation Therapy. *Ann. Coloproctol.* **35**, 3–8 (2019).
 201. Kokelaar, R., Evans, M., Davies, M., Harris, D. & Beynon, J. Locally advanced rectal cancer: management challenges. *Onco. Targets. Ther.* **9**, 6265–6272 (2016).
 202. Yang, J. *et al.* Genome landscapes of rectal cancer before and after preoperative chemoradiotherapy. *Theranostics* **9**, 6856–6866 (2019).
 203. Kitayama, J., Yasuda, K., Kawai, K., Sunami, E. & Nagawa, H. Circulating lymphocyte number has a positive association with tumor response in neoadjuvant chemoradiotherapy for advanced rectal cancer. *Radiat. Oncol.* **5**, (2010).
 204. Choi, C. H., Kim, W. D., Lee, S. J. & Park, W. Y. Clinical predictive factors of pathologic tumor response after preoperative chemoradiotherapy in rectal cancer. *Radiat. Oncol. J.* **30**, 99–107 (2012).
 205. Tada, N. *et al.* Prediction of the preoperative chemoradiotherapy response for rectal cancer by peripheral blood lymphocyte subsets. *World J. Surg. Oncol.* **13**, (2015).
 206. Shinto, E. *et al.* CD8+ and FOXP3+ tumor-infiltrating T cells before and after chemoradiotherapy for rectal cancer. *Ann. Surg. Oncol.* **21**, (2014).

207. Teng, F. *et al.* Tumor infiltrating lymphocytes (TILs) before and after neoadjuvant chemoradiotherapy and its clinical utility for rectal cancer. *Am. J. Cancer Res.* **5**, 2064–2074 (2015).
208. Ogura, A. *et al.* Pattern of programmed cell death-ligand 1 expression and CD8-positive T-cell infiltration before and after chemoradiotherapy in rectal cancer. *Eur. J. Cancer* **91**, 11–20 (2018).
209. Akiyoshi, T. *et al.* Immunogenomic profiles associated with response to neoadjuvant chemoradiotherapy in patients with rectal cancer. *Br. J. Surg.* **106**, 1381–1392 (2019).
210. Yasuda, K., Nirei, T., Sunami, E., Nagawa, H. & Kitayama, J. Density of CD4(+) and CD8(+) T lymphocytes in biopsy samples can be a predictor of pathological response to chemoradiotherapy (CRT) for rectal cancer. *Radiat. Oncol.* **6**, 49 (2011).
211. Li, A. *et al.* Pretreatment blood biomarkers predict pathologic responses to neo-CRT in patients with locally advanced rectal cancer. *Future Oncol.* **15**, 3233–3242 (2019).
212. Deloch, L. *et al.* Modern Radiotherapy Concepts and the Impact of Radiation on Immune Activation. *Front. Oncol.* **6**, (2016).
213. Bracci, L., Schiavoni, G., Sistigu, A. & Belardelli, F. Immune-based mechanisms of cytotoxic chemotherapy: implications for the design of novel and rationale-based combined treatments against cancer. *Cell Death Differ.* **21**, 15–25 (2014).
214. Dworak, O., Keilholz, L. & Hoffmann, A. Pathological features of rectal cancer after preoperative radiochemotherapy. *Int. J. Colorectal Dis.* **12**, 19–23 (1997).
215. Santos, M. D. *et al.* Prognostic Value of Mandard and Dworak Tumor Regression Grading in Rectal Cancer: Study of a Single Tertiary Center. *ISRN Surg.* **2014**, (2014).
216. George, T. J., Allegra, C. J. & Yothers, G. Neoadjuvant Rectal (NAR) Score: a New Surrogate Endpoint in Rectal Cancer Clinical Trials. *Curr. Colorectal Cancer Rep.* **11**, 275–280 (2015).
217. Fokas, E. *et al.* Neoadjuvant rectal score as individual-level surrogate for disease-free survival in rectal cancer in the CAO/ARO/AIO-04 randomized phase III trial. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* **29**, 1521–1527 (2018).
218. Lal, N., Beggs, A. D., Willcox, B. E. & Middleton, G. W. An immunogenomic stratification of colorectal cancer: Implications for development of targeted immunotherapy. *Oncoimmunology* **4**, 1–9 (2015).
219. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
220. Yoshihara, K. *et al.* Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* **4**, 1–11 (2013).
221. Guinney, J. *et al.* The consensus molecular subtypes of colorectal cancer. *Nat. Med.* **21**, 1350–1356 (2015).
222. Lal, N. *et al.* KRAS mutation and consensus molecular subtypes 2 and 3 are independently associated with reduced immune infiltration and reactivity in colorectal cancer. *Clin. Cancer Res.* **24**, 224–233 (2018).
223. Aleskandarany, M. A., Sonbul, S. N., Mukherjee, A. & Rakha, E. A. Molecular Mechanisms Underlying Lymphovascular Invasion in Invasive Breast Cancer. *Pathobiology* **82**, 113–123 (2015).
224. Dicken, B. J. *et al.* Lymphovascular invasion is associated with poor survival in gastric cancer: An application of gene-expression and tissue array techniques. *Ann. Surg.* **243**, 64–73 (2006).
225. Vinay, D. S. *et al.* Immune evasion in cancer: Mechanistic basis and therapeutic strategies. *Seminars in Cancer Biology* **35**, S185–S198 (2015).
226. Bedognetti, D. *et al.* Toward a comprehensive view of cancer immune responsiveness: A synopsis from the SITC workshop. *Journal for ImmunoTherapy of Cancer* **7**, 131 (2019).
227. Mayakonda, A., Lin, D. C., Assenov, Y., Plass, C. & Koeffler, H. P. Maftools: Efficient and comprehensive analysis of somatic variants in cancer. *Genome Res.* **28**, 1747–1756 (2018).
228. Matsutani, S. *et al.* Significance of tumor-infiltrating lymphocytes before and after neoadjuvant therapy for rectal cancer. *Cancer Sci.* **109**, 966–979 (2018).
229. Kong, J. C. H. *et al.* Tumor-Infiltrating Lymphocyte Function Predicts Response to Neoadjuvant Chemoradiotherapy in Locally Advanced Rectal Cancer. *JCO Precis. Oncol.* 1–15 (2018).

230. Middleton, G. *et al.* BRAF-Mutant Transcriptional Subtypes Predict Outcome of Combined BRAF, MEK, and EGFR Blockade with Dabrafenib, Trametinib, and Panitumumab in Patients with Colorectal Cancer. *Clin. Cancer Res.* **26**, 2466–2476 (2020).
231. Wilkins, A. *et al.* Differential and longitudinal immune gene patterns associated with reprogrammed microenvironment and viral mimicry in response to neoadjuvant radiotherapy in rectal cancer. *J. Immunother. Cancer* **9**, 1717 (2021).
232. Korobeinikov, A., Starkov, K. E. & Valle, P. A. Modeling cancer evolution: Evolutionary escape under immune system control. in *Journal of Physics: Conference Series* **811**, 12004 (Institute of Physics Publishing, 2017).
233. Lim, Y. W. *et al.* Germline genetic polymorphisms influence tumor gene expression and immune cell infiltration. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E11701–E11710 (2018).
234. Flemer, B. *et al.* Tumour-associated and non-tumour-associated microbiota in colorectal cancer. *Gut* **66**, 633–643 (2017).
235. Tjalsma, H. *et al.* Profiling the humoral immune response in colon cancer patients: Diagnostic antigens from *Streptococcus bovis*. *Int. J. Cancer* **119**, 2127–2135 (2006).
236. Qiu, F. *et al.* Impacts of cigarette smoking on immune responsiveness: Up and down or upside down? *Oncotarget* **8**, 268–284 (2017).
237. González Maglio, D. H., Paz, M. L. & Leoni, J. Sunlight Effects on Immune System: Is There Something Else in addition to UV-Induced Immunosuppression? *Biomed Res. Int.* **2016**, 1934518 (2016).
238. Maeda, K. *et al.* Impact of chemotherapy for colorectal cancer on regulatory T-cells and tumor immunity. *Anticancer Res.* **31**, 4569–74 (2011).
239. Thompson, S. J., Thompson, S. E. M. & Cazier, J.-B. CaStLeS (Compute and Storage for the Life Sciences): a collection of compute and storage resources for supporting research at the University of Birmingham. (2019). doi:10.5281/ZENODO.3250616
240. Johnson, W. E., Li, C. & Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8**, 118–127 (2007).
241. Aran, D., Hu, Z. & Butte, A. J. xCell: Digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol.* **18**, (2017).
242. Jiang, X., Björkström, N. K. & Melum, E. Intact CD100–CD72 Interaction Necessary for TCR-Induced T Cell Proliferation. *Front. Immunol.* **8**, 30 (2017).
243. Riboldi, E. *et al.* Human C-type lectin domain family 4, member C (CLEC4C/BDCA-2/CD303) is a receptor for asialo-galactosyl-oligosaccharides. *J. Biol. Chem.* **286**, 35329–35333 (2011).
244. Takada, H. *et al.* ADAM23, a possible tumor suppressor gene, is frequently silenced in gastric cancers by homozygous deletion or aberrant promoter hypermethylation. *Oncogene* **24**, 8051–8060 (2005).
245. Burr, M. L. *et al.* An Evolutionarily Conserved Function of Polycomb Silences the MHC Class I Antigen Presentation Pathway and Enables Immune Evasion in Cancer. (2019). doi:10.1016/j.ccell.2019.08.008
246. Wright, K. D. *et al.* The p52 isoform of SHC1 is a key driver of breast cancer initiation. *Breast Cancer Res.* **21**, 74 (2019).
247. Van Pul, K. M. *et al.* Selectively hampered activation of lymph node-resident dendritic cells precedes profound T cell suppression and metastatic spread in the breast cancer sentinel lymph node. *J. Immunother. Cancer* **7**, (2019).
248. Jones, D., Pereira, E. R. & Padera, T. P. Growth and immune evasion of lymph node metastasis. *Frontiers in Oncology* **8**, 36 (2018).
249. Pereira, E. R., Jones, D., Jung, K. & Padera, T. P. The lymph node microenvironment and its role in the progression of metastatic cancer. *Seminars in Cell and Developmental Biology* **38**, 98–105 (2015).
250. Janssen, L. M. E., Ramsay, E. E., Logsdon, C. D. & Overwijk, W. W. The immune system in cancer metastasis: friend or foe? *J. Immunother. Cancer* **5**, 79 (2017).
251. Kitamura, T., Qian, B. Z. & Pollard, J. W. Immune cell promotion of metastasis. *Nature Reviews Immunology* **15**, 73–86 (2015).

252. Mascaux, C. *et al.* Immune evasion before tumour invasion in early lung squamous carcinogenesis. *Nature* **571**, 570–575 (2019).
253. Naxerova, K. *et al.* Origins of lymphatic and distant metastases in human colorectal cancer. *Science* (80-). **357**, 55 (2017).
254. Gulubova, M. & Vlaykova, T. Prognostic significance of mast cell number and microvascular density for the survival of patients with primary colorectal cancer. *J. Gastroenterol. Hepatol.* **24**, 1265–1275 (2009).
255. Mao, Y. *et al.* Low tumor infiltrating mast cell density confers prognostic benefit and reflects immunoactivation in colorectal cancer. *Int. J. Cancer* **143**, 2271–2280 (2018).
256. Malinowsky, K. *et al.* Activation of the PI3K/AKT pathway correlates with prognosis in stage II colon cancer. *Br. J. Cancer* **110**, 2081–2089 (2014).
257. Sheffer, M. *et al.* Association of survival and disease progression with chromosomal instability: A genomic exploration of colorectal cancer. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 7131–7136 (2009).
258. Luke, J. J., Bao, R., Sweis, R. F., Spranger, S. & Gajewski, T. F. WNT/b-catenin pathway activation correlates with immune exclusion across human cancers. *Clin. Cancer Res.* **25**, 3074–3083 (2019).
259. Garriss, C. S. *et al.* Successful Anti-PD-1 Cancer Immunotherapy Requires T Cell-Dendritic Cell Crosstalk Involving the Cytokines IFN- γ and IL-12. *Immunity* **49**, 1148–1161.e7 (2018).
260. Brown, C. C. *et al.* Transcriptional Basis of Mouse and Human Dendritic Cell Heterogeneity In Brief Single-cell analyses of dendritic cells reveals new subsets with distinct pro-and anti-inflammatory potential. *Cell* **179**, 846–863 (2019).
261. Wculek, S. K. *et al.* Dendritic cells in cancer immunology and immunotherapy. *Nat. Rev. Immunol.* doi:10.1038/s41577-019-0210-z
262. Walker, L. S. K. Treg and CTLA-4: two intertwining pathways to immune tolerance. *J. Autoimmun.* **45**, 49–57 (2013).
263. Dobrzanski, M. J. Expanding roles for CD4 T cells and their subpopulations in tumor immunity and therapy. *Frontiers in Oncology* **3 MAR**, 63 (2013).
264. Angelova, M. *et al.* Evolution of Metastases in Space and Time under Immune Selection. *Cell* **175**, 751–765.e16 (2018).
265. Lee, J.-R. *et al.* Transcriptome analysis of paired primary colorectal carcinoma and liver metastases reveals fusion transcripts and similar gene expression profiles in primary carcinoma and liver metastases. *BMC Cancer* **16**, 539 (2016).
266. Kamal, Y., Schmit, S. L., Hoehn, H. J., Amos, C. I. & Frost, H. R. Transcriptomic differences between primary colorectal adenocarcinomas and distant metastases reveal metastatic colorectal cancer subtypes. *Cancer Res.* **79**, 4227–4241 (2019).
267. Tabernero, J. *et al.* Exploratory analysis of the effect of FTD/TPI in patients treated in RECURSE by prognostic factors. *J. Clin. Oncol.* **37**, 677–677 (2019).
268. Granot, T. *et al.* Dendritic Cells Display Subset and Tissue-Specific Maturation Dynamics over Human Life. *Immunity* **46**, 504–515 (2017).
269. A Phase II Trial Assessing Nivolumab in Class II Expressing Microsatellite Stable Colorectal Cancer - Full Text View - ClinicalTrials.gov. Available at: <https://clinicaltrials.gov/ct2/show/NCT03981146>. (Accessed: 29th September 2020)
270. Hunter, M. V., Moncada, R., Weiss, J. M., Yanai, I. & White, R. M. Spatially resolved transcriptomics reveals the architecture of the tumor-microenvironment interface. *Nat. Commun.* **2021 121 12**, 1–16 (2021).
271. Arora, R. *et al.* Spatial transcriptomics reveals distinct and conserved tumor core and edge architectures that predict survival and targeted therapy response. *Nat. Commun.* **2023 141 14**, 1–19 (2023).
272. Alexander, P. G., McMillan, D. C. & Park, J. H. The local inflammatory response in colorectal cancer – Type, location or density? A systematic review and meta-analysis. *Cancer Treat. Rev.* **83**, 101949 (2020).
273. Weng, J. *et al.* Exploring immunotherapy in colorectal cancer. *J. Hematol. Oncol.* **2022 151 15**, 1–28 (2022).

274. Schaarschmidt, S., Fischer, A., Zuther, E. & Hinch, D. K. Evaluation of Seven Different RNA-Seq Alignment Tools Based on Experimental Data from the Model Plant *Arabidopsis thaliana*. *Int. J. Mol. Sci.* **21**, (2020).
275. Musich, R., Cadle-Davidson, L. & Osier, M. V. Comparison of Short-Read Sequence Aligners Indicates Strengths and Weaknesses for Biologists to Consider. *Front. Plant Sci.* **12**, (2021).
276. Guo, Y. *et al.* Improvements and impacts of GRCh38 human reference on high throughput sequencing data analysis. *Genomics* **109**, 83–90 (2017).
277. Lavezzo, E. *et al.* Suppression of a SARS-CoV-2 outbreak in the Italian municipality of Vo'. *Nature* **584**, 425 (2020).
278. COVID-19 Map - Johns Hopkins Coronavirus Resource Center. Available at: <https://coronavirus.jhu.edu/map.html>. (Accessed: 9th October 2023)
279. Wu, C. *et al.* Risk Factors Associated With Acute Respiratory Distress Syndrome and Death in Patients With Coronavirus Disease 2019 Pneumonia in Wuhan, China. *JAMA Intern. Med.* (2020). doi:10.1001/jamainternmed.2020.0994
280. Wang, D. *et al.* Clinical Characteristics of 138 Hospitalized Patients with 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. *JAMA - J. Am. Med. Assoc.* **323**, 1061–1069 (2020).
281. Yang, X. *et al.* Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. *Lancet Respir. Med.* (2020). doi:10.1016/S2213-2600(20)30079-5
282. Pollán, M. *et al.* Prevalence of SARS-CoV-2 in Spain (ENE-COVID): a nationwide, population-based seroepidemiological study. *Lancet* **396**, 535–544 (2020).
283. Várnai, C. *et al.* Mortality Among Adults With Cancer Undergoing Chemotherapy or Immunotherapy and Infected With COVID-19. *JAMA Netw. Open* **5**, (2022).
284. Lee, L. Y. W. *et al.* COVID-19 mortality in patients with cancer on chemotherapy or other anticancer treatments: a prospective cohort study. *Lancet (London, England)* **395**, 1919 (2020).
285. Dai, M. *et al.* Patients with cancer appear more vulnerable to SARS-COV-2: a multi-center study during the COVID-19 outbreak. *Cancer Discov.* CD-20-0422 (2020). doi:10.1158/2159-8290.CD-20-0422
286. Petrilli, C. M. *et al.* Factors associated with hospitalization and critical illness among 4,103 patients with COVID-19 disease in New York City. *medRxiv* 2020.04.08.20057794 (2020). doi:10.1101/2020.04.08.20057794
287. Gt Walker, P. *et al.* *The Global Impact of COVID-19 and Strategies for Mitigation and Suppression*.
288. The Health Protection (Coronavirus, Restrictions) (England) Regulations 2020. Available at: <https://www.legislation.gov.uk/ukSI/2020/350>. (Accessed: 19th October 2023)
289. Liang, W. *et al.* Cancer patients in SARS-CoV-2 infection: a nationwide analysis in China. *The Lancet Oncology* **21**, 335–337 (2020).
290. Yu, J., Ouyang, W., Chua, M. L. K. & Xie, C. SARS-CoV-2 Transmission in Patients With Cancer at a Tertiary Care Hospital in Wuhan, China. *JAMA Oncol.* (2020). doi:10.1001/jamaoncol.2020.0980
291. Zhang, L. *et al.* Clinical characteristics of COVID-19-infected cancer patients: a retrospective case study in three hospitals within Wuhan, China. *Ann. Oncol.* (2020). doi:10.1016/j.annonc.2020.03.296
292. Miyashita, H. *et al.* Do patients with cancer have a poorer prognosis of COVID-19? An experience in New York City. *Ann. Oncol.* **31**, 1088 (2020).
293. Banna, G., Curioni-Fontecedro, A., Friedlaender, A. & Addeo, A. How we treat patients with lung cancer during the SARS-CoV-2 pandemic: primum non nocere. *ESMO Open* **4**, (2019).
294. Guidelines | ESMO. Available at: <https://www.esmo.org/guidelines/cancer-patient-management-during-the-covid-19-pandemic>. (Accessed: 10th October 2023)
295. Spicer, J., Chamberlain, C. & Papa, S. Provision of cancer care during the COVID-19 pandemic. *Nat. Rev. Clin. Oncol.* **17**, 329 (2020).
296. Nepogodiev, D. *et al.* Elective surgery cancellations due to the COVID-19 pandemic: global predictive modelling to inform surgical recovery plans. *Br. J. Surg.* **107**, 1440–1449 (2020).

297. England Summary | Coronavirus (COVID-19) in the UK. Available at: <https://coronavirus.data.gov.uk/>. (Accessed: 9th October 2023)
298. Spicer, J., Chamberlain, C. & Papa, S. Provision of cancer care during the COVID-19 pandemic. *Nat. Rev. Clin. Oncol.* **17**, 329 (2020).
299. Overview | COVID-19 rapid guideline: delivery of systemic anticancer treatments | Guidance | NICE.
300. Coles, C. Guidelines on radiation therapy for breast cancer during the COVID-19 pandemic. doi:10.1016/S1470-2045(14)71221-5
301. Vegivinti, C. T. R. *et al.* Efficacy of antiviral therapies for COVID-19: a systematic review of randomized controlled trials. *BMC Infect. Dis.* **22**, (2022).
302. NHS Test and Trace: how it works - GOV.UK. Available at: <https://www.gov.uk/guidance/nhs-test-and-trace-how-it-works>. (Accessed: 6th October 2020)
303. Mistry, D. A., Wang, J. Y., Moeser, M. E., Starkey, T. & Lee, L. Y. W. A systematic review of the sensitivity and specificity of lateral flow devices in the detection of SARS-CoV-2. *BMC Infect. Dis.* **21**, 1–14 (2021).
304. He, J., Guo, Y., Mao, R. & Zhang, J. Proportion of asymptomatic coronavirus disease 2019: A systematic review and meta-analysis. *J. Med. Virol.* jmv.26326 (2020). doi:10.1002/jmv.26326
305. Singanayagam, A. *et al.* Duration of infectiousness and correlation with RT-PCR cycle threshold values in cases of COVID-19, England, January to May 2020. *Eurosurveillance* **25**, 2001483 (2020).
306. Wiersinga, W. J., Rhodes, A., Cheng, A. C., Peacock, S. J. & Prescott, H. C. Pathophysiology, Transmission, Diagnosis, and Treatment of Coronavirus Disease 2019 (COVID-19): A Review. *JAMA* **324**, 782–793 (2020).
307. COVID-19 Response: Living with COVID-19 - GOV.UK. Available at: <https://www.gov.uk/government/publications/covid-19-response-living-with-covid-19/covid-19-response-living-with-covid-19>. (Accessed: 16th June 2023)
308. Lu, R. *et al.* Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* **395**, 565–574 (2020).
309. Channappanavar, R. & Perlman, S. Pathogenic human coronavirus infections: causes and consequences of cytokine storm and immunopathology. *Seminars in Immunopathology* **39**, 529–539 (2017).
310. He, X. *et al.* Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat. Med.* **2020** 265 **26**, 672–675 (2020).
311. Takahashi, K. *et al.* Duration of Infectious Virus Shedding by SARS-CoV-2 Omicron Variant–Infected Vaccinees. *Emerg. Infect. Dis.* **28**, 998 (2022).
312. *Coronavirus disease 2019 (COVID-19) Situation Report-72 HIGHLIGHTS.*
313. Florindo, H. F. *et al.* Immune-mediated approaches against COVID-19. *Nat. Nanotechnol.* **2020** 158 **15**, 630–645 (2020).
314. Shirbhate, E. *et al.* Understanding the role of ACE-2 receptor in pathogenesis of COVID-19 disease: a potential approach for therapeutic intervention. *Pharmacol. Reports* **73**, 1539 (2021).
315. Gottschalk, G., Knox, K. & Roy, A. ACE2: At the crossroad of COVID-19 and lung cancer. *Gene Reports* **23**, 101077 (2021).
316. Zhong, P. *et al.* COVID-19-associated gastrointestinal and liver injury: clinical features and potential mechanisms. *Signal Transduct. Target. Ther.* **2020** 51 **5**, 1–8 (2020).
317. Zhou, P. *et al.* Discovery of a novel coronavirus associated with the recent pneumonia outbreak in humans and its potential bat origin. *Nature* **2020.01.22.914952** (2020). doi:10.1101/2020.01.22.914952
318. Jiang, Y., Yin, W. & Xu, H. E. RNA-dependent RNA polymerase: Structure, mechanism, and drug discovery for COVID-19. *Biochem. Biophys. Res. Commun.* **538**, 47 (2021).
319. Burki, T. First shared SARS-CoV-2 genome: GISAID vs virological.org. *The Lancet. Microbe* **4**, e395 (2023).
320. Gordon, D. E. *et al.* A SARS-CoV-2 Protein Interaction Map Reveals Targets for Drug-Repurposing.

- Nature* **583**, 459 (2020).
321. Littler, D. R., Gully, B. S., Colson, R. N. & Rossjohn, J. Crystal Structure of the SARS-CoV-2 Non-structural Protein 9, Nsp9. *iScience* **23**, 101258 (2020).
 322. Pillon, M. C. *et al.* Cryo-EM structures of the SARS-CoV-2 endoribonuclease Nsp15 reveal insight into nuclease specificity and dynamics. *Nat. Commun.* **12**, (2021).
 323. Rosas-Lemus, M. *et al.* The crystal structure of nsp10-nsp16 heterodimer from SARS-CoV-2 in complex with S-adenosylmethionine. *bioRxiv* (2020). doi:10.1101/2020.04.17.047498
 324. Bayati, A., Kumar, R., Francis, V. & Mcpherson, P. S. SARS-CoV-2 infects cells after viral entry via clathrin-mediated endocytosis. *J. Biol. Chem.* **296**, 100306 (2021).
 325. Ward, S. E. *et al.* Von Willebrand factor propeptide in severe coronavirus disease 2019 (COVID-19): evidence of acute and sustained endothelial cell activation. *Br. J. Haematol.* **192**, 714–719 (2021).
 326. Leyfman, Y. *et al.* Cancer and COVID-19: unravelling the immunological interplay with a review of promising therapies against severe SARS-CoV-2 for cancer patients. *J. Hematol. Oncol.* **2023 161** **16**, 1–15 (2023).
 327. Townsend, L. *et al.* Severe COVID-19 is characterised by inflammation and immature myeloid cells early in disease progression. *Heliyon* **8**, e09230 (2022).
 328. Hajjo, R., Sabbah, D. A. & Tropsha, A. Analyzing the Systems Biology Effects of COVID-19 mRNA Vaccines to Assess Their Safety and Putative Side Effects. *Pathogens* **11**, (2022).
 329. McGonagle, D., Bridgewood, C. & Meaney, J. F. M. A tricompartamental model of lung oxygenation disruption to explain pulmonary and systemic pathology in severe COVID-19. *Lancet Respir. Med.* **9**, 665–672 (2021).
 330. Redd, W. D. *et al.* Prevalence and Characteristics of Gastrointestinal Symptoms in Patients With Severe Acute Respiratory Syndrome Coronavirus 2 Infection in the United States: A Multicenter Cohort Study. *Gastroenterology* **159**, 765 (2020).
 331. El Moheb, M. *et al.* Gastrointestinal Complications in Critically Ill Patients With and Without COVID-19. *JAMA* **324**, 1899–1901 (2020).
 332. Akuffo-Addo, E., Nicholas, M. N. & Joseph, M. COVID-19 Skin Manifestations in Skin of Colour. *J. Cutan. Med. Surg.* **26**, 189 (2022).
 333. Lee, C. C. E. *et al.* COVID-19-Associated Cardiovascular Complications. *Dis. 2021, Vol. 9, Page 47* **9**, 47 (2021).
 334. Guzik, T. J. *et al.* COVID-19 and the cardiovascular system: implications for risk assessment, diagnosis, and treatment options. *Cardiovasc. Res.* **116**, 1666–1687 (2020).
 335. Boukhris, M. *et al.* Cardiovascular Implications of the COVID-19 Pandemic: A Global Perspective. *Can. J. Cardiol.* **36**, 1068 (2020).
 336. Grimes, J. M. & Grimes, K. V. p38 MAPK inhibition: A promising therapeutic approach for COVID-19. *J. Mol. Cell. Cardiol.* **144**, 63 (2020).
 337. Zhang, L., Han, X. & Shi, Y. Comparative analysis of SARS-CoV-2 receptor ACE2 expression in multiple solid tumors and matched non-diseased tissues. *Infect. Genet. Evol.* **85**, 104428 (2020).
 338. Murphy, K. T. *et al.* MAs receptor activation slows tumor growth and attenuates muscle wasting in cancer. *Cancer Res.* **79**, 706–719 (2019).
 339. Lee, V. Y. *et al.* Bleomycin induces alveolar epithelial cell death through JNK-dependent activation of the mitochondrial death pathway. *Am. J. Physiol. - Lung Cell. Mol. Physiol.* **289**, 521–528 (2005).
 340. Ahmadi, M. *et al.* ACE2 correlates with immune infiltrates in colon adenocarcinoma: Implication for COVID-19. *Int. Immunopharmacol.* **95**, 107568 (2021).
 341. Wang, H. & Yang, J. Colorectal Cancer that Highly Express Both ACE2 and TMPRSS2, Suggesting Severe Symptoms to SARS-CoV-2 Infection. *Pathol. Oncol. Res.* **27**, (2021).
 342. Liu, C. *et al.* High expression of ACE2 and TMPRSS2 and clinical characteristics of COVID-19 in colorectal cancer patients. *npj Precis. Oncol.* **2021 51** **5**, 1–7 (2021).
 343. Kondratiuk, A. L., Pillay, T. D., Kon, O. M. & Lalvani, A. A conceptual framework to accelerate the clinical

- impact of evolving research into long COVID. *Lancet Infect. Dis.* **21**, 756–757 (2021).
344. Su, S. *et al.* Epidemiology, clinical presentation, pathophysiology, and management of long COVID: an update. *Mol. Psychiatry* **2023** 1–14 (2023). doi:10.1038/s41380-023-02171-3
 345. Leyfman, Y. *et al.* REVIEW Open Access Cancer and COVID-19: unravelling the immunological interplay with a review of promising therapies against severe SARS-CoV-2 for cancer patients. *J. Hematol. Oncol.* **16**, 39 (2023).
 346. Chen, C. *et al.* Global Prevalence of Post-Coronavirus Disease 2019 (COVID-19) Condition or Long COVID: A Meta-Analysis and Systematic Review. *J. Infect. Dis.* **226**, 1593–1607 (2022).
 347. A clinical case definition of post COVID-19 condition by a Delphi consensus, 6 October 2021. Available at: https://www.who.int/publications/i/item/WHO-2019-nCoV-Post_COVID-19_condition-Clinical_case_definition-2021.1. (Accessed: 9th October 2023)
 348. Montefusco, L. *et al.* Acute and long-term disruption of glycometabolic control after SARS-CoV-2 infection. *Nat. Metab.* **2021** **3**, 774–785 (2021).
 349. Fernández-Ortega, M. *et al.* Cognitive dysfunction, diabetes mellitus 2 and arterial hypertension: Sequelae up to one year of COVID-19. *Travel Med. Infect. Dis.* **52**, 102553 (2023).
 350. Ahamed, J. & Laurence, J. Long COVID endotheliopathy: hypothesized mechanisms and potential therapeutic approaches. *J. Clin. Invest.* **132**, (2022).
 351. Castanares-Zapatero, D. *et al.* Pathophysiology and mechanism of long COVID: a comprehensive review. *Ann. Med.* **54**, 1473–1487 (2022).
 352. Mantovani, A. *et al.* Long Covid: where we stand and challenges ahead. *Cell Death Differ.* **29**, 1891 (2022).
 353. Monje, M. & Iwasaki, A. The neurobiology of long COVID. *Neuron* **110**, 3484 (2022).
 354. Gold, J. E., Okyay, R. A., Licht, W. E. & Hurley, D. J. Investigation of Long COVID Prevalence and Its Relationship to Epstein-Barr Virus Reactivation. *Pathogens* **10**, (2021).
 355. Su, Y. *et al.* Multiple early factors anticipate post-acute COVID-19 sequelae. *Cell* **185**, 881 (2022).
 356. Wallukat, G. *et al.* Functional autoantibodies against G-protein coupled receptors in patients with persistent Long-COVID-19 symptoms. *J. Transl. Autoimmun.* **4**, 100100 (2021).
 357. Carabelli, A. M. *et al.* SARS-CoV-2 variant biology: immune escape, transmission and fitness. *Nat. Rev. Microbiol.* **2023** **213** **21**, 162–177 (2023).
 358. Markov, P. V. *et al.* The evolution of SARS-CoV-2. *Nat. Rev. Microbiol.* **2023** **216** **21**, 361–379 (2023).
 359. Korber, B. *et al.* Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. *Cell* **182**, 812 (2020).
 360. Dayan, G. H. *et al.* Efficacy of a monovalent (D614) SARS-CoV-2 recombinant protein vaccine with AS03 adjuvant in adults: a phase 3, multi-country study. *eClinicalMedicine* 102168 (2023). doi:10.1016/J.ECLINM.2023.102168
 361. Harvey, W. T. *et al.* SARS-CoV-2 variants, spike mutations and immune escape. *Nat. Rev. Microbiol.* **2021** **197** **19**, 409–424 (2021).
 362. Pascall, D. J. *et al.* Directions of change in intrinsic case severity across successive SARS-CoV-2 variant waves have been inconsistent. *J. Infect.* **87**, 128–135 (2023).
 363. Wolter, N. *et al.* Early assessment of the clinical severity of the SARS-CoV-2 omicron variant in South Africa: a data linkage study. *Lancet* **399**, 437–446 (2022).
 364. Bager, P. *et al.* Risk of hospitalisation associated with infection with SARS-CoV-2 omicron variant versus delta variant in Denmark: an observational cohort study. *Lancet Infect. Dis.* **22**, 967–976 (2022).
 365. Nyberg, T. *et al.* Comparative analysis of the risks of hospitalisation and death associated with SARS-CoV-2 omicron (B.1.1.529) and delta (B.1.617.2) variants in England: a cohort study. *Lancet* **399**, 1303–1312 (2022).
 366. Saito, A. *et al.* Enhanced fusogenicity and pathogenicity of SARS-CoV-2 Delta P681R mutation. *Nat.* **2021** **6027896** **602**, 300–306 (2021).

367. Suzuki, R. *et al.* Attenuated fusogenicity and pathogenicity of SARS-CoV-2 Omicron variant. *Nat.* 2022 6037902 **603**, 700–705 (2022).
368. Hui, K. P. Y. *et al.* SARS-CoV-2 Omicron variant replication in human bronchus and lung ex vivo. *Nat.* 2022 6037902 **603**, 715–720 (2022).
369. Aggarwal, A. *et al.* Platform for isolation and characterization of SARS-CoV-2 variants enables rapid characterization of Omicron in Australia. *Nat. Microbiol.* 2022 76 **7**, 896–908 (2022).
370. Torretta, S. *et al.* Diagnosis of SARS-CoV-2 by RT-PCR Using Different Sample Sources: Review of the Literature. *Ear. Nose. Throat J.* **100**, 131S (2021).
371. Kubina, R. & Dziedzic, A. Molecular and Serological Tests for COVID-19. A Comparative Review of SARS-CoV-2 Coronavirus Laboratory and Point-of-Care Diagnostics. *Diagnostics* **10**, (2020).
372. Mallett, S. *et al.* At what times during infection is SARS-CoV-2 detectable and no longer detectable using RT-PCR-based tests? A systematic review of individual participant data. *BMC Med.* **18**, 1–17 (2020).
373. Dinnes, J. *et al.* Rapid, point-of-care antigen tests for diagnosis of SARS-CoV-2 infection. *Cochrane Database Syst. Rev.* **2021**, (2021).
374. Peto, T. *et al.* COVID-19: Rapid antigen detection for SARS-CoV-2 by lateral flow assay: A national systematic evaluation of sensitivity and specificity for mass-testing. *EClinicalMedicine* **36**, 100924 (2021).
375. Hsiao, W. W. W. *et al.* Recent Advances in Novel Lateral Flow Technologies for Detection of COVID-19. *Biosensors* **11**, (2021).
376. Tang, E. W. & Bobenchik, A. M. The Promise and Peril of Antibody Testing for COVID-19. *JAMA* **323**, 1881–1883 (2020).
377. Tang, E. W. & Bobenchik, A. M. Testing for SARS-CoV-2 (COVID-19): A General Review. doi:10.1128/JVI.01756-13
378. Bauer, G. The variability of the serological response to SARS-corona virus-2: Potential resolution of ambiguity through determination of avidity (functional affinity). *J. Med. Virol.* **93**, 311 (2021).
379. Ravi, N., Cortade, D. L., Ng, E. & Wang, S. X. Diagnostics for SARS-CoV-2 detection: A comprehensive review of the FDA-EUA COVID-19 testing landscape. *Biosens. Bioelectron.* **165**, 112454 (2020).
380. Elecsys® Anti-SARS-CoV-2. Available at: <https://diagnostics.roche.com/gb/en/products/params/electsys-anti-sars-cov-2.html>. (Accessed: 9th October 2023)
381. Al Hajji, Y., Taylor, H., Starkey, T., Lee, L. Y. W. & Tilby, M. Antibody response to a third booster dose of SARS-CoV-2 vaccination in adults with haematological and solid cancer: a systematic review. *Br. J. Cancer* **127**, 1827 (2022).
382. Heinz, F. X. & Stiasny, K. Distinguishing features of current COVID-19 vaccines: knowns and unknowns of antigen presentation and modes of action. *npj Vaccines* 2021 61 **6**, 1–13 (2021).
383. Bellamkonda, N. *et al.* Immune Response to SARS-CoV-2 Vaccines. *Biomedicines* **10**, 1464 (2022).
384. Le, T. T., Cramer, J. P., Chen, R. & Mayhew, S. Evolution of the COVID-19 vaccine development landscape. *Nat. Rev. Drug Discov.* **19**, 667–668 (2020).
385. Arbeitman, C. R., Rojas, P., Ojeda-May, P. & Garcia, M. E. The SARS-CoV-2 spike protein is vulnerable to moderate electric fields. *Nat. Commun.* **12**, (2021).
386. Grifoni, A. *et al.* Targets of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed Individuals. *Cell* **181**, 1489 (2020).
387. Tozinameran: Uses, Interactions, Mechanism of Action | DrugBank Online. Available at: <https://go.drugbank.com/drugs/DB15696>. (Accessed: 9th October 2023)
388. Moderna COVID-19 Vaccine: Uses, Interactions, Mechanism of Action | DrugBank Online. Available at: <https://go.drugbank.com/drugs/DB15654>. (Accessed: 9th October 2023)
389. AstraZeneca COVID-19 Vaccine: Uses, Interactions, Mechanism of Action | DrugBank Online. Available at: <https://go.drugbank.com/drugs/DB15656>. (Accessed: 9th October 2023)
390. Jackson, N. A. C., Kester, K. E., Casimiro, D., Gurunathan, S. & DeRosa, F. The promise of mRNA vaccines: a biotech and industrial perspective. *npj Vaccines* 2020 51 **5**, 1–6 (2020).

391. Mendonça, S. A., Lorincz, R., Boucher, P. & Curiel, D. T. Adenoviral vector vaccine platforms in the SARS-CoV-2 pandemic. *npj Vaccines* 2021 61 **6**, 1–14 (2021).
392. Morris, S. J., Sebastian, S., Spencer, A. J. & Gilbert, S. C. Simian adenoviruses as vaccine vectors. *Future Virol.* **11**, 649–659 (2016).
393. Voysey, M. *et al.* Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. *www.thelancet.com* **397**, 2021 (2020).
394. Folegatti, P. M. *et al.* Safety and immunogenicity of the ChAdOx1 nCoV-19 vaccine against SARS-CoV-2: a preliminary report of a phase 1/2, single-blind, randomised controlled trial. *Lancet* **396**, 467–478 (2020).
395. Dutta, N. K., Mazumdar, K. & Gordy, J. T. The Nucleocapsid Protein of SARS-CoV-2: a Target for Vaccine Development. *J. Virol.* **94**, (2020).
396. Nikolaidis, M., Markoulatos, P., Van De Peer, Y., Oliver, S. G. & Amoutzias, G. D. The Neighborhood of the Spike Gene Is a Hotspot for Modular Intertypic Homologous and Nonhomologous Recombination in Coronavirus Genomes. *Mol. Biol. Evol.* **39**, (2022).
397. Treatments for COVID-19 - NHS. Available at: <https://www.nhs.uk/conditions/covid-19/treatments-for-covid-19/>. (Accessed: 9th October 2023)
398. Bakheit, A. H., Darwish, H., Darwish, I. A. & Al-Ghusn, A. I. Remdesivir. *Profiles Drug Subst. Excip. Relat. Methodol.* **48**, 71 (2023).
399. Akinosoglou, K., Schinas, G. & Gogos, C. Oral Antiviral Treatment for COVID-19: A Comprehensive Review on Nirmatrelvir/Ritonavir. *Viruses* **14**, (2022).
400. Toots, M. *et al.* Characterization of orally efficacious influenza drug with high resistance barrier in ferrets and human airway epithelia. *Sci. Transl. Med.* **11**, (2019).
401. Toots, M. *et al.* Quantitative Efficacy Paradigms of the Influenza Clinical Drug Candidate EIDD-2801 in the Ferret Model. *Transl. Res.* **218**, 16 (2020).
402. Amara, A. *et al.* The development and validation of a novel LC-MS/MS method for the simultaneous quantification of Molnupiravir and its metabolite β -d-N4-hydroxycytidine in human plasma and saliva. *J. Pharm. Biomed. Anal.* **206**, 114356 (2021).
403. Painter, W. P. *et al.* Human Safety, Tolerability, and Pharmacokinetics of Molnupiravir, a Novel Broad-Spectrum Oral Antiviral Agent with Activity against SARS-CoV-2. *Antimicrob. Agents Chemother.* **65**, (2021).
404. Pashaei, Y. Analytical methods for the determination of remdesivir as a promising antiviral candidate drug for the COVID-19 pandemic. *Drug Discov. Ther.* **14**, 273–281 (2020).
405. Gordon, C. J., Tchesnokov, E. P., Feng, J. Y., Porter, D. P. & Götte, M. The antiviral compound remdesivir potently inhibits RNA-dependent RNA polymerase from Middle East respiratory syndrome coronavirus. *J. Biol. Chem.* **295**, 4773 (2020).
406. Ferner, R. E. & Aronson, J. K. Remdesivir in covid-19. *BMJ* **369**, (2020).
407. Tchesnokov, E. P., Feng, J. Y., Porter, D. P. & Götte, M. Mechanism of Inhibition of Ebola Virus RNA-Dependent RNA Polymerase by Remdesivir. *Viruses* 2019, Vol. 11, Page 326 **11**, 326 (2019).
408. Wang, Y. *et al.* Remdesivir in adults with severe COVID-19: a randomised, double-blind, placebo-controlled, multicentre trial. *Lancet (London, England)* **395**, 1569 (2020).
409. Mehta, N., Mazer-Amirshahi, M., Alkindi, N. & Pourmand, A. Pharmacotherapy in COVID-19; A narrative review for emergency providers. *Am. J. Emerg. Med.* **38**, 1488 (2020).
410. Pfizer. FACT SHEET FOR HEALTHCARE PROVIDERS: EMERGENCY USE AUTHORIZATION FOR PAXLOVID™ HIGHLIGHTS OF EMERGENCY USE AUTHORIZATION (EUA) These highlights of the EUA do not include all the information needed to use PAXLOVID™ under the EUA. See the FULL FACT SHEET FOR HEALTHCARE PROVIDERS for PAXLOVID. PAXLOVID (nirmatrelvir tablets; ritonavir tablets), co-packaged for oral use Original EUA Authorized Date.
411. Merck, Co & Inc. FACT SHEET FOR HEALTHCARE PROVIDERS: EMERGENCY USE AUTHORIZATION FOR LAGEVRIO™ (molnupiravir) CAPSULES HIGHLIGHTS OF EMERGENCY USE AUTHORIZATION (EUA) These highlights of the EUA do not include all the information needed to use

LAGEVRIO under the EUA. See the FULL FACT SHEET FOR HEALTHCARE PROVIDERS for LAGEVRIO. LAGEVRIO™ (molnupiravir) capsules, for oral use Original EUA Authorized Date: 12/23/2021 Revised EUA. (2023).

412. Focosi, D. *et al.* Monoclonal antibody therapies against SARS-CoV-2. *Lancet Infect. Dis.* **22**, e311–e326 (2022).
413. Medical Association, A. Tixagevimab and Cilgavimab (Evusheld) for Pre-Exposure Prophylaxis of COVID-19. *JAMA* **327**, 384–385 (2022).
414. Levin, M. J. *et al.* Intramuscular AZD7442 (Tixagevimab–Cilgavimab) for Prevention of Covid-19. *N. Engl. J. Med.* **386**, 2188–2200 (2022).
415. Suribhatla, R. *et al.* Systematic review of the clinical effectiveness of Tixagevimab/Cilgavimab for prophylaxis of COVID-19 in immunocompromised patients. *medRxiv* 2022.11.07.22281786 (2022). doi:10.1101/2022.11.07.22281786
416. Drysdale, M. *et al.* Real-world effectiveness of sotrovimab for the treatment of SARS-CoV-2 infection during Omicron BA.2 subvariant predominance: a systematic literature review. *Infection* **1**, 1–17 (2023).
417. Amani, B. & Amani, B. Efficacy and safety of sotrovimab in patients with COVID-19: A rapid review and meta-analysis. *Rev. Med. Virol.* **32**, (2022).
418. Cox, M. G. *et al.* SARS-CoV-2 variant evasion of monoclonal antibodies based on in vitro studies. *Nat. Rev. Microbiol.* 2022 212 **21**, 112–124 (2022).
419. Arora, P. *et al.* Omicron sublineage BQ.1.1 resistance to monoclonal antibodies. *Lancet. Infect. Dis.* **23**, 22 (2023).
420. Cao, Y. *et al.* Omicron escapes the majority of existing SARS-CoV-2 neutralizing antibodies. *Nature* **602**, 657 (2022).
421. Wang, Q. *et al.* Alarming antibody evasion properties of rising SARS-CoV-2 BQ and XBB subvariants. *Cell* **186**, 279 (2023).
422. First participant dosed in SUPERNOVA Phase I/III trial evaluating AZD5156, a next-generation long-acting antibody combination, for prevention of COVID-19. Available at: <https://www.astrazeneca-us.com/media/statements/2022/first-participant-dosed-in-supernova-phase-I-III-trial-evaluating-azd5156-a-next-generation-long-acting-antibody-combination-for-prevention-of-covid-19.html>. (Accessed: 10th October 2023)
423. ECCMID data reinforces AstraZeneca’s commitment to transform protection for the most vulnerable by advancing science in vaccines and immune therapies. Available at: <https://www.astrazeneca.com/media-centre/press-releases/2023/eccmid-data-reinforces-astrazenecas-commitment-to-transform-protection-for-the-most-vulnerable-by-advancing-science-in-vaccines-and-immune-therapies.html>. (Accessed: 10th October 2023)
424. Guan, W. *et al.* Clinical Characteristics of Coronavirus Disease 2019 in China. *N. Engl. J. Med.* **382**, 1708–1720 (2020).
425. Hu, B., Guo, H., Zhou, P. & Shi, Z. L. Characteristics of SARS-CoV-2 and COVID-19. *Nat. Rev. Microbiol.* **19**, 141 (2021).
426. Thompson, C. N. *et al.* COVID-19 Outbreak — New York City, February 29–June 1, 2020. *Morb. Mortal. Wkly. Rep.* **69**, 1725 (2020).
427. Muhsen, K. *et al.* A nationwide analysis of population group differences in the COVID-19 epidemic in Israel, February 2020–February 2021. *Lancet Reg. Heal. - Eur.* **7**, 100130 (2021).
428. Hashim, M. J., Alsuwaidi, A. R. & Khan, G. Population Risk Factors for COVID-19 Mortality in 93 Countries. (2020). doi:10.2991/jegh.k.200721.001
429. Weinberger, D. M. *et al.* Estimation of Excess Deaths Associated With the COVID-19 Pandemic in the United States, March to May 2020. *JAMA Intern. Med.* **180**, 1336–1344 (2020).
430. Williamson, E. J. *et al.* OpenSAFELY: factors associated with COVID-19 death in 17 million patients. *Nature* **584**, 430 (2020).
431. Lighter, J. *et al.* Obesity in patients younger than 60 years is a risk factor for COVID-19 hospital admission. *Clin. Infect. Dis.* **71**, 896–897 (2020).
432. Simonnet, A. *et al.* High Prevalence of Obesity in Severe Acute Respiratory Syndrome Coronavirus-2

- (SARS-CoV-2) Requiring Invasive Mechanical Ventilation. *Obesity* **28**, 1195–1199 (2020).
433. Rentsch, C. T. *et al.* Covid-19 Testing, Hospital Admission, and Intensive Care Among 2,026,227 United States Veterans Aged 54–75 Years. (2020). doi:10.1101/2020.04.09.20059964
 434. Halpin, D. M. G., Faner, R., Sibila, O., Badia, J. R. & Agusti, A. Do chronic respiratory diseases or their treatment affect the risk of SARS-CoV-2 infection? *Lancet Respir. Med.* **8**, 436–438 (2020).
 435. Shah, M., Sachdeva, M. & Dodiuk-Gad, R. P. COVID-19 and racial disparities. *J. Am. Acad. Dermatol.* **83**, e35 (2020).
 436. Hrusak, O. *et al.* Flash survey on severe acute respiratory syndrome coronavirus-2 infections in paediatric patients on anticancer treatment. *Eur. J. Cancer* **132**, 11–16 (2020).
 437. Fendler, A. *et al.* Functional antibody and T-cell immunity following SARS-CoV-2 infection, including by variants of concern, in patients with cancer: the CAPTURE study. *Res. Sq.* (2021). doi:10.21203/RS.3.RS-916427/V1
 438. Barrière, J. *et al.* Impaired immunogenicity of BNT162b2 anti-SARS-CoV-2 vaccine in patients treated for solid tumors. *Ann. Oncol.* **32**, 1053 (2021).
 439. Monin, L. *et al.* Safety and immunogenicity of one versus two doses of the COVID-19 vaccine BNT162b2 for patients with cancer: interim analysis of a prospective observational study. *Lancet. Oncol.* **22**, 765 (2021).
 440. Madelon, N. *et al.* Robust T cell responses in anti-CD20 treated patients following COVID-19 vaccination: a prospective cohort study. *Clin. Infect. Dis. An Off. Publ. Infect. Dis. Soc. Am.* **75**, E1037–E1045 (2022).
 441. Singer, J. *et al.* Article evaluation of antibody responses to covid-19 vaccines among solid tumor and hematologic patients. *Cancers (Basel)*. **13**, 4312 (2021).
 442. Villarreal-Garza, C. *et al.* Attitudes and Factors Associated With COVID-19 Vaccine Hesitancy Among Patients With Breast Cancer. *JAMA Oncol.* **7**, 1242 (2021).
 443. Ingram, S. A. *et al.* Hesitancy and malignancy: Vaccine hesitancy among individuals with cancer. https://doi.org/10.1200/JCO.2020.39.28_suppl.148 **39**, 148–148 (2021).
 444. Carle, C. *et al.* The risk of contracting SARS-CoV-2 or developing COVID-19 for people with cancer: A systematic review of the early evidence. *J. Cancer Policy* **33**, 100338 (2022).
 445. Poliseño, M. *et al.* Clinical Characteristics, Outcomes, and Risk Factors of Patients Hospitalized for COVID-19 across the Latest Pandemic Waves: Has Something Changed? *BioMed 2023, Vol. 3, Pages 272–281* **3**, 272–281 (2023).
 446. Martin, P. *et al.* COVID-19 in cancer patients: update from the joint analysis of the ESMO-CoCARE, BSMO, and PSMO international databases. *ESMO Open* **8**, 101566 (2023).
 447. Pinato, D. J. *et al.* Outcomes of the SARS-CoV-2 omicron (B.1.1.529) variant outbreak among vaccinated and unvaccinated patients with cancer in Europe: results from the retrospective, multicentre, OnCovid registry study. *Lancet. Oncol.* **23**, 865 (2022).
 448. Lee, M. *et al.* Impact of COVID-19 on case fatality rate of patients with cancer during the Omicron wave. *Cancer Cell* **40**, 343 (2022).
 449. Leiner, J. *et al.* Characteristics and outcomes of COVID-19 patients during B.1.1.529 (Omicron) dominance compared to B.1.617.2 (Delta) in 89 German hospitals. *BMC Infect. Dis.* **22**, (2022).
 450. Kawano, Y. *et al.* Temporal trends in COVID-19 outcomes among patients with systemic autoimmune rheumatic diseases: from the first wave through the initial Omicron wave. *Ann. Rheum. Dis.* **81**, 1742 (2022).
 451. Bernal, J. L. *et al.* Effectiveness of Covid-19 Vaccines against the B.1.617.2 (Delta) Variant. *N. Engl. J. Med.* **385**, 585–594 (2021).
 452. Andrews, N. *et al.* Effectiveness of COVID-19 booster vaccines against COVID-19-related symptoms, hospitalization and death in England. *Nat. Med.* **28**, 831 (2022).
 453. Andrews, N. *et al.* Duration of Protection against Mild and Severe Disease by Covid-19 Vaccines. *N. Engl. J. Med.* **386**, 340–350 (2022).
 454. Twohig, K. A. *et al.* Hospital admission and emergency care attendance risk for SARS-CoV-2 delta (B.1.617.2) compared with alpha (B.1.1.7) variants of concern: a cohort study. *Lancet Infect. Dis.* **22**, 35–

42 (2022).

455. Lee, A. J. X. & Purshouse, K. COVID-19 and cancer registries: learning from the first peak of the SARS-CoV-2 pandemic. *Br. J. Cancer* 2021 12411 **124**, 1777–1784 (2021).
456. Kuderer, N. M. *et al.* Clinical impact of COVID-19 on patients with cancer (CCC19): a cohort study. *Lancet (London, England)* **395**, 1907 (2020).
457. Home. Available at: <https://ccc19.org/index.html>. (Accessed: 24th April 2024)
458. de Joode, K. *et al.* Dutch Oncology COVID-19 consortium: Outcome of COVID-19 in patients with cancer in a nationwide cohort study. *Eur. J. Cancer* **141**, 171–184 (2020).
459. Joerger, M. *et al.* Outcome and Prognostic Factors of COVID-19 Infection in Swiss Cancer Patients: Final Results of SAKK 80/20 (CaSA). *Cancers* 2022, Vol. 14, Page 2191 **14**, 2191 (2022).
460. Lièvre, A. *et al.* Risk factors for Coronavirus Disease 2019 (COVID-19) severity and mortality among solid cancer patients and impact of the disease on anticancer treatment: A French nationwide cohort study (GCO-002 CACOV-19). *Eur. J. Cancer* **141**, 62–81 (2020).
461. Pinato, D. J. *et al.* Presenting Features and Early Mortality from SARS-CoV-2 Infection in Cancer Patients during the Initial Stage of the COVID-19 Pandemic in Europe. *Cancers* 2020, Vol. 12, Page 1841 **12**, 1841 (2020).
462. UK Coronavirus Cancer Programme – Safeguard, evaluate and protect cancer patients during the COVID-19 pandemic. Available at: <https://ukcovidcancerprogramme.org/>. (Accessed: 14th February 2023)
463. Anil, I. *et al.* The UK Coronavirus Cancer Monitoring Project: protecting patients with cancer in the era of COVID-19. *Lancet Oncol.* **21**, 622–624 (2020).
464. Harris, P. A. *et al.* Research electronic data capture (REDCap)-A metadata-driven methodology and workflow process for providing translational research informatics support. *J. Biomed. Inform.* **42**, 377–381 (2009).
465. Harris, P. A. *et al.* The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics* **95**, 103208 (2019).
466. ICD-10 Version:2019. Available at: <https://icd.who.int/browse10/2019/en>. (Accessed: 16th June 2023)
467. Institute for Health, N. & Excellence, C. COVID-19 rapid guideline: critical care in adults. (2020).
468. Cancer registration statistics, England - Office for National Statistics. Available at: <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/cancerregistrationstatisticsengland/2017>. (Accessed: 17th October 2023)
469. Yang, K. *et al.* Clinical characteristics, outcomes, and risk factors for mortality in patients with cancer and COVID-19 in Hubei, China: a multicentre, retrospective, cohort study. *Lancet Oncol.* **21**, 904–913 (2020).
470. He, W. *et al.* COVID-19 in persons with haematological cancers. *Leuk.* 2020 346 **34**, 1637–1645 (2020).
471. Tian, J. *et al.* Clinical characteristics and risk factors associated with COVID-19 disease severity in patients with cancer in Wuhan, China: a multicentre, retrospective, cohort study. *Lancet. Oncol.* **21**, 893 (2020).
472. Garassino, M. C. *et al.* COVID-19 in patients with thoracic malignancies (TERAVOLT): first results of an international, registry-based, cohort study. *Lancet. Oncol.* **21**, 914 (2020).
473. Guan, W. *et al.* Clinical characteristics of 2019 novel coronavirus infection in China. *N. Engl. J. Med.* (2020). doi:10.1101/2020.02.06.20020974
474. Gudbjartsson, D. F. *et al.* Spread of SARS-CoV-2 in the Icelandic Population. *N. Engl. J. Med.* **382**, 2302–2315 (2020).
475. Prabani, K. I. P., Weerasekara, I. & Damayanthi, H. D. W. T. COVID-19 vaccine acceptance and hesitancy among patients with cancer: a systematic review and meta-analysis. *Public Health* **212**, 66 (2022).
476. Dean, N. E., Hogan, J. W. & Schnitzer, M. E. Covid-19 Vaccine Effectiveness and the Test-Negative Design. <https://doi.org/10.1056/NEJMe2113151> **385**, 1431–1433 (2021).
477. NHS England » Landmark moment as first NHS patient receives COVID-19 vaccination. Available at:

- <https://www.england.nhs.uk/2020/12/landmark-moment-as-first-nhs-patient-receives-covid-19-vaccination/>. (Accessed: 17th October 2023)
478. NHS England » NHS begins COVID-19 booster vaccination campaign. Available at: <https://www.england.nhs.uk/2021/09/nhs-begins-covid-19-booster-vaccination-campaign/>. (Accessed: 17th October 2023)
 479. COVID-19 Second Generation Surveillance System (SGSS) - NHS Digital. Available at: <https://digital.nhs.uk/services/data-access-request-service-dars/dars-products-and-services/data-set-catalogue/covid-19-second-generation-surveillance-system-sgss>. (Accessed: 17th October 2023)
 480. National Immunisation and Vaccination System (NIVS) - NHS Digital. Available at: <https://digital.nhs.uk/coronavirus/vaccinations/training-and-onboarding/point-of-care/national-immunisation-and-vaccination-system-nivs>. (Accessed: 17th October 2023)
 481. Secondary Uses Service (SUS) - NHS Digital. Available at: <https://digital.nhs.uk/services/secondary-uses-service-sus>. (Accessed: 17th October 2023)
 482. [MI] Rapid Cancer Registration Data - NHS Digital. Available at: <https://digital.nhs.uk/data-and-information/publications/statistical/mi-rapid-cancer-registration-data/current/current>. (Accessed: 17th October 2023)
 483. [Withdrawn] COVID-19 testing data: methodology note - GOV.UK. Available at: <https://www.gov.uk/government/publications/coronavirus-covid-19-testing-data-methodology/covid-19-testing-data-methodology-note>. (Accessed: 17th October 2023)
 484. English indices of deprivation - GOV.UK. Available at: <https://www.gov.uk/government/collections/english-indices-of-deprivation>. (Accessed: 15th October 2023)
 485. List of ethnic groups - GOV.UK. Available at: <https://www.ethnicity-facts-figures.service.gov.uk/style-guide/ethnic-groups>. (Accessed: 15th October 2023)
 486. SACT Homepage. Available at: <http://www.chemodataset.nhs.uk/home>. (Accessed: 1st April 2020)
 487. Mass asymptomatic Covid-19 testing: Strategy and accuracy - House of Commons Library. Available at: <https://commonslibrary.parliament.uk/research-briefings/cbp-9223/>. (Accessed: 19th October 2023)
 488. COVID-19 variants: genomically confirmed case numbers - GOV.UK. Available at: <https://www.gov.uk/government/publications/covid-19-variants-genomically-confirmed-case-numbers>. (Accessed: 17th October 2023)
 489. Chua, H. *et al.* The use of test-negative controls to monitor vaccine effectiveness: a systematic review of methodology. *Epidemiology* **31**, 43 (2020).
 490. Cortellini, A. *et al.* COVID-19 Sequelae and the Host Proinflammatory Response: An Analysis From the OnCovid Registry. *JNCI J. Natl. Cancer Inst.* **114**, 979–987 (2022).
 491. Satyanarayana, G. *et al.* Coinfections in Patients With Cancer and COVID-19: A COVID-19 and Cancer Consortium (CCC19) Study. *Open Forum Infect. Dis.* **9**, (2022).
 492. Garcia-Gallo, E. *et al.* ISARIC-COVID-19 dataset: A Prospective, Standardized, Global Dataset of Patients Hospitalized with COVID-19. *Sci. Data* **2022** *9*, 1–22 (2022).
 493. Gong, I. Y. *et al.* Association of COVID-19 Vaccination With Breakthrough Infections and Complications in Patients With Cancer. *JAMA Oncol.* **9**, 386–394 (2023).
 494. Brazete, C. *et al.* COVID-19 vaccines effectiveness against symptomatic disease and severe outcomes, 2021–2022: a test-negative case–control study. *Public Health* **218**, 84–91 (2023).
 495. Carazo, S. *et al.* Effectiveness of previous infection-induced and vaccine-induced protection against hospitalisation due to omicron BA subvariants in older adults: a test-negative, case-control study in Quebec, Canada. *Lancet Heal. Longev.* **4**, e409–e420 (2023).
 496. Wu, J. T. Y. *et al.* Association of COVID-19 Vaccination With SARS-CoV-2 Infection in Patients With Cancer: A US Nationwide Veterans Affairs Study. *JAMA Oncol.* **8**, 1 (2022).
 497. Heudel, P. *et al.* Survival and risk of COVID-19 after SARS-COV-2 vaccination in a series of 2391 cancer patients. *Eur. J. Cancer* **165**, 174 (2022).
 498. Bernal, J. L. *et al.* Effectiveness of Covid-19 Vaccines against the B.1.617.2 (Delta) Variant. <https://doi.org/10.1056/NEJMoa2108891> **385**, 585–594 (2021).

499. Bernal, J. L. *et al.* Effectiveness of the Pfizer-BioNTech and Oxford-AstraZeneca vaccines on covid-19 related symptoms, hospital admissions, and mortality in older adults in England: test negative case-control study. *BMJ* **373**, (2021).
500. Lopes de Assis, F. *et al.* Tracking B cell responses to the SARS-CoV-2 mRNA-1273 vaccine. *Cell Rep.* **42**, 112780 (2023).
501. Khoury, D. S. *et al.* Neutralizing antibody levels are highly predictive of immune protection from symptomatic SARS-CoV-2 infection. *Nat. Med.* **2021** *277* **27**, 1205–1211 (2021).
502. Lapuente, D., Winkler, T. H. & Tenbusch, M. B-cell and antibody responses to SARS-CoV-2: infection, vaccination, and hybrid immunity. *Cell. Mol. Immunol.* **2023** *212* **21**, 144–158 (2023).
503. Moss, P. The T cell immune response against SARS-CoV-2. *Nat. Immunol.* **2022** *232* **23**, 186–193 (2022).
504. Shapiro, L. C. *et al.* Efficacy of booster doses in augmenting waning immune responses to COVID-19 vaccine in patients with cancer. *Cancer Cell* **40**, 3 (2022).
505. Ligumsky, H. *et al.* Immunogenicity and safety of BNT162b2 mRNA vaccine booster in actively treated patients with cancer. *Lancet. Oncol.* **23**, 193 (2022).
506. Prendecki, M. *et al.* Immunological responses to SARS-CoV-2 vaccines in kidney transplant recipients. *Lancet (London, England)* **398**, 1482 (2021).
507. Martin, P. *et al.* Comparison of immunogenicity and clinical effectiveness between BNT162b2 and ChAdOx1 SARS-CoV-2 vaccines in people with end-stage kidney disease receiving haemodialysis: A prospective, observational cohort study. *Lancet Reg. Heal. - Eur.* **21**, (2022).
508. Ireland, G. *et al.* Serological responses to COVID-19 Comirnaty booster vaccine, London, United Kingdom, September to December 2021. *Eurosurveillance* **27**, (2022).
509. Röltgen, K. & Boyd, S. D. Antibody and B cell responses to SARS-CoV-2 infection and vaccination. *Cell Host Microbe* **29**, 1063–1075 (2021).
510. Turner, J. S. *et al.* SARS-CoV-2 mRNA vaccines induce persistent human germinal centre responses. *Nature* **596**, 109 (2021).
511. Swadling, L. *et al.* Pre-existing polymerase-specific T cells expand in abortive seronegative SARS-CoV-2. *Nature* **601**, 110 (2022).
512. Ng, T. I. *et al.* Antiviral Drug Discovery for the Treatment of COVID-19 Infections. *Viruses* **14**, (2022).
513. Ambrose, N. *et al.* Neutralizing Monoclonal Antibody Use and COVID-19 Infection Outcomes. *JAMA Netw. Open* **6**, e239694 (2023).
514. Suribhatla, R. *et al.* Systematic review and meta-analysis of the clinical effectiveness of tixagevimab/cilgavimab for prophylaxis of COVID-19 in immunocompromised patients. *Br. J. Haematol.* **201**, 813–823 (2023).
515. Mitchell, H. *et al.* Impact of the COVID-19 pandemic on emergency hospital cancer admissions in a UK region. *BMC Cancer* **22**, 1–8 (2022).
516. Sze, S. *et al.* Ethnicity and clinical outcomes in COVID-19: A systematic review and meta-analysis. *eClinicalMedicine* **29**, 100630 (2020).
517. Albani, V., Welsh, C. E., Brown, H., Matthews, F. E. & Bamba, C. Explaining the deprivation gap in COVID-19 mortality rates: A decomposition analysis of geographical inequalities in England. *Soc. Sci. Med.* **311**, 115319 (2022).
518. Lone, N. I. *et al.* Influence of socioeconomic deprivation on interventions and outcomes for patients admitted with COVID-19 to critical care units in Scotland: A national cohort study. *Lancet Reg. Heal. - Eur.* **1**, (2021).
519. Buikema, A. R. *et al.* Racial and ethnic disparity in clinical outcomes among patients with confirmed COVID-19 infection in a large US electronic health record database. *iScience* **39**, (2021).
520. Ospina, A. V. *et al.* Impact of vaccination against COVID-19 on patients with cancer in ACHOC-C19 study: Real world evidence from one Latin American country. *J. Cancer* **14**, 2410 (2023).
521. Nelli, F. *et al.* Effects of Antibody Response after Booster Vaccination on SARS-CoV-2 Breakthrough Infections and Disease Outcomes in Advanced Cancer Patients: A Prospective Analysis of the Vax-on-

- Third Study. *Curr. Oncol.* **30**, 5103–5115 (2023).
522. Choueiri, T. K. *et al.* Breakthrough SARS-CoV-2 infections among patients with cancer following two and three doses of COVID-19 mRNA vaccines: a retrospective observational study from the COVID-19 and Cancer Consortium. *Lancet Reg. Heal. - Am.* **19**, 100445 (2023).
 523. Q-Q Plot (Quantile to Quantile Plot). *Concise Encycl. Stat.* 437–439 (2008). doi:10.1007/978-0-387-32833-1_331
 524. Ahmad, M. H., Adnan, R. & Adnan, N. A Comparative Study On Some Methods For Handling Multicollinearity Problems. *MATEMATIKA* **22**, 109–119 (2006).
 525. Daoud, J. I. Multicollinearity and Regression Analysis. *J. Phys. Conf. Ser.* **949**, 012009 (2017).
 526. Mahase, E. Covid-19: New “Pirola” variant BA.2.86 continues to spread in UK and US. *BMJ* **382**, p2097 (2023).
 527. Demetriou, P. *et al.* A dynamic CD2-rich compartment at the outer edge of the immunological synapse boosts and integrates signals. *Nat. Immunol.* **2020 2110** **21**, 1232–1243 (2020).
 528. Lee, L. Y. W. *et al.* COVID-19 prevalence and mortality in patients with cancer and the effect of primary tumour subtype and patient demographics: a prospective cohort study. *Lancet. Oncol.* **21**, 1309 (2020).
 529. Lee, L. Y. W. *et al.* Vaccine effectiveness against COVID-19 breakthrough infections in patients with cancer (UKCCEP): a population-based test-negative case-control study. *Lancet. Oncol.* **23**, 748 (2022).
 530. Lee, L. Y. W. *et al.* COVID-19: Third dose booster vaccine effectiveness against breakthrough coronavirus infection, hospitalisations and death in patients with cancer: A population-based study. *Eur. J. Cancer* **175**, 1–10 (2022).
 531. Lee, L. Y. W. *et al.* Association of SARS-CoV-2 Spike Protein Antibody Vaccine Response With Infection Severity in Patients With Cancer: A National COVID Cancer Cross-sectional Evaluation. *JAMA Oncol.* **9**, 188–196 (2023).
 532. Starkey, T. *et al.* A population-scale temporal case–control evaluation of COVID-19 disease phenotype and related outcome rates in patients with cancer in England (UKCCP). *Sci. Reports 2023 131* **13**, 1–8 (2023).
 533. Angelova, M. *et al.* Characterization of the immunophenotypes and antigenomes of colorectal cancers reveals distinct tumor escape mechanisms and novel targets for immunotherapy. *Genome Biol.* **16**, (2015).
 534. Marisa, L. *et al.* The balance between cytotoxic T-cell lymphocytes and immune checkpoint expression in the prognosis of colon tumors. *J. Natl. Cancer Inst.* **110**, 68–77 (2018).
 535. Wherry, E. J. *et al.* Molecular Signature of CD8+ T Cell Exhaustion during Chronic Viral Infection. *Immunity* **27**, 670–684 (2007).
 536. Gardner, A., de Mingo Pulido, Á. & Ruffell, B. Dendritic Cells and Their Role in Immunotherapy. *Frontiers in Immunology* **11**, 924 (2020).
 537. Felker, P. *et al.* TGF-β1 Accelerates Dendritic Cell Differentiation from Common Dendritic Cell Progenitors and Directs Subset Specification toward Conventional Dendritic Cells. *J. Immunol.* **185**, 5326–5335 (2010).
 538. Chan, C. W. *et al.* Interferon-producing killer dendritic cells provide a link between innate and adaptive immunity. *Nat. Med.* **12**, 207–213 (2006).
 539. Jin, P. *et al.* Molecular signatures of maturing dendritic cells: Implications for testing the quality of dendritic cell therapies. *J. Transl. Med.* **8**, 4 (2010).
 540. Liu, W. M., Dennis, J. L., Fowler, D. W. & Dalgleish, A. G. The gene expression profile of unstimulated dendritic cells can be used as a predictor of function. *Int. J. Cancer* **130**, 979–990 (2012).
 541. Schaefer, C. F. *et al.* PID: The pathway interaction database. *Nucleic Acids Res.* **37**, (2009).
 542. Torri, A. *et al.* Gene expression profiles identify inflammatory signatures in dendritic cells. *PLoS One* **5**, 9404 (2010).
 543. Vendelova, E. *et al.* Tolerogenic transcriptional signatures of steady-state and pathogen-induced dendritic cells. *Frontiers in Immunology* **9**, 1 (2018).

SUPPLEMENTARY MATERIAL

Experimental code

Code and scripts corresponding to my analyses are available on GitHub: <https://github.com/tstarkeyphd>. Please contact [REDACTED] with any queries or requests for further information.

Supplementary tables

Stage	Contents	Cancer Immunity Cycle Stage
TIGER 1	Cancer cell death (apoptosis)	Cancer cell death (1)
TIGER 2	MHC Class I antigen presentation MHC Class II antigen presentation	Cancer antigen presentation (2)
TIGER 3	T cell priming and activation	Priming and activation (3)
TIGER 4	Endothelial cell activation	Trafficking of T cells to tumours (4)
TIGER 5	Adaptive CD4 immune cells Adaptive CD8 immune cells Cancer cell killing immune cells	Infiltration of T cells into tumours (5)
TIGER 6	Immune checkpoint activation	Recognition of cancer cells by T cells (6)
TIGER 7	Effector CD8 v CD8 exhausted cell gene comparison Effector / Lymphoid cell activation	Killing of cancer cells (7)

Supplementary Table 1: Tumour Immunogenomic Report (TIGER) overview for each Cancer Immunity Cycle stage ¹⁹⁶.

Geneset name	Stage	Process	Genes	Source / Reference
Apoptosis	TIGER 1	Cancer cell death	161	Hallmark gene collection
Class I MHC antigen presentation	TIGER 2	MHC Class I antigen presentation	251	REACT 75820
Class II MHC antigen presentation	TIGER 2	MHC Class II antigen presentation	91	REACT 121399
Activated CD4	TIGER 3	T cell priming and activation	35	533
Activated CD8			24	
T cells			86	
Endothelial cell activation	TIGER 4	Endothelial cell activation	9	GO:0042118
Central memory CD4	TIGER 5	Infiltration of adaptive CD4 immune cells	21	533
Effector memory CD4			14	
Th1			37	
Th17			36	
Treg			37	
Central memory CD8	TIGER 5	Infiltration of adaptive CD8 immune cells	39	533
Cytotoxic cells			9	
Effector memory CD8			42	
Gamma delta T cells	TIGER 5	Infiltration of cancer cell killing immune cells	43	533
Natural killer T cells			12	
Immune Checkpoint	TIGER 6	Immune checkpoint activation	19	534
Effector v Exhausted CD8 T cells	TIGER 7	Effector CD8 v CD8 exhausted cell	200	⁵³⁵ (GSE9650 1256 200 UP)
NK cell mediated cytotoxicity	TIGER 7	Effector / Lymphoid cell activation	137	hsa04650
+ Dendritic cells (DC)	TIGER 2	"Professional" antigen presenting cells (cancer antigen presentation)	34	533
+ Macrophages			22	
+ Activated B cells			18	
+ Memory B cells			14	
+ Eosinophils	TIGER 5	Innate immune Cells (Infiltration of T cells into tumours)	15	533
+ Mast cells			42	
+ Neutrophils			16	
+ Natural killer cells (NK)			17	
+ Macrophage activation	TIGER 7	Effector / Lymphoid cell activation (Killing of cancer cells)	31	GO:0042116

Supplementary Table 2: TIGER stages and constituent genesets. + denotes that genesets are not used when calculating expression of Cancer Immunity Cycle stages.

Process	Geneset name	Genes	Source / Reference
Negative regulation of the immune system (Neg reg)	Negative regulation of adaptive immune response	45	GO:0002820
	Negative regulation of antigen presentation	8	GO:0002578
	Negative regulation of cytokine production	25	GO:0002719
	Negative regulation of cytokine secretion	6	GO:0002740
	Negative regulation of immune effector process	118	GO:0002698
	Negative regulation of immune response	142	GO:0050777
	Negative regulation of innate immune response	54	GO:0045824
	Negative regulation of molecular mediator production	36	GO:0002701
Positive regulation of the immune system (Pos reg)	Positive regulation of adaptive immune response	105	GO:0002821
	Positive regulation of antigen presentation	9	GO:0002579
	Positive regulation of cytokine production	53	GO:0002720
	Positive regulation of cytokine secretion	8	GO:0002741
	Positive regulation of immune effector process	215	GO:0002699
	Positive regulation of immune response	888	GO:0050778
	Positive regulation of molecular mediator production	96	GO:0002702
Activation of the immune system	Activation of immune response	708	GO:0002253
	Activation of innate immune response	315	GO:0002218
	Activation of lymphocytes	176	GO:0002285

Supplementary Table 3: TIGER immune regulation and activation genesets. These genesets are not used when calculating expression of Cancer Immunity Cycle stages.

Cancer Immunity Cycle (CIC)		Death (OS Event) (n/N, %)		Median time to death, days (IQR)	
Constituent signature	Stage	Expression High	Expression Low	Expression High	Expression Low
Apoptosis	1	60/320 (18.8%)	73/318 (23.0%)	433.0 (185.0-1095.25)	472.0 (206.0-1162.0)
MHC I antigen presentation	2	63/318 (19.8%)	70/320 (21.9%)	424.0 (252.0-884.0)	508.0 (173.75-1240.5)
MHC II antigen presentation	2	63/320 (19.7%)	70/318 (22.0%)	405.0 (183.5-845.5)	511.5 (211.5-1240.5)
Activated CD4	3	60/319 (18.8%)	73/319 (22.9%)	434.5 (172.0-1111.5)	472.0 (206.0-1096.0)
Activated CD8	3	67/319 (21.0%)	66/319 (20.7%)	442.0 (175.5-1096.0)	460.5 (208.0-1131.75)
T cells	3	60/320 (18.8%)	73/318 (23.0%)	399.5 (171.75-774.0)	512.0 (228.0-1314.0)
Endothelial activation	4	52/318 (16.4%)	81/320 (25.3%)	396.0 (145.25-833.75)	512.0 (242.0-1185.0)
Central memory CD4	5	61/319 (19.1%)	72/319 (22.6%)	403.0 (214.0-869.0)	530.5 (177.25-1205.5)
Effector memory CD4	5	52/319 (16.3%)	81/319 (25.4%)	604.0 (157.5-1307.25)	424.0 (214.0-869.0)
Th1	5	67/319 (21.0%)	66/319 (20.7%)	424.0 (201.0-884.0)	508.0 (185.75-1240.5)
Th17	5	57/319 (17.9%)	76/319 (23.8%)	424.0 (188.0-1095.0)	507.5 (205.25-1167.75)
Treg	5	60/320 (18.8%)	73/318 (23.0%)	424.0 (283.0-833.75)	504.0 (156.0-1305.0)
Central memory CD8	5	60/320 (18.8%)	73/318 (23.0%)	399.5 (235.0-833.75)	512.0 (172.0-1258.0)
Cytotoxic cells	5	57/320 (17.8%)	76/318 (23.9%)	368.0 (159.0-682.0)	567.5 (234.0-1318.25)
Effector memory CD8	5	62/320 (19.4%)	71/318 (22.3%)	396.0 (179.0-806.0)	549.0 (217.0-1286.0)
Gamma Delta T Cells	5	63/319 (19.7%)	70/319 (21.9%)	424.0 (182.0-884.0)	500.0 (211.5-1187.25)
Natural Killer T Cells	5	63/320 (19.7%)	70/318 (22.0%)	396.0 (210.0-935.5)	511.5 (173.75-1162.75)
Immune checkpoint	6	60/319 (18.8%)	73/319 (22.9%)	396.0 (185.0-1002.75)	504.0 (206.0-1188.0)
Effector v exhausted CD8 T cell	7	72/319 (22.6%)	61/319 (19.1%)	424.0 (207.5-809.25)	512.0 (172.0-1348.0)
NK cell cytotoxicity	7	64/320 (20.0%)	69/318 (21.7%)	382.0 (158.25-701.0)	610.0 (236.0-1314.0)

Supplementary Table 4: Deaths (overall survival/OS events) and time to death for each Cancer Immunity Cycle (CIC) constituent gene signature. Signature expression bifurcated at the median and classified as being “high” or “low” for each CRC patient within TCGA COADREAD.

Geneset	TML - High expression (n=77)	TML - Low exp- ression (n=416)	TML - Difference	p value*
TIGER 1	661.1	640.7	20.4	NS
TIGER 2	804.8	763.0	41.8	NS
TIGER 3	523.6	484.2	39.4	NS
TIGER 4	344.4	329.4	15.0	NS
TIGER 5	285.3	233.1	52.2	6.16E-07
TIGER 6	85.9	45.9	39.9	7.44E-09
TIGER 7	442.5	396.8	45.8	1.73E-03
Apoptosis	661.1	640.7	20.4	NS
MHC I antigen presentation	671.3	676.4	-5.1	NS
MHC II antigen presentation	938.3	849.7	88.6	0.036
Activated CD4	643.4	486.3	157.1	5.36E-10
Activated CD8	876.7	932.4	-55.7	NS
T cells	50.8	34.0	16.8	0.003
Endothelial activation	344.4	329.4	15.0	NS
Central memory CD4	280.6	210.2	70.4	0.001
Effector memory CD4	1439.3	1349.3	89.9	NS
Th1	204.2	143.2	60.9	0.003
Th17	27.5	27.1	0.4	NS
Treg	49.9	47.5	2.4	NS
Central memory CD8	131.1	114.6	16.5	NS
Cytotoxic cells	206.7	140.1	66.6	9.99E-04
Effector memory CD8	210.7	117.5	93.2	1.34E-10
Gamma Delta T Cells	274.4	174.8	99.6	6.16E-05
Natural Killer T Cells	28.7	6.8	21.8	4.11E-14
Immune checkpoint	85.9	45.9	39.9	2.12E-08
Effector v exhausted CD8 T cell	725.6	682.4	43.1	NS
NK cell cytotoxicity	159.5	111.1	48.4	1.57E-07
Activated B cells	42.9	37.4	5.5	NS
Dendritic cells	150.3	96.4	54.0	4.32E-05
Macrophages	32.3	33.5	-1.3	NS
Memory B cells	58.2	64.8	-6.7	NS
Eosinophil	117.4	122.4	-5.0	NS
Mast cells	32.4	30.0	2.3	NS
Neutrophils	116.8	71.0	45.8	8.81E-04
Natural killer cells	69.8	53.5	16.3	0.021
Macrophage activation	96.0	76.8	19.3	NS
Neg reg adaptive immune response	83.5	78.2	5.3	NS
Neg reg antigen presentation	413.1	287.4	125.8	0.001
Neg reg cytokine production	53.4	41.4	11.9	NS
Neg reg cytokine secretion	19.7	18.5	1.2	NS
Neg reg immune effector process	128.7	112.1	16.5	NS
Neg reg immune response	89.1	88.3	0.8	NS
Neg reg innate immune response	297.4	248.9	48.5	NS
Neg reg molecular mediator production	61.0	53.2	7.8	NS

Geneset	TML - High expression	TML - Low expression	TML - Difference	p value*
Pos reg adaptive immune response	144.7	134.3	10.4	NS
Pos reg antigen presentation	224.7	200.9	23.8	NS
Pos reg cytokine production	293.2	228.9	64.2	0.004
Pos reg cytokine secretion	953.6	985.7	-32.1	NS
Pos reg immune effector process	122.0	113.2	8.8	NS
Pos reg immune response	144.3	129.7	14.5	NS
Pos reg molecular mediator production	153.8	135.4	18.4	NS
Immune response activation	162.2	145.3	16.9	NS
Innate immune response activation	275.7	232.5	43.3	0.001
Lymphocyte activation	136.1	125.1	11.1	NS

Supplementary Table 5: Association of TIGER geneset expression with tumour mutational load (TML).

*Wilcoxon-rank sum tests were performed. *p values corrected by applying a Bonferroni coefficient.*

xCell Gene signature	Group	Parent/Child
CLP (Common lymphoid progenitors)	HSC	HSC
CMP (Common myeloid progenitors)	HSC	HSC
Erythrocytes	HSC	HSC
GMP (Granulocyte-macrophage progenitors)	HSC	HSC
MEP (Megakaryocyte-erythroid progenitors)	HSC	HSC
Megakaryocytes	HSC	HSC
HSC (Hematopoietic stem cells)	HSC	Parent
MPP (Multipotent progenitors)	HSC	Parent
Class-switched memory B-cells	Lymphoid	B-cells
Memory B-cells	Lymphoid	B-cells
naive B-cells	Lymphoid	B-cells
Plasma cells	Lymphoid	B-cells
pro B-cells	Lymphoid	B-cells
CD4+ memory T-cells	Lymphoid	CD4+ memory T-cells
CD4+ Tcm (CD4+ central memory T-cells)	Lymphoid	CD4+ T-cells
CD4+ Tem (CD4+ effector memory T-cells)	Lymphoid	CD4+ T-cells
CD4+ naive T-cells	Lymphoid	CD4+ T-cells
Tgd cells (Gamma delta T-cells)	Lymphoid	CD4+ T-cells
Tregs (Regulatory T-cells)	Lymphoid	CD4+ T-cells
Th1 cells (Type 1 T-helper cells)	Lymphoid	CD4+ T-cells
Th2 cells (Type 2 T-helper cells)	Lymphoid	CD4+ T-cells
CD8+ Tcm (CD8+ central memory T-cells)	Lymphoid	CD8+ T-cells
CD8+ Tem (CD8+ effector memory T-cells)	Lymphoid	CD8+ T-cells
CD8+ naive T-cells	Lymphoid	CD8+ T-cells
B-cells	Lymphoid	Parent
CD4+ T-cells	Lymphoid	Parent
CD8+ T-cells	Lymphoid	Parent
NKT (Natural killer T-cells)	Lymphoid	Parent
NK cells (Natural Killer cells)	Lymphoid	Parent
Neurons	Epithelial	Astrocytes
Keratinocytes	Epithelial	Epithelial cells
Sebocytes	Epithelial	Epithelial cells
Astrocytes	Epithelial	Parent
Epithelial cells	Epithelial	Parent
Hepatocytes	Epithelial	Parent
Melanocytes	Epithelial	Parent

xCell Gene signature	Group	Parent/Child
Preadipocytes	Stroma	Adipocytes
ly Endothelial cells (Lymphatic endothelial cells)	Stroma	Endothelial cells
mv Endothelial cells (Microvascular endothelial cells)	Stroma	Endothelial cells
Mesangial cells	Stroma	F broblasts
Pericytes	Stroma	F broblasts
Adipocytes	Stroma	Parent
Chondrocytes	Stroma	Parent
Endothelial cells	Stroma	Parent
Fibroblasts	Stroma	Parent
MSC (Mesenchymal stem cells)	Stroma	Parent
Skeletal muscle	Stroma	Parent
Smooth muscle	Stroma	Parent
Myocytes	Stroma	Skeletal muscle
Osteoblast	Stroma	Skeletal muscle
Platelets	HSC	Parent
aDC (Activated dendritic cells)	Myeloid	DC
iDC (Immature dendritic cells)	Myeloid	DC
pDC (Plasmacytoid dendritic cells)	Myeloid	DC
cDC (Conventional dendritic cells)	Myeloid	DC
Macrophages M1	Myeloid	Macrophages
Macrophages M2	Myeloid	Macrophages
Basophils	Myeloid	Parent
DC	Myeloid	Parent
Eosinophils	Myeloid	Parent
Macrophages	Myeloid	Parent
Mast cells	Myeloid	Parent
Monocytes	Myeloid	Parent
Neutrophils	Myeloid	Parent

Supplementary Table 6: xCell gene signatures from the xCell R package²⁴¹.

Gene Signature	Source / Reference
Co-ordinate Immune Response Cluster (CIRC)	218
CTLA4 PATHWAY	BIOCARTA CTLA4
PD 1 SIGNALING	R-HSA-389948
TGF BETA SIGNALING	Hallmark gene collection
XVEGF PATHWAY	BIOCARTA VEGF
DENDRITIC CELL ANTIGEN PROCESSING AND PRESENTATION	GO:0002468
DENDRITIC CELL APOPTOTIC PROCESS	GO:0097048
DENDRITIC CELL CYTOKINE PRODUCTION	GO:0002371
DENDRITIC CELL DIFFERENTIATION	GO:0097028
DENDRITIC CELL MIGRATION	GO:0036336
XYELOID DENDRITIC CELL ACTIVATION	GO:0001773
CLASS I MHC MEDIATED ANTIGEN PROCESSING PRESENTATION	R-HSA-983169
MHC CLASS II ANTIGEN PRESENTATION	R-HSA-2132295
NEGATIVE REGULATION OF ANTIGEN PROCESSING AND PRESENTATION	GO:0002578
POSITIVE REGULATION OF ANTIGEN PROCESSING AND PRESENTATION	GO:0002579
ANTIGEN PRESENTATION: FOLDING ASSEMBLY AND PEPTIDE LOADING OF CLASS I MHC	R-HSA-983170
ANTIGEN PROCESSING AND PRESENTATION ENDOGENOUS LIPID	GO:0048006
ANTIGEN VIA MHC CLASS IB	GO:0042590
ANTIGEN PROCESSING AND PRESENTATION OF EXOGENOUS PEPTIDE	GO:0042590
ANTIGEN VIA MHC CLASS I	GO:0002480
ANTIGEN PROCESSING AND PRESENTATION OF EXOGENOUS PEPTIDE	GO:0002480
ANTIGEN VIA MHC CLASS I TAP INDEPENDENT	GO:0002475
ANTIGEN PROCESSING AND PRESENTATION VIA MHC CLASS IB	GO:0002475
MHC CLASS I BIOSYNTHETIC PROCESS	GO:0045341
MHC CLASS I PEPTIDE LOADING COMPLEX	GO:0042824
MHC CLASS I PROTEIN BINDING	GO:0042288
MHC CLASS I PROTEIN COMPLEX	GO:0042612
MHC CLASS I PROTEIN COMPLEX BINDING	GO:0023024
MHC CLASS I RECEPTOR ACTIVITY	GO:0032393
MHC CLASS IB PROTEIN BINDING	GO:0023029
MHC CLASS IB RECEPTOR ACTIVITY	GO:0032394
NEF MEDIATED DOWNREGULATION OF MHC CLASS I COMPLEX CELL SURFACE EXPRESSION	R-HSA-164940
ANTIGEN PROCESSING AND PRESENTATION OF PEPTIDE OR POLYSACCHARIDE ANTIGEN VIA MHC CLASS II	GO:0002504
MHC CLASS II BIOSYNTHETIC PROCESS	GO:0045342
MHC CLASS II PROTEIN BINDING	GO:0042289
MHC CLASS II PROTEIN COMPLEX	GO:0042613
MHC CLASS II PROTEIN COMPLEX BINDING	GO:0023026
MHC CLASS II RECEPTOR ACTIVITY	GO:0032395
IFNG PATHWAY	BIOCARTA IFNG
IL12 PATHWAY	BIOCARTA IL12
CDC1	260
CDC2	260
CDC2A	260

Gene Signature	Source / Reference
CDC2B	260
CDC1 Interaction Inhibition	536
CDC1 Interaction Stimulation	536
CDC1 Marker Inhibition	536
CDC1 Marker Stimulation	536
NF-KAPPAB SIGNALING	GO:0038061
CONVENTIONAL CDC VS PLASMACYTOID PDC UP	537 (GSE22432 3612 200 UP)
CONVENTIONAL VS PLASMACYTOID DC SPLEEN UP	538 (GSE3691 2657 200 UP)
IFN PRODUCING KILLER DC VS CONVENTIONAL DC SPLEEN UP	538 (GSE3691 2655 200 UP)
DC Maturation Early	539
DC Maturation Late	539
DC Maturation Th1	539
DC Surface Expression	539
Bad DC	540
Good DC	540
NF-KAPPAB ATYPICAL PATHWAY	541
NF-KAPPAB CANONICAL PATHWAY	541
Inflammatory DC	542
ssmDC	543
CDC1 PRR	261
CDC1 SM	261
CDC2 PRR	261
CDC2 SM	261
DC Cancer Targets	261
DC Cancer Targets Inhibition	261
DC T cell Induction Negative	261
DC T cell Induction Positive	261
MoDC SM	261
pDC PRR	261
pDC SM	261
ALLOGRAFT REJECTION	Hallmark gene collection
IL6 JAK STAT3 SIGNALING	Hallmark gene collection
INFLAMMATORY RESPONSE	Hallmark gene collection
INTERFERON GAMMA RESPONSE	Hallmark gene collection
NOTCH SIGNALING	Hallmark gene collection
TNFA SIGNALING VIA NFKB	Hallmark gene collection
WNT BETA CATENIN SIGNALING	Hallmark gene collection

Supplementary Table 7: Immune cell and pathway gene signatures analysed in addition to TIGER and xCell for the distant metastasis results section.

Geneset	PC1	PC2
Endothelial activation	0.353	0.455
Activated CD4	0.289	-0.115
TGD (gamma-delta T)	0.255	0.156
Immune Checkpoint	0.252	0.095
NKT (natural killer T)	0.245	-0.233
Memory B cells	0.244	-0.087
Mast cells	0.220	-0.048
Effector memory CD8	0.216	0.090
Treg	0.204	-0.019
Central memory CD8	0.201	0.012
Macrophage activation	0.178	0.002
Activated CD8	0.174	0.059
Macrophages	0.174	-0.176
Th17	0.162	-0.671
MHC Class I antigen presentation	0.145	0.075
NK cytotoxicity	0.143	-0.006
Cytotoxic cells	0.137	0.201
DC	0.133	-0.135
Central memory CD4	0.130	0.134
T cells	0.129	-0.014
Th1	0.127	0.071
Eosinophil	0.121	0.053
Effector memory CD4	0.120	-0.150
Effector v Exhausted CD8	0.118	0.000
Apoptosis	0.112	0.015
NK (Natural Killer cells)	0.103	-0.030
MHC Class II antigen presentation	0.099	-0.084
Activated B cells	0.089	0.190
Neutrophils	0.077	-0.020

Supplementary Table 8: Cancer Immunity Cycle (TIGER) determinants of LN stromal variation in stage 2 CRC patients by principal component analysis (PCA) for PCA axes 1 and 2.

Geneset	Aggressive primary	Non-aggressive primary	Expression difference	Adjusted p value
TIGER 1	10.74	10.50	0.24	6.56E-09
TIGER 2	10.76	10.49	0.27	1.03E-10
TIGER 3	10.53	10.15	0.38	9.21E-08
TIGER 4	10.18	9.56	0.62	3.19E-05
TIGER 5	10.64	10.30	0.34	3.48E-08
TIGER 6	10.60	10.03	0.57	8.04E-08
TIGER 7	10.70	10.40	0.29	3.60E-11
Apoptosis	10.74	10.50	0.24	1.87E-08
MHC Class I antigen presentation	10.78	10.42	0.36	6.89E-12
MHC Class II antigen presentation	10.75	10.57	0.18	6.05E-05
Activated CD4	10.27	9.74	0.53	3.80E-05
Activated CD8	10.67	10.32	0.35	1.63E-06
T cells	10.66	10.40	0.26	1.85E-05
Endothelial activation	10.18	9.56	0.62	9.11E-05
Central memory CD4	10.89	10.51	0.38	3.12E-09
Effector memory CD4	10.75	10.45	0.30	1.61E-03
Th1	10.81	10.48	0.33	6.08E-11
Th17	10.32	10.26	0.05	NS
Treg	10.52	10.15	0.37	1.69E-04
Central memory CD8	10.51	10.09	0.42	5.08E-05
Cytotoxic cells	10.86	10.65	0.22	1.95E-02
Effector memory CD8	10.56	10.13	0.43	5.06E-07
TGD	10.60	10.01	0.59	7.72E-10
NKT	10.56	10.23	0.33	NS
Immune Checkpoint	10.60	10.03	0.57	2.30E-07
Effector v exhausted CD8	10.75	10.48	0.27	3.51E-11
NK cytotoxicity	10.65	10.33	0.32	1.94E-08
Activated B cells	10.90	10.60	0.30	8.95E-07
Dendritic cells	10.45	10.25	0.20	3.98E-02
Macrophages	10.63	10.35	0.29	4.23E-03
Memory B cells	10.39	9.97	0.42	7.30E-04
Eosinophil	10.67	10.36	0.31	3.52E-04
Mast cells	10.50	10.01	0.48	9.31E-08
Neutrophils	10.84	10.68	0.16	NS
Natural killer cells	10.74	10.56	0.18	NS
Macrophage activation	10.53	10.20	0.33	2.59E-05

Supplementary Table 9: Cancer Immunity Cycle expression changes in normal lymph nodes based on aggressiveness (proliferative capacity) of paired CRC primary tumours. Wilcoxon-rank sum tests were performed, and p values corrected by applying a Bonferroni coefficient.

Gene	baseMean	log2FoldChange	lfcSE	stat	pvalue	padj
OXSRI	94.861	2.356	0.155	-15.239	1.95E-52	3.84E-48
PAF1	68.061	1.905	0.145	-13.112	2.82E-39	2.77E-35
TRIO	100.756	1.520	0.120	-12.629	1.46E-36	9.57E-33
ABCC1	42.933	2.397	0.211	-11.363	6.37E-30	3.13E-26
POLR2A	208.747	1.229	0.108	-11.325	9.90E-30	3.89E-26
GDAP1	20.971	2.632	0.234	-11.229	2.92E-29	9.58E-26
PITPNC1	32.315	2.475	0.224	-11.059	1.98E-28	5.55E-25
VPS72	104.124	1.706	0.154	-11.046	2.29E-28	5.64E-25
TSG101	40.197	1.642	0.149	-11.005	3.63E-28	7.93E-25
HNRNPAB	97.861	1.726	0.160	-10.808	3.16E-27	6.21E-24
STK39	65.329	1.264	0.118	-10.688	1.16E-26	2.06E-23
C1orf63	178.866	1.153	0.110	-10.526	6.54E-26	1.07E-22
DBN1	28.460	1.869	0.182	-10.278	8.86E-25	1.13E-21
TATDN3	33.030	2.867	0.279	-10.274	9.22E-25	1.13E-21
KMT2D	58.204	1.780	0.173	-10.268	9.84E-25	1.14E-21
LRRRC8A	35.562	1.856	0.181	-10.243	1.27E-24	1.39E-21
PACSLN2	69.097	1.465	0.144	-10.183	2.36E-24	2.44E-21
KIAA0020	122.140	1.013	0.100	-10.093	5.96E-24	5.85E-21
WI2-81516E3.1	3.521	4.590	0.456	-10.067	7.71E-24	7.21E-21
IGKV1-8	3.878	4.375	0.438	-9.983	1.82E-23	1.62E-20
C19orf43	201.343	1.473	0.148	-9.966	2.15E-23	1.80E-20
PNP	14.722	3.261	0.327	-9.964	2.20E-23	1.80E-20
DOCK9	38.646	1.716	0.172	-9.953	2.45E-23	1.93E-20
PYCR1	17.473	4.074	0.411	-9.907	3.88E-23	2.93E-20
SIPA1L3	16.380	2.368	0.241	-9.807	1.05E-22	7.62E-20
RTN4IP1	15.608	2.383	0.244	-9.757	1.72E-22	1.21E-19
STX18	17.394	1.984	0.204	-9.727	2.30E-22	1.56E-19
HLA-C	273.094	1.265	0.130	-9.716	2.57E-22	1.69E-19
HERC2P3	5.571	3.638	0.375	-9.710	2.75E-22	1.74E-19
HNRNPD	252.460	0.701	0.072	-9.675	3.86E-22	2.30E-19
C21orf2	3.636	3.790	0.392	-9.676	3.82E-22	2.30E-19
TAS1R3	33.414	4.258	0.442	-9.635	5.68E-22	3.28E-19
RN7SL1	4028.712	2.317	0.242	-9.591	8.75E-22	4.91E-19
SH3GL1P3	26.969	4.159	0.434	-9.576	1.01E-21	5.50E-19
RN7SL674P	25.493	3.354	0.352	-9.527	1.62E-21	8.40E-19
RNA5SP429	21.758	3.547	0.372	-9.527	1.62E-21	8.40E-19
TOMM5	9.385	2.509	0.265	-9.466	2.92E-21	1.44E-18
RNA5S9	1.655	3.848	0.407	-9.465	2.93E-21	1.44E-18
DLD	61.206	0.969	0.103	-9.429	4.13E-21	1.98E-18
TNFRSF25	8.747	3.782	0.401	-9.423	4.37E-21	2.04E-18
IGF2	117.620	2.643	0.282	-9.377	6.77E-21	3.10E-18
NONO	172.017	0.697	0.075	-9.274	1.79E-20	7.64E-18
TNS3	60.182	1.231	0.133	-9.258	2.09E-20	8.74E-18
DFFA	73.894	1.226	0.133	-9.184	4.14E-20	1.69E-17
SLC35B2	13.201	1.946	0.212	-9.181	4.27E-20	1.71E-17
TLE1	71.864	1.408	0.154	-9.165	4.94E-20	1.94E-17
SRRM2	932.241	0.969	0.106	-9.141	6.18E-20	2.38E-17
TCTEX1D4	3.055	3.785	0.415	-9.123	7.33E-20	2.77E-17
FOXO4	4.959	3.927	0.433	-9.074	1.14E-19	4.24E-17
ZDHHC20-IT1	8.082	3.912	0.433	-9.037	1.61E-19	5.86E-17
EMC7	29.246	1.707	0.189	-9.019	1.90E-19	6.77E-17
GJA6P	1.825	3.834	0.426	-9.007	2.11E-19	7.42E-17
GPR124	15.672	3.865	0.431	-8.959	3.27E-19	1.11E-16
ATP5G2P3	4.053	3.546	0.396	-8.947	3.64E-19	1.21E-16
RRP12	36.143	1.703	0.191	-8.941	3.86E-19	1.26E-16
PLXND1	35.108	1.291	0.145	-8.909	5.15E-19	1.66E-16
ARHGEF25	18.405	1.722	0.194	-8.872	7.18E-19	2.27E-16
TCTN2	16.806	3.178	0.358	-8.870	7.33E-19	2.29E-16
EFCC1	21.444	3.140	0.355	-8.852	8.61E-19	2.64E-16
ANO2	19.048	3.601	0.408	-8.835	1.01E-18	3.04E-16
SNHG15	22.534	1.693	0.193	-8.788	1.52E-18	4.52E-16
FAM175A	10.247	2.276	0.260	-8.767	1.83E-18	5.30E-16
FKBP8	3.363	2.292	0.261	-8.769	1.81E-18	5.30E-16
ZZEF1	28.817	1.408	0.162	-8.715	2.91E-18	8.28E-16
VPS9D1-AS1	15.503	2.251	0.260	-8.672	4.24E-18	1.17E-15
NCF1	15.067	3.340	0.386	-8.663	4.59E-18	1.25E-15
BCS1L	26.389	1.855	0.215	-8.624	6.45E-18	1.74E-15
U2AF2	35.103	1.374	0.159	-8.621	6.62E-18	1.76E-15
IFT140	6.063	2.590	0.302	-8.585	9.11E-18	2.39E-15
HIST1H3H	6.769	3.366	0.392	-8.578	9.65E-18	2.50E-15
CTD-2007H13.3	8.565	3.359	0.392	-8.575	9.89E-18	2.52E-15
P4HB	74.930	1.406	0.165	-8.534	1.41E-17	3.42E-15
CLDND1	10.055	1.806	0.212	-8.537	1.38E-17	3.42E-15
RAC2	24.888	3.521	0.413	-8.534	1.41E-17	3.42E-15
PIGG	23.367	1.817	0.214	-8.481	2.24E-17	5.29E-15
APOL6	84.494	0.898	0.106	-8.460	2.66E-17	6.23E-15
POLR2D	16.788	1.700	0.201	-8.435	3.32E-17	7.67E-15
MIR663A	38.108	2.826	0.335	-8.430	3.44E-17	7.87E-15
SGSH	15.596	2.129	0.253	-8.413	4.00E-17	9.05E-15
PFKFB3	52.480	1.143	0.136	-8.406	4.25E-17	9.49E-15
TENM3	7.192	1.994	0.237	-8.404	4.31E-17	9.52E-15
NEURL4	11.115	3.090	0.368	-8.391	4.80E-17	1.05E-14
RN7SL5P	3.642	2.826	0.338	-8.365	6.02E-17	1.30E-14
DCAF13P3	2.386	3.378	0.404	-8.361	6.22E-17	1.33E-14
DNM2	141.579	0.998	0.119	-8.353	6.67E-17	1.41E-14

VARs	45.700	1.237	0.149	-8.311	9.51E-17	1.99E-14
CGN	31.893	1.255	0.152	-8.283	1.20E-16	2.46E-14
DDX56	105.932	1.271	0.154	-8.270	1.34E-16	2.72E-14
UNC5CL	5.292	2.277	0.276	-8.250	1.59E-16	3.18E-14
RNA5SP202	187.843	2.802	0.340	-8.248	1.61E-16	3.19E-14
FAM219B	20.076	1.936	0.235	-8.236	1.78E-16	3.49E-14
FOXJ3	24.524	1.106	0.134	-8.224	1.96E-16	3.81E-14
PLEC	254.101	1.062	0.129	-8.221	2.02E-16	3.89E-14
LENG8	25.433	2.240	0.273	-8.214	2.15E-16	4.09E-14
DPP9	41.568	1.328	0.162	-8.207	2.27E-16	4.24E-14
C9orf3	23.527	1.450	0.177	-8.207	2.26E-16	4.24E-14
RBM12B-AS1	5.934	2.630	0.321	-8.200	2.40E-16	4.45E-14
VIP	1.862	-2.727	0.334	8.175	2.97E-16	5.45E-14
CDRT15P1	2.699	3.366	0.414	-8.140	3.97E-16	7.09E-14
MINK1	21.421	2.168	0.267	-8.121	4.64E-16	8.21E-14
GPR132	39.694	2.260	0.279	-8.093	5.83E-16	1.01E-13
MYH9	341.667	0.633	0.078	-8.082	6.35E-16	1.10E-13
PRKRIR	18.191	1.429	0.177	-8.061	7.56E-16	1.29E-13
WHSC1	111.260	0.622	0.077	-8.059	7.67E-16	1.30E-13
CTSB	454.691	0.797	0.099	-8.048	8.43E-16	1.42E-13
VAV2	11.357	2.693	0.335	-8.030	9.72E-16	1.62E-13
RN7SL4P	82.093	2.403	0.300	-8.020	1.06E-15	1.73E-13
HSD11B1	3.633	2.832	0.353	-8.013	1.12E-15	1.82E-13
RNA5SP141	11.427	3.161	0.395	-8.012	1.13E-15	1.82E-13
RN7SL54P	5.839	3.055	0.382	-7.992	1.32E-15	2.11E-13
XAB2	5.897	2.702	0.339	-7.979	1.47E-15	2.33E-13
CXCL17	2.680	2.892	0.363	-7.961	1.71E-15	2.69E-13
ADAMTSL4	15.571	2.317	0.291	-7.959	1.74E-15	2.71E-13
ZDHHC5	31.425	1.089	0.137	-7.954	1.80E-15	2.76E-13
TRAPPC11	11.257	1.446	0.182	-7.955	1.79E-15	2.76E-13
THBS1	63.139	0.915	0.115	-7.932	2.16E-15	3.29E-13
ACTA2-AS1	2.888	3.116	0.393	-7.929	2.21E-15	3.35E-13
HIST1H4B	19.523	2.774	0.350	-7.924	2.30E-15	3.46E-13
DROSHA	77.088	0.998	0.126	-7.915	2.48E-15	3.70E-13
ZNF507	35.436	1.152	0.146	-7.901	2.77E-15	4.10E-13
PRPF39	26.455	1.145	0.145	-7.875	3.40E-15	4.99E-13
GDF7	11.295	2.950	0.375	-7.864	3.73E-15	5.43E-13
CTB-88F18.3	1.912	3.540	0.451	-7.855	3.99E-15	5.76E-13
RNA5SP357	7.272	3.173	0.405	-7.833	4.77E-15	6.84E-13
RN7SL128P	4.599	2.906	0.373	-7.795	6.42E-15	9.14E-13
SRP68	34.038	1.245	0.160	-7.779	7.31E-15	1.03E-12
RN7SL396P	17.263	3.051	0.393	-7.770	7.86E-15	1.10E-12
SEL1L3	159.815	0.860	0.111	-7.762	8.39E-15	1.17E-12
C1orf51	3.947	2.550	0.329	-7.747	9.41E-15	1.29E-12
FTO	9.462	2.191	0.283	-7.741	9.85E-15	1.34E-12
ZNF286A	7.097	2.308	0.299	-7.729	1.09E-14	1.47E-12
ADAM23	14.543	2.375	0.308	-7.718	1.18E-14	1.58E-12
YIPF2	53.179	0.803	0.104	-7.709	1.27E-14	1.68E-12
SERINC5	56.886	1.436	0.186	-7.705	1.30E-14	1.71E-12
PLBD2	12.251	2.758	0.358	-7.702	1.33E-14	1.74E-12
SNORA74A	18.436	2.948	0.383	-7.688	1.49E-14	1.93E-12
ADCY3	26.781	1.340	0.176	-7.630	2.35E-14	3.01E-12
HMHA1	65.682	1.541	0.202	-7.623	2.48E-14	3.16E-12
PRPS1P2	1.669	3.441	0.453	-7.602	2.92E-14	3.70E-12
RN7SKP255	2.246	3.482	0.459	-7.583	3.37E-14	4.24E-12
C5orf66	6.231	2.434	0.321	-7.580	3.46E-14	4.33E-12
ZGPAT	3.799	2.909	0.384	-7.576	3.57E-14	4.44E-12
WNK4	9.655	2.100	0.277	-7.571	3.70E-14	4.57E-12
AGO2	32.465	1.082	0.143	-7.552	4.29E-14	5.26E-12
HAVCR1	14.920	2.894	0.385	-7.520	5.47E-14	6.67E-12
UBALD2	59.167	0.714	0.095	-7.501	6.34E-14	7.69E-12
RNA5SP161	8.499	2.782	0.371	-7.491	6.81E-14	8.21E-12
VPS37B	65.844	1.325	0.177	-7.485	7.18E-14	8.55E-12
PREX1	34.729	1.250	0.167	-7.472	7.87E-14	9.32E-12
MIR941-3	3.123	3.127	0.419	-7.463	8.45E-14	9.89E-12
DNAAF1	8.239	2.943	0.395	-7.441	9.97E-14	1.16E-11
TRAF7	44.367	1.101	0.148	-7.435	1.04E-13	1.21E-11
CNNM2	10.294	1.720	0.232	-7.416	1.20E-13	1.38E-11
DOCK9-AS1	2.244	3.020	0.408	-7.407	1.29E-13	1.47E-11
HIST1H2AG	4.851	2.336	0.315	-7.405	1.31E-13	1.48E-11
ABO	2.513	3.127	0.423	-7.393	1.43E-13	1.60E-11
LINC00338	26.143	1.938	0.263	-7.366	1.76E-13	1.94E-11
SRCRB4D	4.213	2.536	0.346	-7.335	2.22E-13	2.43E-11
ZNF592	24.008	1.363	0.186	-7.330	2.30E-13	2.50E-11
FLJ00104	4.697	2.570	0.352	-7.294	3.02E-13	3.26E-11
HERC2P2	3.959	1.949	0.267	-7.291	3.07E-13	3.28E-11
HIST1H2AE	23.323	2.250	0.309	-7.292	3.06E-13	3.28E-11
HIST3H2A	3.059	2.849	0.391	-7.286	3.18E-13	3.38E-11
TRMT61B	4.402	2.372	0.326	-7.283	3.26E-13	3.44E-11
KLF13	34.275	0.820	0.113	-7.281	3.30E-13	3.47E-11
ZEB2	56.584	1.053	0.145	-7.278	3.38E-13	3.53E-11
KIAA1456	7.256	1.905	0.262	-7.264	3.75E-13	3.90E-11
TMEM259	52.502	1.372	0.189	-7.258	3.94E-13	4.07E-11
SRCAP	43.290	0.701	0.097	-7.243	4.38E-13	4.51E-11
ACAA1	11.366	2.650	0.367	-7.230	4.83E-13	4.94E-11
LRRN2	15.560	2.456	0.341	-7.212	5.51E-13	5.58E-11

CLSTN2	10.148	2.218	0.308	-7.197	6.16E-13	6.21E-11
IK	116.812	0.860	0.120	-7.191	6.42E-13	6.44E-11
KCNC4	22.193	1.762	0.245	-7.183	6.84E-13	6.82E-11
UNC119B	6.467	2.030	0.283	-7.173	7.36E-13	7.27E-11
PSMC1P10	14.368	1.415	0.198	-7.141	9.25E-13	9.09E-11
ZBTB45	18.689	1.007	0.142	-7.102	1.23E-12	1.20E-10
RN7SK	18.997	2.086	0.294	-7.086	1.38E-12	1.34E-10
ARAP1	15.579	2.284	0.322	-7.086	1.38E-12	1.34E-10
RASGEF1C	1.534	2.786	0.394	-7.076	1.49E-12	1.43E-10
FAM160B2	20.596	1.640	0.232	-7.069	1.56E-12	1.50E-10
SNORA7B	108.057	2.919	0.413	-7.064	1.62E-12	1.55E-10
RNA28S5	3790.216	1.147	0.162	-7.061	1.65E-12	1.56E-10
IGHV4-34	2.847	2.864	0.406	-7.061	1.65E-12	1.56E-10
USP32	23.603	1.397	0.198	-7.054	1.74E-12	1.63E-10
DIP2A	86.269	1.002	0.142	-7.043	1.88E-12	1.75E-10
PRICKLE2-AS3	4.990	2.211	0.315	-7.022	2.19E-12	2.03E-10
PRKRIRP7	5.736	2.285	0.326	-7.019	2.23E-12	2.06E-10
CEP97	15.052	1.238	0.178	-6.970	3.17E-12	2.91E-10
C2orf72	15.381	1.418	0.204	-6.956	3.50E-12	3.20E-10
HOMER2P1	5.495	2.663	0.383	-6.953	3.59E-12	3.26E-10
SIK1	11.670	1.683	0.242	-6.948	3.70E-12	3.35E-10
HSCB	8.777	1.705	0.246	-6.939	3.95E-12	3.56E-10
RN7SL767P	11.088	2.484	0.358	-6.935	4.08E-12	3.64E-10
CCDC124	26.919	1.325	0.191	-6.926	4.34E-12	3.86E-10
MDN1	102.683	0.669	0.097	-6.900	5.20E-12	4.60E-10
SEP1	33.892	2.160	0.313	-6.897	5.31E-12	4.68E-10
MAP7D1	70.632	0.933	0.135	-6.889	5.62E-12	4.93E-10
ERN1	32.696	1.157	0.168	-6.886	5.74E-12	5.02E-10
MLLT1	3.725	2.602	0.378	-6.879	6.03E-12	5.24E-10
ARMC8	16.891	0.987	0.144	-6.852	7.29E-12	6.31E-10
PHF13	11.328	1.256	0.184	-6.843	7.73E-12	6.66E-10
TAPBPL	19.223	1.735	0.254	-6.827	8.69E-12	7.45E-10
HIP1R	19.521	1.379	0.202	-6.817	9.28E-12	7.93E-10
CASC7	46.897	0.709	0.104	-6.811	9.66E-12	8.22E-10
CHD2	245.688	0.584	0.086	-6.809	9.85E-12	8.35E-10
CTD-2287O16.5	11.236	1.740	0.256	-6.804	1.02E-11	8.58E-10
ZNF70	7.496	2.133	0.314	-6.803	1.02E-11	8.58E-10
MIR497HG	1.445	3.150	0.463	-6.803	1.03E-11	8.58E-10
CLASP2	53.832	1.027	0.151	-6.794	1.09E-11	9.11E-10
MARS	45.130	0.897	0.133	-6.771	1.28E-11	1.06E-09
SRCIN1	5.121	2.253	0.333	-6.764	1.34E-11	1.11E-09
TSR3	7.778	2.014	0.298	-6.758	1.40E-11	1.15E-09
ARFGEF2	73.467	0.950	0.141	-6.756	1.42E-11	1.15E-09
PTGFRN	22.513	0.993	0.147	-6.756	1.42E-11	1.15E-09
PRR12	2.778	2.184	0.323	-6.756	1.42E-11	1.15E-09
DHODH	17.320	1.471	0.219	-6.734	1.65E-11	1.33E-09
LRRC41	8.710	1.472	0.219	-6.720	1.82E-11	1.47E-09
AP2S1	208.384	0.993	0.148	-6.711	1.93E-11	1.55E-09
ACIN1	102.553	0.808	0.120	-6.708	1.97E-11	1.57E-09
ABR	33.271	1.285	0.192	-6.697	2.12E-11	1.69E-09
BPTF	270.598	0.601	0.090	-6.693	2.18E-11	1.73E-09
MON1B	133.481	0.675	0.101	-6.680	2.39E-11	1.89E-09
CHTF18	21.164	1.493	0.223	-6.678	2.42E-11	1.90E-09
EMILIN1	49.781	1.256	0.188	-6.665	2.65E-11	2.07E-09
KIAA0513	18.414	1.551	0.233	-6.661	2.73E-11	2.13E-09
SMCO1	1.492	2.834	0.426	-6.660	2.74E-11	2.13E-09
C1RL-AS1	62.870	1.244	0.187	-6.648	2.97E-11	2.29E-09
CSF2RB	12.670	2.056	0.309	-6.648	2.97E-11	2.29E-09
TCERG1	85.220	0.536	0.081	-6.633	3.29E-11	2.53E-09
RRP8	25.827	1.002	0.151	-6.624	3.49E-11	2.67E-09
USP17L7	1.295	2.652	0.400	-6.621	3.58E-11	2.72E-09
RNA5SP370	4.400	2.576	0.389	-6.615	3.71E-11	2.82E-09
DIS3	84.810	0.739	0.112	-6.607	3.93E-11	2.97E-09
SNED1	5.303	1.837	0.278	-6.598	4.18E-11	3.15E-09
GPT	9.192	1.498	0.227	-6.593	4.30E-11	3.23E-09
EEF1GP5	5.839	1.925	0.292	-6.592	4.34E-11	3.24E-09
ACHE	14.292	1.557	0.236	-6.585	4.56E-11	3.39E-09
EIF4EBP3	3.401	2.631	0.400	-6.580	4.71E-11	3.49E-09
ARPC5L	59.760	0.875	0.133	-6.576	4.85E-11	3.58E-09
NCEH1	18.636	0.910	0.138	-6.575	4.87E-11	3.59E-09
SMAD3	66.804	0.938	0.143	-6.571	4.98E-11	3.65E-09
CAD	83.792	0.851	0.130	-6.555	5.57E-11	4.07E-09
CBR4	106.119	0.939	0.143	-6.552	5.68E-11	4.12E-09
RELL2	3.024	2.470	0.377	-6.552	5.66E-11	4.12E-09
OLFM4	82.243	-1.196	0.183	6.546	5.92E-11	4.28E-09
LYPD6B	11.700	2.102	0.321	-6.543	6.05E-11	4.35E-09
SNORD15B	14.275	2.329	0.356	-6.538	6.24E-11	4.48E-09
LETM1	61.606	0.719	0.110	-6.536	6.31E-11	4.51E-09
PLIN3	1.310	2.730	0.418	-6.532	6.48E-11	4.62E-09
CNN1	55.279	-1.062	0.163	6.531	6.54E-11	4.64E-09
SHC1	17.203	1.635	0.251	-6.524	6.85E-11	4.84E-09
AIMP2	3.839	2.089	0.321	-6.515	7.28E-11	5.13E-09
INPP5E	7.009	1.639	0.252	-6.504	7.81E-11	5.44E-09
ZSCAN23	1.175	2.786	0.428	-6.504	7.80E-11	5.44E-09
NTN1	3.925	-1.824	0.281	6.500	8.01E-11	5.56E-09
EIF4E2	110.365	0.820	0.126	-6.484	8.90E-11	6.14E-09

CAMKK1	9.733	1.363	0.210	-6.477	9.33E-11	6.41E-09
PKD2	23.452	1.183	0.183	-6.464	1.02E-10	6.98E-09
ZNF207	275.469	0.552	0.086	-6.451	1.11E-10	7.59E-09
FBXW7	167.597	0.811	0.126	-6.436	1.22E-10	8.31E-09
CNPY4	14.693	1.281	0.199	-6.424	1.33E-10	8.95E-09
DNAJC11	9.281	1.699	0.265	-6.409	1.47E-10	9.88E-09
GRB2	119.622	0.864	0.135	-6.405	1.51E-10	1.01E-08
CASC4	48.640	0.810	0.127	-6.402	1.53E-10	1.02E-08
TYW5	19.211	1.013	0.158	-6.400	1.55E-10	1.03E-08
C1orf43	50.879	0.701	0.110	-6.398	1.58E-10	1.05E-08
INSR	80.418	0.646	0.101	-6.375	1.83E-10	1.21E-08
ZNF609	186.405	1.036	0.163	-6.375	1.83E-10	1.21E-08
PRDM10	12.123	1.566	0.246	-6.359	2.02E-10	1.33E-08
DDX19A	42.725	0.681	0.107	-6.352	2.12E-10	1.39E-08
RGMA	9.975	-1.380	0.217	6.352	2.13E-10	1.39E-08
LYAR	39.826	0.955	0.150	-6.349	2.16E-10	1.41E-08
RECQL4	40.963	0.766	0.121	-6.342	2.27E-10	1.48E-08
KIAA0430	129.402	0.685	0.108	-6.341	2.29E-10	1.48E-08
CSK	124.706	1.205	0.190	-6.331	2.44E-10	1.57E-08
TMPRSS11CP	5.391	2.424	0.383	-6.330	2.45E-10	1.57E-08
SNORD89	120.339	1.411	0.223	-6.320	2.61E-10	1.67E-08
FBXL14	4.312	2.186	0.347	-6.302	2.93E-10	1.87E-08
USP38	10.083	1.184	0.188	-6.298	3.01E-10	1.92E-08
PANK4	5.743	1.827	0.290	-6.291	3.16E-10	1.99E-08
RNF187	226.508	0.660	0.105	-6.288	3.22E-10	2.01E-08
CHST7	4.528	1.759	0.280	-6.280	3.40E-10	2.12E-08
SCARNA4	5.761	1.986	0.316	-6.279	3.41E-10	2.12E-08
POM121	17.415	1.850	0.295	-6.277	3.44E-10	2.13E-08
ESPNL	7.674	2.094	0.334	-6.277	3.45E-10	2.13E-08
ZNF224	72.279	1.092	0.174	-6.274	3.52E-10	2.17E-08
PKDCC	11.094	1.079	0.172	-6.272	3.56E-10	2.19E-08
DHX38	28.323	1.058	0.169	-6.268	3.67E-10	2.24E-08
YIPF4	147.551	0.603	0.096	-6.265	3.73E-10	2.27E-08
PLEKHG3	52.599	0.797	0.127	-6.265	3.73E-10	2.27E-08
ZNF491	3.140	2.371	0.378	-6.264	3.75E-10	2.27E-08
UNKL	22.695	1.347	0.215	-6.259	3.88E-10	2.33E-08
CTD-3232M19.2	0.987	2.861	0.457	-6.259	3.87E-10	2.33E-08
FAUP1	113.340	1.020	0.163	-6.254	4.01E-10	2.40E-08
FLII	45.260	0.975	0.156	-6.248	4.15E-10	2.48E-08
AGAP2-AS1	2.889	2.318	0.371	-6.247	4.18E-10	2.49E-08
TRMT10B	36.797	0.878	0.141	-6.231	4.63E-10	2.74E-08
PNN	274.839	0.449	0.072	-6.225	4.82E-10	2.83E-08
IGLV1-47	2.242	2.460	0.395	-6.224	4.84E-10	2.84E-08
SNX29	45.028	1.452	0.233	-6.222	4.92E-10	2.87E-08
GNAS	103.612	0.603	0.097	-6.219	5.01E-10	2.91E-08
ABCF3	16.817	0.904	0.146	-6.213	5.19E-10	3.01E-08
WNT5A	7.887	-1.204	0.194	6.203	5.53E-10	3.20E-08
SNRPA1	34.987	1.036	0.167	-6.197	5.77E-10	3.32E-08
TMEM63A	159.224	0.851	0.137	-6.192	5.94E-10	3.41E-08
COL18A1-AS2	1.161	2.867	0.463	-6.186	6.17E-10	3.53E-08
RAB1F	11.220	1.262	0.204	-6.183	6.27E-10	3.58E-08
KDM1B	10.194	1.107	0.179	-6.173	6.71E-10	3.82E-08
LONP2	176.221	0.802	0.130	-6.158	7.38E-10	4.19E-08
TEP1	22.481	0.975	0.158	-6.156	7.44E-10	4.21E-08
COG6	7.643	1.170	0.190	-6.154	7.56E-10	4.25E-08
DDB1	107.415	0.582	0.095	-6.147	7.91E-10	4.44E-08
MED29	35.244	0.814	0.132	-6.144	8.03E-10	4.49E-08
SAMD4A	17.473	1.155	0.188	-6.143	8.11E-10	4.53E-08
ZBED4	20.766	1.040	0.170	-6.126	9.00E-10	5.00E-08
SCARNA21	70.367	2.095	0.342	-6.126	9.01E-10	5.00E-08
C20orf166-AS1	1.991	2.826	0.462	-6.123	9.19E-10	5.09E-08
SH3RF3-AS1	4.747	2.212	0.361	-6.120	9.36E-10	5.17E-08
SPTBN1	367.776	0.320	0.052	-6.117	9.52E-10	5.24E-08
CLSTN1	56.982	0.633	0.104	-6.104	1.03E-09	5.66E-08
DACT1	4.184	1.871	0.307	-6.102	1.05E-09	5.73E-08
DNAJC5	138.659	0.517	0.085	-6.097	1.08E-09	5.89E-08
TACC1	436.185	0.696	0.114	-6.096	1.09E-09	5.92E-08
HSPB7	6.038	-1.669	0.274	6.094	1.10E-09	5.99E-08
TNC	27.784	1.247	0.205	-6.083	1.18E-09	6.39E-08
VCP	154.234	0.399	0.066	-6.062	1.35E-09	7.28E-08
CPSF3L	21.726	1.126	0.186	-6.054	1.41E-09	7.61E-08
ARMCX1	5.368	1.796	0.297	-6.039	1.55E-09	8.31E-08
TKTL1	8.432	2.089	0.346	-6.035	1.59E-09	8.50E-08
CD72	13.083	2.241	0.372	-6.024	1.70E-09	9.07E-08
RN7SL2	607.324	1.718	0.286	-6.013	1.82E-09	9.65E-08
ANP32B	412.147	0.345	0.057	-6.009	1.87E-09	9.84E-08
SWAP70	67.641	0.834	0.139	-6.009	1.86E-09	9.84E-08
COG4	43.597	0.947	0.158	-6.005	1.91E-09	1.01E-07
NDUFA10	90.681	0.515	0.086	-6.002	1.94E-09	1.02E-07
ABI3BP	12.350	1.331	0.222	-5.996	2.03E-09	1.06E-07
SNORA71D	1.199	2.744	0.458	-5.990	2.10E-09	1.09E-07
ADAM15	174.290	0.531	0.089	-5.969	2.39E-09	1.24E-07
IBA57	19.859	1.223	0.205	-5.966	2.43E-09	1.26E-07
DHTKD1	92.528	0.712	0.119	-5.965	2.44E-09	1.26E-07
MAMSTR	3.893	1.779	0.298	-5.960	2.52E-09	1.29E-07
ALDH8A1	3.820	1.928	0.324	-5.959	2.54E-09	1.30E-07

HPS6	12.006	1.036	0.174	-5.945	2.77E-09	1.42E-07
BCCIP	84.919	0.791	0.133	-5.943	2.80E-09	1.43E-07
RNA5SP145	6.906	2.777	0.468	-5.934	2.96E-09	1.51E-07
STAT3	172.915	0.707	0.119	-5.932	3.00E-09	1.52E-07
HNRNPA1P2	2.359	2.276	0.384	-5.922	3.17E-09	1.61E-07
LRP1	69.768	0.700	0.118	-5.919	3.23E-09	1.63E-07
SH3BGR3	7.377	1.470	0.248	-5.920	3.23E-09	1.63E-07
EEF1A1P9	1.609	2.029	0.343	-5.913	3.36E-09	1.69E-07
PTPRA	37.683	0.792	0.134	-5.909	3.44E-09	1.73E-07
TP53I11	13.717	1.235	0.209	-5.908	3.45E-09	1.73E-07
RBM23	53.178	0.875	0.148	-5.907	3.48E-09	1.73E-07
DHDDS	74.679	1.016	0.172	-5.907	3.49E-09	1.73E-07
TCHP	59.993	1.066	0.181	-5.897	3.70E-09	1.83E-07
ZNF710	17.402	1.117	0.189	-5.893	3.80E-09	1.88E-07
TIMP3	177.377	-0.584	0.099	5.891	3.84E-09	1.89E-07
HUWE1	134.386	0.350	0.060	-5.888	3.91E-09	1.92E-07
STAMPB	36.001	0.788	0.134	-5.882	4.05E-09	1.98E-07
B4GALT5	59.534	0.614	0.105	-5.877	4.18E-09	2.04E-07
MAT2A	6.050	1.497	0.255	-5.873	4.27E-09	2.08E-07
BTBD19	10.298	1.205	0.205	-5.871	4.34E-09	2.11E-07
BCL2L2	10.224	1.280	0.218	-5.865	4.50E-09	2.18E-07
ZNF770	6.497	1.377	0.235	-5.864	4.51E-09	2.18E-07
SFI1	34.249	1.151	0.196	-5.858	4.68E-09	2.26E-07
TNK2	113.015	0.980	0.167	-5.854	4.80E-09	2.31E-07
CNDP2	89.027	0.735	0.126	-5.851	4.88E-09	2.35E-07
MAD1L1	23.557	1.260	0.215	-5.849	4.95E-09	2.37E-07
TNXB	1.673	2.803	0.479	-5.848	4.97E-09	2.38E-07
DOCK4	8.382	1.229	0.210	-5.843	5.12E-09	2.44E-07
LTBP3	155.471	0.736	0.126	-5.836	5.35E-09	2.54E-07
R3HDM4	57.321	0.586	0.100	-5.834	5.41E-09	2.56E-07
C16orf59	3.355	1.779	0.305	-5.825	5.71E-09	2.70E-07
BX322557.10	4.141	1.492	0.256	-5.819	5.93E-09	2.79E-07
ARHGDIA	2.942	2.205	0.379	-5.816	6.02E-09	2.83E-07
IDH3A	14.346	1.352	0.232	-5.816	6.04E-09	2.83E-07
SNORD8	49.762	2.046	0.353	-5.803	6.51E-09	3.04E-07
ACTL6A	12.729	1.163	0.200	-5.803	6.52E-09	3.04E-07
GAPDH	7.641	1.454	0.251	-5.800	6.64E-09	3.09E-07
ALPL	2.701	1.682	0.290	-5.795	6.85E-09	3.18E-07
MORC2	10.880	1.039	0.180	-5.786	7.21E-09	3.34E-07
SSBP2	10.946	1.215	0.210	-5.783	7.34E-09	3.39E-07
TRAM1	24.639	0.892	0.154	-5.779	7.53E-09	3.47E-07
RBM19	30.484	0.868	0.150	-5.774	7.74E-09	3.56E-07
BRD1	15.443	0.831	0.144	-5.756	8.62E-09	3.95E-07
CTB-89H12.4	39.661	0.780	0.136	-5.753	8.74E-09	4.00E-07
RMRP	28.275	1.692	0.294	-5.748	9.02E-09	4.10E-07
OR1X5P	2.881	2.117	0.368	-5.748	9.05E-09	4.11E-07
TNFAIP8L1	54.808	0.819	0.143	-5.746	9.15E-09	4.14E-07
HNRNPKP2	7.146	1.891	0.329	-5.743	9.33E-09	4.21E-07
MKI67	139.799	0.599	0.104	-5.735	9.76E-09	4.40E-07
SP1	53.746	0.717	0.125	-5.730	1.00E-08	4.50E-07
LA16c-358B7.3	31.570	1.165	0.203	-5.730	1.00E-08	4.50E-07
ZNF791	95.949	0.911	0.159	-5.726	1.03E-08	4.61E-07
PTMA	2882.915	0.368	0.064	-5.720	1.07E-08	4.75E-07
EXOC7	31.645	1.054	0.185	-5.712	1.12E-08	4.97E-07
VPS18	46.118	1.009	0.177	-5.699	1.21E-08	5.34E-07
MMP11	43.523	-0.946	0.166	5.690	1.27E-08	5.60E-07
C16orf91	4.789	1.473	0.259	-5.689	1.28E-08	5.63E-07
HIST1H2BC	14.126	1.529	0.269	-5.689	1.28E-08	5.63E-07
COX5BP1	2.531	2.109	0.371	-5.681	1.34E-08	5.88E-07
MYBPC2	4.106	1.978	0.348	-5.680	1.35E-08	5.90E-07
SERPINE1	33.950	0.923	0.163	-5.677	1.37E-08	6.00E-07
NFIX	136.194	0.541	0.096	-5.661	1.50E-08	6.55E-07
TNNI1	5.357	1.743	0.308	-5.660	1.52E-08	6.59E-07
MCM4	76.898	0.555	0.098	-5.651	1.59E-08	6.91E-07
LDB1	5.541	1.295	0.229	-5.645	1.66E-08	7.17E-07
PLEKHA2	78.265	0.522	0.093	-5.643	1.67E-08	7.20E-07
ARHGAP35	38.328	0.570	0.101	-5.627	1.83E-08	7.90E-07
HDAC7	401.058	0.464	0.082	-5.624	1.87E-08	8.03E-07
LRRK1	17.004	1.129	0.201	-5.616	1.95E-08	8.37E-07
TYRO3P	1.178	2.421	0.432	-5.602	2.12E-08	9.05E-07
TUFM	135.946	0.612	0.109	-5.597	2.18E-08	9.26E-07
MGP	208.725	0.860	0.154	-5.597	2.18E-08	9.26E-07
DDIT4	8.146	1.793	0.320	-5.596	2.19E-08	9.28E-07
CALR	140.426	0.738	0.132	-5.596	2.20E-08	9.28E-07
POLR3E	50.709	0.817	0.146	-5.595	2.20E-08	9.29E-07
P2RY6	5.612	1.538	0.275	-5.591	2.25E-08	9.49E-07
PALB2	6.041	1.129	0.202	-5.585	2.33E-08	9.80E-07
RAB10	71.867	0.492	0.088	-5.576	2.46E-08	1.03E-06
C2CD2	52.033	0.775	0.139	-5.576	2.46E-08	1.03E-06
PCSK7	362.873	0.759	0.136	-5.573	2.51E-08	1.05E-06
BTNL2	1.659	2.283	0.410	-5.568	2.57E-08	1.07E-06
ATF6B	23.651	0.998	0.179	-5.565	2.63E-08	1.09E-06
HIST1H2AJ	3.419	1.663	0.299	-5.553	2.81E-08	1.16E-06
RAC1P2	190.011	0.902	0.163	-5.546	2.93E-08	1.21E-06
TANK	12.234	0.994	0.179	-5.542	2.99E-08	1.23E-06
CALU	268.266	0.519	0.094	-5.540	3.03E-08	1.25E-06

CAPN15	60.862	0.776	0.140	-5.533	3.14E-08	1.29E-06
MEF2C	114.193	0.997	0.180	-5.527	3.26E-08	1.34E-06
TECR	28.590	1.173	0.213	-5.515	3.48E-08	1.42E-06
EIF4H	19.631	1.042	0.189	-5.507	3.64E-08	1.49E-06
POGZ	81.558	0.508	0.092	-5.501	3.78E-08	1.54E-06
TNFSF10	66.242	0.695	0.126	-5.500	3.79E-08	1.54E-06
HIST1H2BD	27.344	1.060	0.193	-5.498	3.84E-08	1.56E-06
TMEM151B	0.951	2.517	0.458	-5.490	4.01E-08	1.63E-06
SCARNA2	26.546	1.482	0.270	-5.483	4.17E-08	1.69E-06
IGLC2	14.303	1.410	0.258	-5.474	4.40E-08	1.78E-06
SMG1	114.006	0.603	0.110	-5.468	4.56E-08	1.84E-06
PCDHGA12	1.676	2.194	0.402	-5.462	4.71E-08	1.89E-06
SAT2	53.349	0.607	0.111	-5.461	4.74E-08	1.90E-06
RBPMS2	2.379	-1.837	0.337	5.455	4.90E-08	1.96E-06
CRAMP1L	19.621	0.709	0.130	-5.454	4.92E-08	1.96E-06
PUF60	94.269	0.677	0.124	-5.448	5.08E-08	2.02E-06
TRAK1	55.848	0.528	0.097	-5.442	5.27E-08	2.08E-06
ARRB1	30.925	0.959	0.176	-5.442	5.27E-08	2.08E-06
C14orf2	412.843	0.436	0.080	-5.437	5.42E-08	2.14E-06
CRK	58.453	0.596	0.110	-5.433	5.54E-08	2.18E-06
MAP3K6	11.498	1.181	0.218	-5.424	5.82E-08	2.29E-06
OPA3	42.345	0.774	0.143	-5.423	5.86E-08	2.30E-06
POMT1	16.525	0.987	0.182	-5.410	6.30E-08	2.47E-06
TMEM104	12.632	1.231	0.228	-5.405	6.48E-08	2.53E-06
ABCE1	46.104	0.576	0.107	-5.380	7.43E-08	2.89E-06
TSR2	44.091	0.564	0.105	-5.374	7.71E-08	3.00E-06
LIMD1	45.832	0.937	0.174	-5.371	7.82E-08	3.03E-06
ANKRD29	39.614	1.215	0.226	-5.370	7.85E-08	3.04E-06
CTNBN1	268.362	0.349	0.065	-5.366	8.03E-08	3.10E-06
TGM2	34.771	0.730	0.136	-5.362	8.24E-08	3.18E-06
BOD1L1	126.058	0.474	0.088	-5.359	8.36E-08	3.21E-06
LRPAP1	37.960	0.693	0.129	-5.356	8.52E-08	3.27E-06
C21orf33	191.446	0.778	0.145	-5.351	8.76E-08	3.36E-06
SFRP1	6.461	1.519	0.284	-5.349	8.83E-08	3.37E-06
C1orf132	41.232	1.015	0.190	-5.349	8.84E-08	3.38E-06
CBWD5	4.117	1.372	0.257	-5.345	9.07E-08	3.45E-06
WDR13	58.775	0.755	0.141	-5.338	9.38E-08	3.57E-06
OR1X1P	2.962	2.024	0.380	-5.326	1.00E-07	3.81E-06
TRMT2B	24.916	0.760	0.143	-5.323	1.02E-07	3.87E-06
EID2	5.356	-1.136	0.213	5.322	1.02E-07	3.87E-06
POLR1B	27.719	0.865	0.163	-5.321	1.03E-07	3.88E-06
CTD-2328D6.1	16.266	1.514	0.285	-5.320	1.04E-07	3.89E-06
PPP1R32	4.425	1.920	0.361	-5.320	1.04E-07	3.89E-06
OR55B1P	0.861	2.717	0.511	-5.320	1.04E-07	3.89E-06
RBM15	26.559	0.746	0.140	-5.316	1.06E-07	3.96E-06
KLHL24	53.569	0.550	0.103	-5.316	1.06E-07	3.96E-06
NME6	20.019	0.977	0.184	-5.311	1.09E-07	4.07E-06
SFMBT2	27.746	1.355	0.255	-5.307	1.11E-07	4.14E-06
RFXANK	1.634	1.890	0.357	-5.296	1.18E-07	4.39E-06
STC2	21.821	0.948	0.180	-5.283	1.27E-07	4.71E-06
TMEM72	1.469	2.060	0.390	-5.281	1.29E-07	4.75E-06
DNHD1	36.855	1.478	0.280	-5.279	1.30E-07	4.79E-06
ANAPC16	142.361	0.666	0.126	-5.278	1.31E-07	4.81E-06
NACC2	18.747	0.907	0.172	-5.277	1.31E-07	4.81E-06
GALNT18	5.941	1.249	0.237	-5.278	1.31E-07	4.81E-06
IGLC3	193.344	1.200	0.228	-5.272	1.35E-07	4.94E-06
PTPRJ	13.024	1.264	0.240	-5.271	1.36E-07	4.96E-06
UXS1	83.451	0.495	0.094	-5.269	1.37E-07	5.00E-06
GTF2A2	28.622	0.673	0.128	-5.258	1.46E-07	5.31E-06
FBXO21	38.606	0.632	0.120	-5.253	1.50E-07	5.43E-06
KIAA0100	73.353	0.526	0.100	-5.238	1.62E-07	5.88E-06
LLNLR-246C6.1	16.846	1.308	0.250	-5.236	1.64E-07	5.93E-06
TCFL5	8.591	1.153	0.220	-5.234	1.66E-07	5.97E-06
BTN2A1	12.166	0.863	0.165	-5.230	1.70E-07	6.09E-06
NT5C2	11.187	1.312	0.251	-5.218	1.81E-07	6.46E-06
IST1	103.152	0.372	0.071	-5.211	1.88E-07	6.69E-06
SPTAN1	77.510	0.596	0.114	-5.203	1.96E-07	6.97E-06
ADPGK	69.616	0.704	0.135	-5.203	1.96E-07	6.97E-06
C9orf139	3.340	1.698	0.326	-5.202	1.97E-07	7.00E-06
ERVW-1	6.400	1.539	0.296	-5.197	2.03E-07	7.18E-06
PMEL	2.161	1.551	0.299	-5.195	2.04E-07	7.22E-06
IGHA2	38.897	1.596	0.307	-5.190	2.10E-07	7.40E-06
HNRNPA3P5	2.149	1.743	0.336	-5.190	2.10E-07	7.40E-06
HNRNPCP1	3.330	1.447	0.279	-5.189	2.12E-07	7.45E-06
L3MBTL4	9.394	1.478	0.285	-5.186	2.15E-07	7.55E-06
ARGFXP2	0.926	2.419	0.467	-5.185	2.16E-07	7.58E-06
TSTA3	71.589	0.821	0.159	-5.174	2.29E-07	8.01E-06
METTL8	67.110	0.778	0.151	-5.171	2.32E-07	8.08E-06
PRKXP1	1.063	2.435	0.471	-5.170	2.34E-07	8.12E-06
XIAP	81.847	0.476	0.092	-5.170	2.34E-07	8.13E-06
TLN2	42.887	0.709	0.137	-5.165	2.41E-07	8.32E-06
PDIA2	2.144	2.054	0.398	-5.158	2.50E-07	8.61E-06
FBXW5	311.183	0.680	0.132	-5.154	2.54E-07	8.76E-06
TSPAN14	112.140	0.509	0.099	-5.154	2.56E-07	8.78E-06
NFKB2	48.569	1.268	0.246	-5.150	2.60E-07	8.93E-06
MACF1	318.808	0.451	0.088	-5.150	2.61E-07	8.93E-06

DDX6	43.143	0.650	0.126	-5.142	2.72E-07	9.31E-06
MAB21L2	9.145	-1.174	0.228	5.139	2.77E-07	9.44E-06
ACTG2	15.842	-0.998	0.195	5.127	2.94E-07	1.00E-05
PDPN	29.463	1.009	0.197	-5.127	2.94E-07	1.00E-05
PPP1R15B	66.135	0.696	0.136	-5.121	3.04E-07	1.03E-05
AGPAT3	85.454	0.465	0.091	-5.113	3.17E-07	1.07E-05
OTUD6B	21.833	0.779	0.152	-5.112	3.19E-07	1.08E-05
PIPSL	2.818	1.784	0.349	-5.109	3.23E-07	1.09E-05
FLCN	40.457	0.684	0.134	-5.107	3.28E-07	1.10E-05
MIR27A	3.298	1.805	0.354	-5.106	3.29E-07	1.10E-05
SERPINA13P	1.039	2.472	0.484	-5.104	3.33E-07	1.11E-05
FAM196B	5.202	1.713	0.336	-5.097	3.45E-07	1.15E-05
PPP6R1	5.530	1.211	0.238	-5.096	3.46E-07	1.15E-05
ASAH1	79.221	0.530	0.104	-5.093	3.53E-07	1.18E-05
BRCA1	47.545	0.670	0.132	-5.091	3.55E-07	1.18E-05
CPNE6	1.343	2.105	0.413	-5.091	3.56E-07	1.18E-05
ANKFY1	49.776	0.839	0.165	-5.090	3.57E-07	1.18E-05
TSSC1	13.949	0.752	0.148	-5.089	3.60E-07	1.19E-05
NBR2	12.337	1.039	0.204	-5.085	3.67E-07	1.21E-05
CCDC57	49.238	0.867	0.170	-5.082	3.73E-07	1.23E-05
PRKCE	37.454	1.010	0.199	-5.080	3.77E-07	1.24E-05
SP3	26.631	0.671	0.132	-5.066	4.06E-07	1.33E-05
PVRL2	46.153	0.559	0.110	-5.066	4.07E-07	1.34E-05
GNB1	159.926	0.373	0.074	-5.063	4.12E-07	1.35E-05
MMACHC	9.962	1.161	0.229	-5.059	4.21E-07	1.38E-05
LRP5L	9.807	1.171	0.232	-5.053	4.34E-07	1.42E-05
TBC1D3P5	0.857	2.661	0.528	-5.038	4.70E-07	1.53E-05
ANP32A	62.012	0.463	0.092	-5.038	4.71E-07	1.53E-05
FAM32A	76.914	0.481	0.095	-5.036	4.74E-07	1.54E-05
GSR	76.036	0.514	0.102	-5.030	4.89E-07	1.59E-05
SLC12A7	49.487	0.643	0.128	-5.030	4.90E-07	1.59E-05
PPP2R5E	100.945	0.456	0.091	-5.023	5.08E-07	1.64E-05
DENND4C	76.278	0.629	0.125	-5.020	5.16E-07	1.67E-05
SUGP2	48.572	0.911	0.182	-5.016	5.26E-07	1.70E-05
NME1-NME2	12.454	1.086	0.217	-5.011	5.40E-07	1.74E-05
TNIP2	35.507	0.821	0.164	-5.004	5.61E-07	1.80E-05
CNTNAP2	11.634	1.098	0.220	-5.001	5.69E-07	1.82E-05
MLX	17.382	0.844	0.169	-4.995	5.89E-07	1.88E-05
KCTD17	40.502	0.748	0.150	-4.993	5.95E-07	1.90E-05
HAUS3	22.414	0.860	0.172	-4.989	6.06E-07	1.93E-05
RNA5SP242	0.719	2.684	0.538	-4.988	6.11E-07	1.94E-05
C22orf29	22.596	0.950	0.191	-4.986	6.17E-07	1.96E-05
CCDC91	29.712	0.635	0.128	-4.983	6.25E-07	1.98E-05
POLG	13.506	1.065	0.214	-4.975	6.51E-07	2.06E-05
SLC35E3	108.611	0.850	0.171	-4.972	6.61E-07	2.08E-05
HLA-DRB5	31.674	1.324	0.266	-4.972	6.64E-07	2.09E-05
STK10	80.998	0.998	0.201	-4.969	6.74E-07	2.12E-05
ST13P10	2.063	2.060	0.415	-4.966	6.85E-07	2.15E-05
ARL11	19.725	1.053	0.212	-4.963	6.94E-07	2.17E-05
ABCA2	12.248	1.038	0.209	-4.961	7.00E-07	2.19E-05
SAP130	20.802	0.843	0.170	-4.956	7.21E-07	2.25E-05
ARHGEF1	82.015	0.804	0.162	-4.951	7.39E-07	2.30E-05
NOC2L	25.212	0.802	0.162	-4.949	7.47E-07	2.32E-05
ACAP2	47.856	0.582	0.118	-4.928	8.29E-07	2.57E-05
CHMP4A	71.864	0.733	0.149	-4.925	8.44E-07	2.62E-05
RSBN1L	108.775	0.554	0.113	-4.922	8.59E-07	2.66E-05
LAMB1	57.950	0.504	0.103	-4.915	8.87E-07	2.74E-05
CASC4P1	1.133	2.360	0.480	-4.914	8.94E-07	2.76E-05
IGHG2	31.037	1.583	0.322	-4.912	9.04E-07	2.78E-05
FGD5-AS1	49.358	0.517	0.105	-4.902	9.46E-07	2.91E-05
LINC00710	1.191	2.115	0.431	-4.902	9.48E-07	2.91E-05
PSMC1P1	23.992	0.915	0.187	-4.901	9.54E-07	2.93E-05
AGT	6.155	-1.312	0.268	4.900	9.60E-07	2.94E-05
MRC2	27.114	-0.665	0.136	4.891	1.00E-06	3.06E-05
EIF3J	140.736	0.455	0.093	-4.888	1.02E-06	3.11E-05
FAM21C	8.420	1.058	0.217	-4.887	1.02E-06	3.12E-05
MAU2	67.490	0.775	0.159	-4.877	1.08E-06	3.28E-05
MAP3K10	4.487	1.118	0.229	-4.875	1.09E-06	3.31E-05
DES	48.824	-1.058	0.217	4.867	1.13E-06	3.43E-05
DDX20	15.222	0.938	0.193	-4.865	1.14E-06	3.46E-05
SEC14L4	1.170	1.893	0.389	-4.863	1.15E-06	3.48E-05
ITGBL1	10.280	1.169	0.240	-4.863	1.16E-06	3.48E-05
CRLF3	106.030	0.717	0.148	-4.858	1.19E-06	3.56E-05
BFAR	30.968	0.544	0.112	-4.856	1.20E-06	3.58E-05
TAP2	42.681	0.602	0.124	-4.856	1.20E-06	3.58E-05
ALOX5	8.752	1.182	0.243	-4.857	1.19E-06	3.58E-05
PIK3C2A	122.226	0.472	0.097	-4.854	1.21E-06	3.61E-05
MARCH6	248.320	0.607	0.125	-4.853	1.22E-06	3.62E-05
PCDHGB6	2.388	1.668	0.345	-4.839	1.31E-06	3.89E-05
TRAPPC6A	154.601	0.624	0.129	-4.836	1.32E-06	3.93E-05
CLK1	95.912	0.484	0.100	-4.834	1.34E-06	3.97E-05
ANGPTL2	10.664	-1.062	0.220	4.828	1.38E-06	4.08E-05
IL17RA	110.993	0.713	0.148	-4.827	1.39E-06	4.09E-05
SNORA34	15.407	1.881	0.390	-4.827	1.39E-06	4.09E-05
PHIP	160.387	0.383	0.079	-4.826	1.40E-06	4.11E-05
GARNL3	19.212	1.238	0.257	-4.824	1.41E-06	4.14E-05

DDX50	31.412	0.500	0.104	-4.816	1.47E-06	4.31E-05
SP100	123.654	0.532	0.111	-4.813	1.49E-06	4.36E-05
MALAT1	26125.323	0.323	0.067	-4.812	1.49E-06	4.37E-05
SPG11	43.009	0.697	0.145	-4.803	1.56E-06	4.56E-05
UBR4	69.914	0.671	0.140	-4.795	1.63E-06	4.73E-05
MMP3	5.156	-1.363	0.285	4.787	1.69E-06	4.91E-05
NIPAL3	75.573	0.730	0.153	-4.785	1.71E-06	4.96E-05
PYROXD1	2.161	1.542	0.322	-4.783	1.73E-06	5.00E-05
ANKRD11	211.718	0.386	0.081	-4.780	1.75E-06	5.06E-05
GMEB2	26.313	0.630	0.132	-4.777	1.78E-06	5.12E-05
ANKRD13D	34.558	0.888	0.186	-4.767	1.87E-06	5.39E-05
HSPBAP1	16.040	1.072	0.225	-4.766	1.88E-06	5.40E-05
NFATC3	22.330	1.175	0.247	-4.765	1.89E-06	5.43E-05
MFSD12	34.866	0.702	0.147	-4.761	1.93E-06	5.50E-05
PREPL	24.827	-0.639	0.134	4.759	1.94E-06	5.54E-05
MAP3K14	1.606	1.782	0.375	-4.756	1.98E-06	5.63E-05
XIST	180.914	1.030	0.217	-4.754	2.00E-06	5.66E-05
ZBTB4	60.673	0.600	0.126	-4.752	2.02E-06	5.71E-05
ARFRP1	92.394	0.730	0.154	-4.751	2.02E-06	5.72E-05
EIF2AK2	63.273	0.439	0.092	-4.745	2.09E-06	5.89E-05
PNKP	51.731	0.652	0.138	-4.743	2.10E-06	5.92E-05
IFITM1	3.793	1.536	0.324	-4.740	2.13E-06	6.01E-05
C1R	153.689	0.505	0.107	-4.737	2.17E-06	6.10E-05
WDR47	8.651	1.163	0.246	-4.735	2.19E-06	6.14E-05
KIAA0907	49.276	0.442	0.093	-4.734	2.20E-06	6.17E-05
FOXF1	3.068	-1.161	0.245	4.732	2.23E-06	6.23E-05
GPR141	0.622	2.522	0.533	-4.727	2.28E-06	6.36E-05
SYNRG	322.975	0.769	0.163	-4.726	2.29E-06	6.37E-05
DNASE1	39.437	0.788	0.167	-4.718	2.38E-06	6.64E-05
PLCB2-AS1	0.585	2.547	0.540	-4.716	2.41E-06	6.69E-05
EXOC6	18.378	0.855	0.181	-4.713	2.44E-06	6.77E-05
DMXL2	22.992	0.713	0.151	-4.712	2.45E-06	6.77E-05
USP42	38.987	0.760	0.161	-4.713	2.44E-06	6.77E-05
VAMP2	133.321	0.827	0.176	-4.711	2.47E-06	6.82E-05
P2RX7	19.096	1.100	0.234	-4.707	2.51E-06	6.93E-05
PCSK5	7.691	0.878	0.187	-4.704	2.55E-06	7.02E-05
HNRNPA1P5	1.419	1.802	0.383	-4.703	2.57E-06	7.06E-05
SSTR3	4.616	1.633	0.347	-4.702	2.58E-06	7.07E-05
LSM7	70.732	-0.397	0.084	4.697	2.64E-06	7.23E-05
CPVL	22.804	1.017	0.217	-4.695	2.66E-06	7.29E-05
PNMA1	2.301	1.381	0.294	-4.695	2.67E-06	7.30E-05
ISY1	55.850	0.494	0.105	-4.693	2.69E-06	7.35E-05
PAXIP1	18.911	0.682	0.145	-4.688	2.77E-06	7.54E-05
MLLT6	122.684	0.401	0.086	-4.687	2.77E-06	7.55E-05
GAREML	8.727	1.084	0.231	-4.685	2.81E-06	7.63E-05
ACTR5	7.727	0.931	0.199	-4.681	2.86E-06	7.76E-05
SLC37A2	2.114	1.579	0.338	-4.674	2.95E-06	7.99E-05
ZNF445	20.476	0.572	0.123	-4.673	2.97E-06	8.04E-05
PRR24	41.468	1.145	0.245	-4.668	3.05E-06	8.21E-05
CCDC134	3.639	1.553	0.333	-4.666	3.07E-06	8.26E-05
FBXW4	50.773	0.558	0.120	-4.653	3.26E-06	8.76E-05
DCLK1	3.667	1.203	0.259	-4.653	3.27E-06	8.77E-05
PALD1	37.262	0.794	0.171	-4.649	3.33E-06	8.93E-05
CACFD1	68.512	0.647	0.139	-4.647	3.37E-06	9.00E-05
KRI1	42.104	0.774	0.166	-4.646	3.38E-06	9.02E-05
WAC-AS1	51.653	0.628	0.135	-4.645	3.39E-06	9.05E-05
C19orf52	29.340	0.827	0.178	-4.645	3.41E-06	9.07E-05
UBE2G2	111.984	0.380	0.082	-4.639	3.51E-06	9.32E-05
RRP7B	26.981	0.943	0.203	-4.636	3.55E-06	9.41E-05
TRIM17	2.162	1.703	0.367	-4.635	3.57E-06	9.45E-05
GRIPAP1	192.667	0.503	0.109	-4.633	3.61E-06	9.54E-05
TBCD	6.025	1.254	0.271	-4.633	3.61E-06	9.54E-05
SNHG14	66.890	0.876	0.190	-4.619	3.85E-06	1.01E-04
ZNF169	16.374	0.918	0.199	-4.614	3.95E-06	1.04E-04
SH3PXD2A	90.253	0.413	0.090	-4.614	3.96E-06	1.04E-04
AP1S3	35.361	0.696	0.151	-4.613	3.96E-06	1.04E-04
ARID5B	69.614	0.467	0.101	-4.613	3.97E-06	1.04E-04
TJAP1	17.355	0.744	0.161	-4.612	3.99E-06	1.04E-04
IPP	27.978	0.807	0.175	-4.606	4.11E-06	1.07E-04
B3GNT3	11.416	-0.800	0.174	4.600	4.22E-06	1.10E-04
CAND1	71.674	0.349	0.076	-4.594	4.35E-06	1.13E-04
CC2D1A	36.672	0.656	0.143	-4.594	4.35E-06	1.13E-04
C1QTNF6	17.361	1.270	0.277	-4.591	4.41E-06	1.14E-04
MED13L	79.314	0.518	0.113	-4.589	4.45E-06	1.15E-04
MPRIIP	57.850	0.390	0.085	-4.586	4.53E-06	1.17E-04
ATP5G2	165.415	-0.394	0.086	4.578	4.69E-06	1.21E-04
METTTL16	45.285	0.694	0.152	-4.575	4.77E-06	1.23E-04
GPRASP1	23.398	1.061	0.232	-4.574	4.79E-06	1.23E-04
CREB1	41.839	0.549	0.120	-4.570	4.87E-06	1.25E-04
HPS1	105.187	0.620	0.136	-4.571	4.86E-06	1.25E-04
METTTL7A	125.559	0.606	0.133	-4.569	4.90E-06	1.25E-04
IGFBP3	62.094	0.713	0.156	-4.570	4.88E-06	1.25E-04
RN7SL3	14.960	1.315	0.288	-4.569	4.89E-06	1.25E-04
WDR45	13.173	0.872	0.191	-4.564	5.02E-06	1.28E-04
GIPC1	83.653	0.436	0.095	-4.564	5.02E-06	1.28E-04
VPS39	24.204	0.800	0.175	-4.563	5.03E-06	1.28E-04

MYL9	0.928	2.206	0.483	-4.562	5.06E-06	1.28E-04
CCDC80	124.545	0.633	0.139	-4.559	5.15E-06	1.30E-04
ADAMTS17	4.855	1.381	0.303	-4.553	5.28E-06	1.33E-04
KCTD13	11.265	1.001	0.220	-4.552	5.31E-06	1.34E-04
FKBP5	185.881	0.794	0.174	-4.550	5.36E-06	1.35E-04
ZNF200	13.113	0.850	0.187	-4.549	5.40E-06	1.36E-04
SMCHD1	121.277	0.587	0.129	-4.547	5.44E-06	1.36E-04
TRIM38	179.853	0.750	0.165	-4.545	5.49E-06	1.38E-04
EIF3J-AS1	35.333	0.827	0.182	-4.544	5.51E-06	1.38E-04
C6orf89	85.470	0.334	0.074	-4.542	5.58E-06	1.40E-04
RNF220	105.620	-0.308	0.068	4.539	5.66E-06	1.41E-04
DNAJC7	70.288	0.566	0.125	-4.538	5.67E-06	1.41E-04
SMEK1	58.829	0.422	0.093	-4.537	5.70E-06	1.42E-04
GAS7	18.221	0.842	0.186	-4.536	5.73E-06	1.42E-04
GATS	19.492	0.892	0.197	-4.531	5.87E-06	1.46E-04
TNNT2	4.651	1.340	0.296	-4.529	5.92E-06	1.47E-04
EHMT1	30.369	0.530	0.117	-4.521	6.15E-06	1.52E-04
RAB11FIP1	166.017	0.369	0.082	-4.518	6.26E-06	1.55E-04
PTCHD4	3.336	1.416	0.313	-4.518	6.25E-06	1.55E-04
ASCC2	19.924	0.768	0.170	-4.512	6.41E-06	1.58E-04
KDM3B	42.230	0.572	0.127	-4.507	6.57E-06	1.62E-04
ANKRD40	13.461	0.719	0.160	-4.499	6.83E-06	1.68E-04
SAA1	6.434	1.234	0.274	-4.499	6.83E-06	1.68E-04
ACKR3	6.519	1.083	0.241	-4.496	6.91E-06	1.69E-04
EIF5	155.564	0.324	0.072	-4.496	6.93E-06	1.70E-04
UBA5	18.297	0.686	0.153	-4.492	7.05E-06	1.72E-04
NPM1P39	5.511	1.652	0.368	-4.490	7.11E-06	1.73E-04
TMEM258	130.462	-0.299	0.067	4.487	7.22E-06	1.76E-04
LAMTOR1	98.203	-0.319	0.071	4.482	7.41E-06	1.80E-04
CSF2RA	21.456	0.975	0.217	-4.482	7.40E-06	1.80E-04
AP2A1	27.337	0.946	0.211	-4.479	7.51E-06	1.82E-04
RICTOR	55.603	0.571	0.127	-4.476	7.59E-06	1.84E-04
RN7SL409P	0.843	1.890	0.422	-4.476	7.61E-06	1.84E-04
ELF4	21.308	-0.487	0.109	4.474	7.67E-06	1.85E-04
C17orf70	91.996	0.853	0.191	-4.467	7.95E-06	1.92E-04
RNF213	353.181	0.498	0.112	-4.464	8.05E-06	1.94E-04
SYK	44.131	0.525	0.118	-4.464	8.04E-06	1.94E-04
ZNF786	9.303	1.037	0.232	-4.464	8.06E-06	1.94E-04
CRBN	63.637	0.855	0.192	-4.462	8.11E-06	1.95E-04
COTL1	119.550	0.425	0.095	-4.460	8.20E-06	1.97E-04
SCAND2P	25.562	0.806	0.181	-4.452	8.53E-06	2.04E-04
MAPK12	12.444	1.025	0.230	-4.450	8.59E-06	2.06E-04
H19	335.796	0.832	0.187	-4.450	8.61E-06	2.06E-04
PPP5C	19.341	1.104	0.248	-4.448	8.67E-06	2.07E-04
LAX1	13.461	1.133	0.255	-4.445	8.78E-06	2.09E-04
MCM3	87.018	0.417	0.094	-4.442	8.91E-06	2.12E-04
RALGAPA2	76.342	0.668	0.150	-4.441	8.93E-06	2.13E-04
DHX33	37.036	0.542	0.122	-4.441	8.96E-06	2.13E-04
snoU13	370.692	0.566	0.128	-4.440	8.98E-06	2.13E-04
PTPRM	9.620	0.977	0.220	-4.438	9.08E-06	2.15E-04
LIG3	27.666	0.657	0.148	-4.437	9.13E-06	2.16E-04
ZNF225	4.043	1.268	0.286	-4.436	9.16E-06	2.17E-04
ZNF442	1.920	1.524	0.344	-4.434	9.27E-06	2.19E-04
GPR123	0.757	2.039	0.460	-4.430	9.42E-06	2.22E-04
PAQR3	6.188	1.042	0.235	-4.429	9.46E-06	2.23E-04
TANGO6	15.511	0.727	0.164	-4.427	9.54E-06	2.24E-04
ERLIN1	21.763	0.689	0.156	-4.420	9.89E-06	2.32E-04
GABPB1-AS1	115.285	0.763	0.173	-4.418	9.98E-06	2.34E-04
GUCY1A2	5.762	1.052	0.238	-4.414	1.02E-05	2.38E-04
C9orf38	24.441	1.012	0.230	-4.410	1.03E-05	2.41E-04
SNRNP200	129.115	0.334	0.076	-4.410	1.04E-05	2.42E-04
SMARCE1P1	13.437	1.430	0.324	-4.407	1.05E-05	2.44E-04
RCE1	5.855	0.922	0.209	-4.404	1.06E-05	2.47E-04
FUT11	83.461	0.490	0.111	-4.403	1.07E-05	2.48E-04
ZBTB47	12.468	0.834	0.190	-4.401	1.08E-05	2.50E-04
RAPH1	3.675	1.263	0.287	-4.400	1.08E-05	2.51E-04
BRSK1	1.605	1.531	0.348	-4.400	1.08E-05	2.51E-04
RCSD1	137.335	1.138	0.259	-4.398	1.09E-05	2.52E-04
PRPF6	404.373	0.384	0.087	-4.395	1.11E-05	2.56E-04
YIPF5	32.814	0.482	0.110	-4.394	1.11E-05	2.56E-04
CCZ1	25.771	0.865	0.197	-4.392	1.13E-05	2.59E-04
KRT2	3.426	1.568	0.357	-4.389	1.14E-05	2.62E-04
SNX14	74.159	0.387	0.088	-4.389	1.14E-05	2.62E-04
IRF1	114.422	0.556	0.127	-4.389	1.14E-05	2.62E-04
CERS3	0.659	2.212	0.504	-4.386	1.16E-05	2.64E-04
IGLV2-11	0.688	2.497	0.569	-4.386	1.16E-05	2.64E-04
SKIV2L2	80.713	0.518	0.118	-4.385	1.16E-05	2.65E-04
RABGAP1L-IT1	45.112	0.893	0.204	-4.382	1.17E-05	2.68E-04
TMEM206	8.641	1.055	0.241	-4.380	1.19E-05	2.70E-04
PHGDH	21.173	0.845	0.193	-4.380	1.19E-05	2.70E-04
TMPO	89.630	0.334	0.076	-4.379	1.19E-05	2.71E-04
MMP1	10.408	-1.098	0.251	4.375	1.21E-05	2.76E-04
CTD-2583A14.8	10.926	0.904	0.207	-4.373	1.22E-05	2.78E-04
SLITRK2	0.742	2.102	0.481	-4.372	1.23E-05	2.79E-04
PLXNB2	48.897	0.563	0.129	-4.371	1.23E-05	2.79E-04
MADD	31.633	0.551	0.126	-4.370	1.24E-05	2.81E-04

TM9SF4	25.619	0.591	0.135	-4.368	1.25E-05	2.83E-04
ZHX2	11.063	1.060	0.243	-4.368	1.25E-05	2.83E-04
SMARCD2	208.271	-0.259	0.059	4.368	1.26E-05	2.83E-04
ANKRD33B	0.740	2.460	0.564	-4.365	1.27E-05	2.86E-04
SNRNP35	45.733	0.460	0.106	-4.364	1.28E-05	2.87E-04
NFE2L1	69.862	0.382	0.088	-4.364	1.28E-05	2.87E-04
MED28	51.300	0.401	0.092	-4.359	1.31E-05	2.93E-04
ZNF324B	4.891	1.262	0.290	-4.357	1.32E-05	2.96E-04
TAF1A	9.133	0.911	0.209	-4.355	1.33E-05	2.97E-04
LINC00963	45.961	0.500	0.115	-4.354	1.33E-05	2.98E-04
PLCE1	21.554	0.654	0.150	-4.351	1.35E-05	3.02E-04
CCDC102B	8.296	0.829	0.191	-4.350	1.36E-05	3.03E-04
CEACAM6	168.639	-0.567	0.130	4.345	1.39E-05	3.10E-04
INTS9	4.717	1.062	0.244	-4.345	1.40E-05	3.10E-04
PPARGC1B	12.492	0.837	0.193	-4.344	1.40E-05	3.11E-04
SLCO3A1	38.635	0.686	0.158	-4.341	1.42E-05	3.15E-04
DDI2	103.719	0.485	0.112	-4.340	1.42E-05	3.15E-04
HECTD4	41.618	0.598	0.138	-4.339	1.43E-05	3.17E-04
CHD8	83.388	0.467	0.108	-4.336	1.45E-05	3.21E-04
TMEM219	42.813	-0.446	0.103	4.335	1.46E-05	3.22E-04
DOCK10	36.592	0.806	0.186	-4.334	1.46E-05	3.22E-04
ST13P6	2.445	1.311	0.302	-4.334	1.46E-05	3.22E-04
IGLV3-10	1.063	1.771	0.409	-4.334	1.46E-05	3.22E-04
HINFP	9.886	0.720	0.166	-4.333	1.47E-05	3.23E-04
CSF1R	76.260	0.976	0.225	-4.330	1.49E-05	3.28E-04
BRICD5	2.695	1.401	0.324	-4.325	1.52E-05	3.34E-04
MBD3	98.394	0.709	0.164	-4.324	1.53E-05	3.36E-04
IKZF3	7.773	1.363	0.315	-4.322	1.55E-05	3.37E-04
INE1	5.678	1.094	0.253	-4.318	1.57E-05	3.43E-04
RN7SL801P	1.227	1.696	0.393	-4.318	1.57E-05	3.43E-04
BECN1	16.086	0.754	0.175	-4.318	1.58E-05	3.43E-04
MAP3K12	4.004	1.281	0.297	-4.317	1.58E-05	3.44E-04
FAM178A	30.034	0.573	0.133	-4.316	1.59E-05	3.45E-04
STX6	16.679	0.654	0.151	-4.315	1.59E-05	3.45E-04
PCGF6	3.629	1.078	0.250	-4.310	1.63E-05	3.53E-04
NR3C2	8.321	1.138	0.264	-4.309	1.64E-05	3.54E-04
MFS11	30.830	0.618	0.144	-4.308	1.65E-05	3.55E-04
PRKD3	45.746	0.596	0.139	-4.296	1.74E-05	3.74E-04
TFF2	6.022	1.347	0.314	-4.296	1.74E-05	3.75E-04
PPFIA4	3.955	1.195	0.278	-4.295	1.75E-05	3.76E-04
ZNF254	2.735	1.151	0.268	-4.295	1.75E-05	3.76E-04
AZI1	23.165	0.755	0.176	-4.294	1.76E-05	3.77E-04
UHRF1	20.256	0.800	0.186	-4.293	1.76E-05	3.78E-04
LMOD1	34.471	-0.855	0.199	4.291	1.78E-05	3.80E-04
SRSF4	80.374	0.462	0.108	-4.289	1.79E-05	3.83E-04
SUPV3L1	14.171	0.939	0.219	-4.289	1.79E-05	3.83E-04
INO80E	1.344	1.762	0.411	-4.289	1.80E-05	3.83E-04
CD163L1	5.326	1.201	0.280	-4.288	1.80E-05	3.84E-04
ACTBP2	0.739	1.988	0.464	-4.285	1.82E-05	3.88E-04
AKIRIN1	45.363	0.458	0.107	-4.285	1.83E-05	3.89E-04
STARD4	4.764	-0.961	0.224	4.283	1.84E-05	3.91E-04
LSM1	2.020	1.437	0.335	-4.283	1.85E-05	3.92E-04
SKAP2	44.864	0.530	0.124	-4.282	1.85E-05	3.93E-04
MIR3687	82.485	1.158	0.270	-4.282	1.86E-05	3.93E-04
MLH1	39.824	0.917	0.214	-4.279	1.87E-05	3.96E-04
BLVRA	23.735	0.705	0.165	-4.279	1.88E-05	3.97E-04
EEF2K	18.535	0.709	0.166	-4.272	1.94E-05	4.08E-04
C7	64.016	0.975	0.228	-4.268	1.98E-05	4.16E-04
FKBP1A	234.462	-0.316	0.074	4.266	1.99E-05	4.19E-04
AHDC1	22.051	0.590	0.138	-4.265	2.00E-05	4.19E-04
NLRP1	48.407	0.861	0.202	-4.265	2.00E-05	4.19E-04
SPEG	9.022	-0.939	0.220	4.264	2.01E-05	4.20E-04
EFEMP1	10.587	0.842	0.197	-4.263	2.02E-05	4.21E-04
UBXN11	2.573	1.321	0.310	-4.262	2.03E-05	4.23E-04
PRIMPOL	4.029	1.144	0.269	-4.261	2.03E-05	4.23E-04
EIF2S3L	1.421	1.506	0.353	-4.260	2.04E-05	4.25E-04
CYBA	52.261	0.766	0.180	-4.256	2.08E-05	4.32E-04
ZBTB7C	11.989	1.011	0.237	-4.256	2.09E-05	4.33E-04
SMARCE1	65.968	0.559	0.131	-4.253	2.11E-05	4.36E-04
C12orf43	20.350	0.721	0.170	-4.248	2.16E-05	4.47E-04
TSPO	73.450	-0.379	0.089	4.247	2.17E-05	4.49E-04
SGTA	123.839	0.568	0.134	-4.241	2.22E-05	4.59E-04
INF2	19.462	0.661	0.156	-4.241	2.23E-05	4.60E-04
ZNF652	94.171	0.555	0.131	-4.240	2.23E-05	4.60E-04
TOR2A	1.783	1.593	0.376	-4.239	2.25E-05	4.63E-04
TCP11L2	25.759	0.885	0.209	-4.237	2.27E-05	4.66E-04
TMEM217	2.361	1.267	0.299	-4.236	2.28E-05	4.68E-04
PHLDB2	14.738	0.734	0.173	-4.235	2.28E-05	4.69E-04
CPXM2	7.716	0.934	0.221	-4.235	2.29E-05	4.69E-04
XRCC3	9.183	0.996	0.235	-4.234	2.30E-05	4.70E-04
STAG2	79.263	0.341	0.081	-4.232	2.31E-05	4.73E-04
ZNF792	4.082	-1.061	0.251	4.232	2.32E-05	4.74E-04
CRIP2	67.332	1.024	0.242	-4.228	2.36E-05	4.81E-04
MICAL3	39.144	0.655	0.155	-4.216	2.48E-05	5.05E-04
HSPG2	16.326	0.897	0.213	-4.216	2.49E-05	5.05E-04
GTF2H3	19.929	0.522	0.124	-4.216	2.49E-05	5.05E-04

RNU11	48.551	1.175	0.279	-4.216	2.49E-05	5.05E-04
FAM168B	26.921	0.581	0.138	-4.215	2.50E-05	5.06E-04
OSBPL10-AS1	5.789	1.103	0.262	-4.215	2.50E-05	5.06E-04
DNAJC21	58.271	0.395	0.094	-4.215	2.50E-05	5.06E-04
PI16	3.774	1.251	0.297	-4.213	2.52E-05	5.10E-04
PCIF1	5.424	0.984	0.234	-4.212	2.53E-05	5.11E-04
APOL1	115.011	0.509	0.121	-4.209	2.57E-05	5.17E-04
HLA-DRB6	1.726	1.518	0.361	-4.206	2.60E-05	5.22E-04
DDHD2	11.829	0.680	0.162	-4.204	2.62E-05	5.26E-04
MARK3	81.251	0.297	0.071	-4.204	2.63E-05	5.27E-04
MGLL	141.442	0.545	0.130	-4.202	2.64E-05	5.29E-04
HMGXB4	49.860	0.632	0.150	-4.203	2.64E-05	5.29E-04
PDLIM3	35.840	0.654	0.156	-4.203	2.64E-05	5.29E-04
TMEM144	6.459	-0.963	0.229	-4.198	2.69E-05	5.37E-04
BMPR2	39.554	0.452	0.108	-4.198	2.70E-05	5.38E-04
NPIP11	1.775	1.779	0.424	-4.193	2.75E-05	5.49E-04
COL11A1	38.671	-0.808	0.193	-4.193	2.76E-05	5.50E-04
JMJD8	23.842	0.724	0.173	-4.188	2.82E-05	5.60E-04
GIMAP4	97.783	0.919	0.219	-4.188	2.82E-05	5.60E-04
ZNF33B	32.972	0.474	0.113	-4.186	2.83E-05	5.63E-04
WDR19	8.595	0.825	0.197	-4.186	2.84E-05	5.63E-04
DNMT1	84.198	0.399	0.095	-4.182	2.88E-05	5.71E-04
INIP	35.243	0.454	0.109	-4.179	2.93E-05	5.79E-04
CLEC4C	1.249	1.788	0.428	-4.178	2.94E-05	5.82E-04
ADAM12	55.114	0.568	0.136	-4.177	2.95E-05	5.83E-04
PSAP	1118.588	0.263	0.063	-4.176	2.96E-05	5.84E-04
SNX1	53.569	0.325	0.078	-4.175	2.99E-05	5.89E-04
PPM1G	172.931	0.371	0.089	-4.172	3.02E-05	5.93E-04
LAP3	40.737	0.514	0.123	-4.170	3.05E-05	5.99E-04
IQSEC1	41.686	0.652	0.156	-4.170	3.05E-05	5.99E-04
ZNF724P	7.001	0.871	0.209	-4.169	3.06E-05	6.00E-04
TUG1	54.301	0.546	0.131	-4.162	3.15E-05	6.16E-04
SUGCT	1.161	1.919	0.461	-4.162	3.16E-05	6.17E-04
ABL2	37.819	0.463	0.111	-4.156	3.23E-05	6.31E-04
HIST1H4H	7.501	0.970	0.234	-4.155	3.25E-05	6.34E-04
TRPA1	1.882	-1.219	0.294	-4.145	3.40E-05	6.62E-04
ZNF37A	117.793	0.514	0.124	-4.144	3.41E-05	6.63E-04
TAOK1	103.161	0.414	0.100	-4.141	3.46E-05	6.71E-04
HIGD1C	1.288	-1.509	0.364	-4.140	3.47E-05	6.72E-04
PID1	2.516	-1.073	0.259	-4.140	3.47E-05	6.72E-04
CIRBP	506.149	0.559	0.135	-4.138	3.51E-05	6.79E-04
ATM	247.416	0.738	0.178	-4.135	3.55E-05	6.87E-04
AARSD1	23.963	0.963	0.233	-4.134	3.56E-05	6.89E-04
TMEM50A	17.783	0.563	0.136	-4.133	3.58E-05	6.92E-04
LPIN1	93.473	0.695	0.168	-4.132	3.60E-05	6.94E-04
SCN3A	2.752	-1.174	0.284	-4.130	3.63E-05	6.99E-04
IL6R	61.535	0.687	0.166	-4.128	3.65E-05	7.03E-04
RAPGEF6	17.364	0.797	0.193	-4.127	3.67E-05	7.06E-04
MMAB	33.623	0.587	0.143	-4.120	3.79E-05	7.27E-04
ATF7IP	128.959	0.672	0.163	-4.118	3.82E-05	7.33E-04
IL10RB	54.706	0.498	0.121	-4.112	3.92E-05	7.51E-04
LRCH4	121.050	0.585	0.142	-4.111	3.93E-05	7.53E-04
OBSCN	26.625	1.101	0.268	-4.108	3.99E-05	7.62E-04
CYTH4	94.493	0.994	0.242	-4.107	4.00E-05	7.64E-04
RALGAPA1	73.764	0.560	0.136	-4.107	4.01E-05	7.65E-04
ZNF711	11.663	-0.646	0.157	-4.106	4.03E-05	7.68E-04
CTC-479C5.12	14.902	0.792	0.193	-4.103	4.07E-05	7.74E-04
ENY2	59.145	0.360	0.088	-4.102	4.09E-05	7.77E-04
CXCL14	108.459	-0.617	0.151	-4.101	4.11E-05	7.80E-04
NSMF	102.287	0.477	0.116	-4.100	4.13E-05	7.81E-04
ABCF1	137.086	0.344	0.084	-4.099	4.14E-05	7.83E-04
KB-226F1.1	18.364	1.016	0.248	-4.099	4.15E-05	7.85E-04
SAFB2	83.192	0.341	0.083	-4.097	4.19E-05	7.90E-04
CTA-211A9.5	42.262	0.823	0.201	-4.096	4.20E-05	7.92E-04
C19orf55	22.439	0.937	0.229	-4.095	4.21E-05	7.93E-04
REG1A	20.209	-0.798	0.195	-4.093	4.26E-05	8.00E-04
LINC00176	9.085	1.397	0.341	-4.091	4.30E-05	8.08E-04
EMD	68.467	-0.337	0.082	-4.090	4.31E-05	8.09E-04
FARP2	34.810	0.480	0.117	-4.085	4.42E-05	8.27E-04
METAP2	334.319	0.396	0.097	-4.083	4.44E-05	8.31E-04
TTLL3	51.310	0.732	0.179	-4.083	4.44E-05	8.31E-04
LAMTOR4	168.776	0.509	0.125	-4.082	4.47E-05	8.34E-04
MAP1B	40.321	0.537	0.132	-4.082	4.47E-05	8.34E-04
CCR4	14.012	0.864	0.212	-4.077	4.56E-05	8.51E-04
MAPKBP1	36.782	0.902	0.221	-4.076	4.58E-05	8.52E-04
EXOC3	52.299	0.521	0.128	-4.073	4.64E-05	8.62E-04
WDR81	2.247	1.355	0.333	-4.073	4.65E-05	8.63E-04
PAK4	2.500	1.265	0.311	-4.071	4.68E-05	8.69E-04
ODF2	21.014	0.550	0.135	-4.070	4.70E-05	8.71E-04
ZNF423	6.044	0.857	0.211	-4.070	4.71E-05	8.72E-04
RHOQP1	0.545	1.920	0.472	-4.069	4.72E-05	8.73E-04
CYP1B1	20.193	0.811	0.200	-4.065	4.81E-05	8.88E-04
ZNF367	4.040	0.928	0.228	-4.064	4.83E-05	8.91E-04
CFLAR-AS1	45.704	1.017	0.250	-4.064	4.83E-05	8.91E-04
PSMC6	49.136	0.443	0.109	-4.062	4.86E-05	8.96E-04
DISP2	3.209	1.251	0.308	-4.062	4.87E-05	8.97E-04

SFRP2	11.478	0.959	0.236	-4.061	4.89E-05	8.98E-04
ANKS1A	24.286	0.562	0.138	-4.060	4.91E-05	9.01E-04
SFMBT1	61.993	0.613	0.151	-4.060	4.92E-05	9.01E-04
GMEB1	31.811	0.478	0.118	-4.059	4.93E-05	9.03E-04
APEX1	27.763	0.615	0.151	-4.059	4.94E-05	9.03E-04
KHSRP	339.644	-0.202	0.050	4.058	4.94E-05	9.04E-04
ITGA4	15.107	0.960	0.237	-4.048	5.18E-05	9.43E-04
REXO1	4.940	1.185	0.293	-4.045	5.23E-05	9.53E-04
C9orf78	73.151	0.380	0.094	-4.043	5.27E-05	9.59E-04
RABEP2	38.989	0.849	0.210	-4.042	5.29E-05	9.62E-04
ATP2A2	89.319	0.337	0.084	-4.040	5.35E-05	9.71E-04
MIR142	566.361	0.969	0.240	-4.040	5.35E-05	9.71E-04
NPM1P25	0.886	1.865	0.462	-4.038	5.39E-05	9.77E-04
BAZ2B	142.247	0.367	0.091	-4.037	5.41E-05	9.79E-04
PPP1R12A	116.969	0.446	0.110	-4.033	5.50E-05	9.95E-04
NEIL1	30.492	0.691	0.171	-4.030	5.57E-05	1.01E-03
C10orf10	63.388	0.616	0.153	-4.029	5.60E-05	1.01E-03
TRIP12	105.012	0.336	0.083	-4.027	5.66E-05	1.02E-03
SEC24D	38.027	0.504	0.125	-4.025	5.69E-05	1.02E-03
CASK	47.628	0.420	0.104	-4.023	5.75E-05	1.03E-03
MYO1F	72.712	0.596	0.148	-4.022	5.77E-05	1.04E-03
TYMP	84.800	0.616	0.153	-4.019	5.84E-05	1.05E-03
YME1L1	55.016	0.329	0.082	-4.017	5.91E-05	1.06E-03
OGFR	108.786	0.276	0.069	-4.015	5.94E-05	1.06E-03
PCDH18	1.623	-1.493	0.372	4.013	5.99E-05	1.07E-03
GZF1	10.101	0.710	0.177	-4.011	6.05E-05	1.08E-03
HNRNPA2B1	207.483	0.334	0.083	-4.010	6.07E-05	1.08E-03
UBASH3B	11.722	0.782	0.195	-4.009	6.10E-05	1.09E-03
MAFB	36.818	0.567	0.142	-4.008	6.11E-05	1.09E-03
HOXB3	32.072	0.524	0.131	-4.008	6.14E-05	1.09E-03
KMT2A	240.901	0.426	0.106	-4.006	6.18E-05	1.10E-03
CTA-357J21.1	12.493	0.682	0.171	-4.000	6.34E-05	1.13E-03
CCP110	38.650	0.502	0.126	-3.996	6.45E-05	1.14E-03
MAVS	35.587	0.585	0.146	-3.995	6.46E-05	1.15E-03
LAMC2	43.956	0.597	0.149	-3.995	6.46E-05	1.15E-03
SLC25A3	83.572	0.385	0.097	-3.993	6.52E-05	1.15E-03
ITSN2	94.173	0.401	0.100	-3.993	6.52E-05	1.15E-03
SLC4A2	0.867	1.732	0.434	-3.992	6.55E-05	1.16E-03
SERTAD4	3.246	-1.147	0.287	3.991	6.58E-05	1.16E-03
PTGR2	6.944	0.881	0.221	-3.986	6.73E-05	1.19E-03
PNPLA6	34.552	0.715	0.180	-3.982	6.84E-05	1.20E-03
RNF167	119.791	0.371	0.093	-3.981	6.85E-05	1.20E-03
RECQL	7.026	0.829	0.208	-3.981	6.86E-05	1.21E-03
CTC-205M6.5	1.984	1.384	0.348	-3.979	6.93E-05	1.21E-03
ALKBH6	5.352	-0.686	0.173	3.978	6.95E-05	1.22E-03
TRIP6	32.093	-0.531	0.133	3.978	6.96E-05	1.22E-03
VGLL4	26.116	0.703	0.177	-3.978	6.96E-05	1.22E-03
IL7	3.989	1.057	0.266	-3.975	7.04E-05	1.23E-03
GUSBP2	16.385	0.859	0.216	-3.972	7.13E-05	1.24E-03
ARHGAP17	7.623	0.971	0.245	-3.970	7.19E-05	1.25E-03
ASB13	8.645	0.954	0.240	-3.968	7.24E-05	1.26E-03
MFN2	30.663	0.460	0.116	-3.966	7.32E-05	1.27E-03
ZBTB8OS	24.684	0.465	0.117	-3.963	7.39E-05	1.28E-03
CBX7	31.335	0.588	0.148	-3.963	7.40E-05	1.28E-03
SLIT3	9.188	0.956	0.241	-3.961	7.46E-05	1.29E-03
HIST1H3G	2.094	1.343	0.339	-3.959	7.53E-05	1.30E-03
USP25	21.117	0.528	0.134	-3.954	7.67E-05	1.32E-03
ZNF354B	5.520	0.898	0.227	-3.953	7.72E-05	1.33E-03
GYPB	1.299	1.575	0.399	-3.951	7.80E-05	1.34E-03
SRRM1	188.352	0.322	0.081	-3.950	7.81E-05	1.34E-03
CEBPD	25.215	-0.512	0.130	3.948	7.87E-05	1.35E-03
PSIP1	95.994	0.453	0.115	-3.948	7.88E-05	1.35E-03
BSN	2.699	1.247	0.316	-3.946	7.93E-05	1.36E-03
EEF2	1361.086	0.222	0.056	-3.940	8.14E-05	1.39E-03
FDPS	20.675	0.629	0.160	-3.939	8.17E-05	1.40E-03
ENG	60.707	0.727	0.185	-3.937	8.25E-05	1.41E-03
NUCB1	127.108	0.612	0.156	-3.934	8.36E-05	1.43E-03
PRLR	4.740	-0.957	0.243	3.931	8.47E-05	1.44E-03
PRPF38B	305.879	0.318	0.081	-3.929	8.54E-05	1.45E-03
CACNA1C-AS1	1.040	1.687	0.430	-3.927	8.60E-05	1.46E-03
ISG20L2	37.683	0.473	0.121	-3.924	8.70E-05	1.48E-03
IRS2	29.711	-0.566	0.144	3.924	8.73E-05	1.48E-03
TRPC3	0.845	-1.516	0.387	3.922	8.79E-05	1.49E-03
DNASE2	58.411	0.438	0.112	-3.921	8.82E-05	1.49E-03
TMEM41A	50.652	0.505	0.129	-3.918	8.92E-05	1.51E-03
KIAA1211	2.049	1.247	0.319	-3.914	9.07E-05	1.53E-03
ACAN	1.417	-1.444	0.369	3.912	9.16E-05	1.55E-03
ADAT1	35.101	0.665	0.170	-3.909	9.26E-05	1.56E-03
CTD-2528A14.5	9.013	1.095	0.280	-3.908	9.30E-05	1.57E-03
RBM26	105.242	0.325	0.083	-3.908	9.32E-05	1.57E-03
ADAMTS2	17.624	0.650	0.166	-3.907	9.33E-05	1.57E-03
RN7SL811P	0.907	1.630	0.417	-3.907	9.34E-05	1.57E-03
SNORD3C	194.833	1.288	0.330	-3.907	9.35E-05	1.57E-03
TBC1D9B	31.913	0.541	0.139	-3.904	9.45E-05	1.58E-03
ERGIC3	103.916	0.398	0.102	-3.903	9.48E-05	1.59E-03
KNOP1	40.977	0.497	0.127	-3.903	9.49E-05	1.59E-03

COX7A2	78.701	-0.332	0.085	3.902	9.53E-05	1.59E-03
LINC00657	222.520	0.354	0.091	-3.901	9.57E-05	1.60E-03
VHL	72.245	0.489	0.125	-3.901	9.58E-05	1.60E-03
AGPAT6	54.988	0.420	0.108	-3.900	9.61E-05	1.60E-03
PRR7	47.088	0.401	0.103	-3.900	9.63E-05	1.61E-03
AP1G2	8.913	1.076	0.276	-3.893	9.91E-05	1.65E-03
TRAP1	5.001	1.039	0.267	-3.888	1.01E-04	1.68E-03
RBMX	353.016	0.436	0.112	-3.882	1.04E-04	1.72E-03
RASIP1	38.435	0.863	0.222	-3.881	1.04E-04	1.73E-03
LINC-PINT	24.567	0.612	0.158	-3.880	1.04E-04	1.73E-03
TMEM30B	11.030	0.697	0.180	-3.879	1.05E-04	1.74E-03
MCRS1	21.302	-0.489	0.126	3.874	1.07E-04	1.77E-03
C16orf70	6.324	0.912	0.235	-3.873	1.07E-04	1.77E-03
CHIT1	8.604	0.998	0.258	-3.870	1.09E-04	1.79E-03
CTD-2012K14.7	0.994	1.597	0.413	-3.867	1.10E-04	1.81E-03
SOX10	0.812	1.660	0.430	-3.866	1.11E-04	1.82E-03
TEX26-AS1	4.363	1.056	0.273	-3.864	1.12E-04	1.84E-03
NDUFV3	40.431	0.476	0.123	-3.862	1.13E-04	1.85E-03
CHKA	24.036	0.805	0.208	-3.861	1.13E-04	1.86E-03
ZNF563	21.089	1.063	0.275	-3.860	1.13E-04	1.86E-03
MAPK8IP3	31.671	0.751	0.195	-3.854	1.16E-04	1.90E-03
LA16c-431H6.6	6.295	0.951	0.247	-3.853	1.17E-04	1.91E-03
FTX	30.616	0.612	0.159	-3.853	1.17E-04	1.91E-03
TEX261	34.670	-0.366	0.095	3.850	1.18E-04	1.93E-03
CP	1.710	1.274	0.331	-3.850	1.18E-04	1.93E-03
EIF3A	373.594	0.320	0.083	-3.847	1.19E-04	1.95E-03
PAWR	83.980	0.311	0.081	-3.845	1.21E-04	1.96E-03
CTS2	127.687	0.385	0.100	-3.845	1.21E-04	1.96E-03
ZNF430	67.234	0.650	0.169	-3.845	1.21E-04	1.96E-03
CTC-444N24.11	66.309	0.339	0.088	-3.841	1.23E-04	1.99E-03
FAM213A	80.013	0.401	0.104	-3.841	1.23E-04	1.99E-03
PXDN	103.780	0.476	0.124	-3.841	1.23E-04	1.99E-03
TRIB3	14.730	-0.627	0.163	3.840	1.23E-04	1.99E-03
LRRFIP1	280.221	0.275	0.072	-3.840	1.23E-04	1.99E-03
MAP4K4	361.268	0.334	0.087	-3.839	1.24E-04	2.00E-03
GREM2	0.779	-1.941	0.506	3.835	1.26E-04	2.03E-03
NDUFB5	10.489	-0.674	0.176	3.833	1.26E-04	2.04E-03
ORAI2	74.305	0.672	0.175	-3.831	1.28E-04	2.06E-03
ASH1L	91.099	0.343	0.090	-3.830	1.28E-04	2.06E-03
GSK3B	34.341	0.372	0.097	-3.829	1.29E-04	2.07E-03
CRISPLD2	29.214	0.501	0.131	-3.829	1.29E-04	2.07E-03
FAM111A	108.522	0.352	0.092	-3.829	1.29E-04	2.07E-03
PARP16	2.650	-0.916	0.240	3.823	1.32E-04	2.11E-03
HERC2	60.709	0.404	0.106	-3.824	1.32E-04	2.11E-03
MYO1G	46.272	1.172	0.307	-3.823	1.32E-04	2.11E-03
SRGAP2	21.167	0.684	0.179	-3.819	1.34E-04	2.14E-03
CPSF2	31.508	0.457	0.120	-3.818	1.34E-04	2.15E-03
SMARCA4	184.649	0.420	0.110	-3.818	1.34E-04	2.15E-03
TMEM35	1.147	-1.442	0.378	3.817	1.35E-04	2.15E-03
TLK1	83.809	0.384	0.101	-3.816	1.36E-04	2.16E-03
FLNC	2.718	-1.151	0.302	3.815	1.36E-04	2.17E-03
CXCL10	4.429	0.973	0.255	-3.814	1.36E-04	2.17E-03
ALS2CR12	1.204	1.507	0.396	-3.811	1.39E-04	2.20E-03
SNHG17	21.834	-0.487	0.128	3.809	1.39E-04	2.21E-03
MRI1	2.557	1.116	0.293	-3.807	1.40E-04	2.23E-03
ZNF714	37.098	0.559	0.147	-3.806	1.41E-04	2.23E-03
IGHV1-24	0.819	1.635	0.430	-3.806	1.41E-04	2.23E-03
COQ5	15.608	0.600	0.158	-3.804	1.43E-04	2.25E-03
FMNL1	74.962	0.898	0.236	-3.804	1.43E-04	2.25E-03
ZNF493	10.161	-0.654	0.172	3.802	1.44E-04	2.27E-03
GRP	1.280	-1.525	0.401	3.801	1.44E-04	2.27E-03
DHX36	82.592	0.384	0.101	-3.799	1.45E-04	2.28E-03
ANKHD1	28.129	0.632	0.166	-3.799	1.45E-04	2.28E-03
SUGP1	40.742	0.807	0.212	-3.800	1.45E-04	2.28E-03
TLR9	3.025	1.284	0.338	-3.799	1.45E-04	2.28E-03
ZNF511	18.801	-0.623	0.164	3.797	1.46E-04	2.29E-03
DPAGT1	23.524	-0.465	0.123	3.797	1.46E-04	2.29E-03
PAXBP1	71.220	0.475	0.125	-3.794	1.48E-04	2.32E-03
TCF7	266.588	0.739	0.195	-3.792	1.49E-04	2.34E-03
ATP5J	55.478	-0.395	0.104	3.791	1.50E-04	2.34E-03
HEG1	21.620	0.581	0.153	-3.790	1.51E-04	2.35E-03
VSIG4	12.059	0.736	0.194	-3.789	1.52E-04	2.36E-03
SLC44A2	160.068	0.671	0.177	-3.788	1.52E-04	2.37E-03
ZNF101	50.440	0.673	0.178	-3.788	1.52E-04	2.37E-03
MYO18A	35.570	0.683	0.180	-3.786	1.53E-04	2.38E-03
EYS	2.606	1.121	0.296	-3.786	1.53E-04	2.38E-03
PRMT10	25.124	0.760	0.201	-3.785	1.54E-04	2.38E-03
IGFBP2	149.259	-0.682	0.180	3.780	1.57E-04	2.43E-03
CYTIP	87.463	0.791	0.210	-3.778	1.58E-04	2.45E-03
NR2C2	8.867	0.814	0.216	-3.777	1.59E-04	2.46E-03
GP1BA	3.714	1.113	0.295	-3.776	1.59E-04	2.46E-03
ANKH	53.994	0.399	0.106	-3.776	1.60E-04	2.46E-03
EWSR1	104.244	0.340	0.090	-3.775	1.60E-04	2.47E-03
SFR1	3.318	-1.034	0.274	3.773	1.61E-04	2.48E-03
PRELP	33.318	0.741	0.196	-3.773	1.61E-04	2.48E-03
PDE7A	108.163	0.754	0.200	-3.769	1.64E-04	2.52E-03

EML6	6.824	0.908	0.241	-3.769	1.64E-04	2.52E-03
FTSJ3	55.565	0.517	0.137	-3.764	1.67E-04	2.56E-03
MGMT	48.680	0.553	0.147	-3.764	1.67E-04	2.56E-03
CD81	562.267	0.429	0.114	-3.762	1.69E-04	2.58E-03
ZNHIT3	10.845	0.659	0.175	-3.761	1.69E-04	2.58E-03
NARFL	40.775	0.743	0.198	-3.759	1.71E-04	2.61E-03
ADAR	150.790	0.237	0.063	-3.758	1.71E-04	2.61E-03
BAG1	74.729	-0.275	0.073	3.756	1.72E-04	2.62E-03
PSMA6	23.570	0.662	0.176	-3.756	1.72E-04	2.62E-03
PXDC1	2.629	0.883	0.235	-3.757	1.72E-04	2.62E-03
CTD-2363C16.2	5.432	1.312	0.349	-3.756	1.72E-04	2.62E-03
CDC25B	179.376	0.458	0.122	-3.755	1.73E-04	2.64E-03
C9orf131	2.048	1.345	0.358	-3.753	1.75E-04	2.66E-03
C11orf58	96.507	0.353	0.094	-3.752	1.75E-04	2.66E-03
AMIGO1	7.874	0.924	0.246	-3.751	1.76E-04	2.67E-03
SETBP1	32.442	0.938	0.250	-3.751	1.76E-04	2.67E-03
FAM163A	2.290	1.393	0.372	-3.749	1.77E-04	2.69E-03
SCARNA6	3.182	1.301	0.347	-3.749	1.78E-04	2.69E-03
MIR3916	22.832	0.633	0.169	-3.748	1.78E-04	2.70E-03
PRAF2	45.725	0.640	0.171	-3.748	1.78E-04	2.70E-03
ZNF275	32.424	0.639	0.171	-3.745	1.81E-04	2.73E-03
AHCYL1	165.783	-0.267	0.071	3.743	1.82E-04	2.74E-03
DIO2	14.437	-0.760	0.203	3.743	1.82E-04	2.74E-03
MOB3A	74.094	0.809	0.216	-3.741	1.83E-04	2.76E-03
ZBP1	5.372	0.835	0.223	-3.741	1.83E-04	2.76E-03
MMP17	3.244	1.296	0.347	-3.741	1.83E-04	2.76E-03
CTA-217C2.1	26.421	0.642	0.172	-3.740	1.84E-04	2.77E-03
PPM1H	3.768	0.929	0.249	-3.734	1.89E-04	2.83E-03
SORBS2	20.062	0.627	0.168	-3.733	1.89E-04	2.84E-03
VIPR1	20.631	0.849	0.228	-3.731	1.91E-04	2.86E-03
TMEM186	7.579	0.914	0.245	-3.730	1.92E-04	2.87E-03
APOBEC3A	2.372	1.218	0.327	-3.728	1.93E-04	2.89E-03
CYP11B1-AS1	0.827	1.385	0.372	-3.725	1.95E-04	2.92E-03
HPS5	23.009	0.623	0.167	-3.723	1.97E-04	2.94E-03
KCNMA1	11.512	-0.774	0.208	3.722	1.98E-04	2.94E-03
TIMM13	218.631	0.271	0.073	-3.720	1.99E-04	2.96E-03
CTBP2	109.337	0.329	0.088	-3.720	1.99E-04	2.96E-03
BRF1	20.305	0.583	0.157	-3.713	2.04E-04	3.04E-03
TRPM4	20.217	0.573	0.154	-3.713	2.05E-04	3.04E-03
SNX30	9.206	0.649	0.175	-3.713	2.05E-04	3.04E-03
TRIM44	127.169	0.348	0.094	-3.712	2.06E-04	3.05E-03
TPM3P9	11.822	0.731	0.197	-3.711	2.06E-04	3.05E-03
CTGF	67.587	0.500	0.135	-3.706	2.10E-04	3.11E-03
ABHD3	23.728	0.725	0.196	-3.706	2.11E-04	3.11E-03
ATP6V0A2	29.203	0.477	0.129	-3.701	2.14E-04	3.16E-03
NRDE2	17.279	0.587	0.159	-3.701	2.14E-04	3.16E-03
NXPE3	48.726	0.708	0.191	-3.701	2.15E-04	3.16E-03
ZCCHC11	45.644	0.413	0.112	-3.698	2.17E-04	3.19E-03
NUMA1	54.015	0.528	0.143	-3.698	2.17E-04	3.19E-03
C19orf35	7.156	1.096	0.296	-3.698	2.18E-04	3.20E-03
CHCHD6	6.568	0.742	0.201	-3.696	2.19E-04	3.22E-03
FBLN2	3.285	1.094	0.296	-3.695	2.20E-04	3.22E-03
CTD-2026C7.1	0.529	1.929	0.522	-3.695	2.20E-04	3.22E-03
PNPLA7	53.468	1.159	0.314	-3.695	2.20E-04	3.22E-03
ZFP2	0.809	1.564	0.423	-3.693	2.21E-04	3.23E-03
GIT2	36.096	0.609	0.165	-3.691	2.23E-04	3.26E-03
ANKRD61	0.623	2.007	0.544	-3.691	2.24E-04	3.26E-03
SIK3	21.351	0.618	0.167	-3.690	2.25E-04	3.27E-03
RN7SL99P	2.396	1.172	0.318	-3.688	2.26E-04	3.29E-03
CCR1	17.272	0.635	0.172	-3.688	2.26E-04	3.29E-03
RAB7L1	45.694	0.590	0.160	-3.687	2.27E-04	3.30E-03
YBX1P2	14.608	0.598	0.162	-3.686	2.28E-04	3.32E-03
NAGK	51.209	0.557	0.151	-3.684	2.30E-04	3.33E-03
GPR155	94.192	0.822	0.223	-3.684	2.29E-04	3.33E-03
TMPRSS11E	0.589	1.666	0.452	-3.684	2.30E-04	3.33E-03
RORC	1.131	-1.478	0.401	3.683	2.31E-04	3.35E-03
FAM127B	43.401	-0.512	0.139	3.682	2.32E-04	3.36E-03
LENG8-AS1	2.912	1.015	0.276	-3.680	2.33E-04	3.37E-03
TIGIT	38.102	0.778	0.211	-3.680	2.33E-04	3.37E-03
DIABLO	2.299	1.153	0.313	-3.680	2.34E-04	3.37E-03
APOBR	16.863	0.592	0.161	-3.678	2.35E-04	3.39E-03
ZNF638	90.215	0.325	0.088	-3.678	2.35E-04	3.39E-03
NAA40	22.456	0.644	0.175	-3.677	2.36E-04	3.40E-03
MRS2	8.487	0.642	0.175	-3.673	2.40E-04	3.44E-03
HID1	7.389	0.804	0.219	-3.673	2.40E-04	3.44E-03
CDC14A	3.499	0.973	0.265	-3.673	2.40E-04	3.44E-03
SLC45A3	12.035	-0.641	0.175	3.673	2.40E-04	3.44E-03
RUSC1	34.096	0.397	0.108	-3.672	2.40E-04	3.45E-03
ZNF587	51.938	0.480	0.131	-3.672	2.41E-04	3.45E-03
RNF113A	14.231	0.683	0.186	-3.672	2.41E-04	3.45E-03
SLFN5	62.050	0.551	0.150	-3.671	2.41E-04	3.45E-03
GLB1L	17.634	0.721	0.196	-3.671	2.42E-04	3.45E-03
CTD-2126E3.3	3.481	1.108	0.302	-3.670	2.42E-04	3.46E-03
TMEM120B	26.290	0.409	0.111	-3.668	2.44E-04	3.48E-03
IFIT5	4.990	0.762	0.208	-3.665	2.47E-04	3.52E-03
ERCC8	5.599	0.843	0.230	-3.665	2.47E-04	3.52E-03

CTD-2303H24.2	0.566	1.826	0.498	-3.664	2.48E-04	3.54E-03
LZTS3	15.012	0.803	0.219	-3.663	2.49E-04	3.55E-03
CHRA1	12.646	0.618	0.169	-3.662	2.51E-04	3.56E-03
EXTL3	12.226	0.663	0.181	-3.662	2.50E-04	3.56E-03
SNORD105B	1.300	1.449	0.396	-3.662	2.50E-04	3.56E-03
TBC1D8B	11.464	0.702	0.192	-3.659	2.54E-04	3.60E-03
RASA2-IT1	13.772	0.826	0.226	-3.655	2.57E-04	3.64E-03
PITRM1	21.503	0.449	0.123	-3.654	2.58E-04	3.65E-03
GPS1	46.630	0.546	0.149	-3.654	2.58E-04	3.65E-03
ZNF557	5.021	0.894	0.245	-3.654	2.58E-04	3.66E-03
TRIM56	225.507	0.481	0.132	-3.653	2.59E-04	3.66E-03
PLK3	15.333	0.721	0.197	-3.653	2.59E-04	3.66E-03
PANK2	62.052	0.439	0.120	-3.653	2.59E-04	3.66E-03
DET1	2.091	1.265	0.346	-3.653	2.59E-04	3.66E-03
PPP2R5A	31.737	-0.349	0.096	3.651	2.61E-04	3.68E-03
POGLUT1	73.744	0.626	0.171	-3.649	2.63E-04	3.70E-03
DPYSL2	50.395	0.417	0.114	-3.647	2.65E-04	3.73E-03
C1orf27	26.688	0.422	0.116	-3.646	2.66E-04	3.74E-03
ORTA19P	4.811	0.977	0.268	-3.645	2.67E-04	3.75E-03
ZBTB16	91.605	1.010	0.277	-3.645	2.67E-04	3.75E-03
NAAA	32.347	0.608	0.167	-3.645	2.67E-04	3.75E-03
PGPEP1	106.079	0.420	0.115	-3.641	2.71E-04	3.80E-03
PHB2	159.413	0.340	0.093	-3.640	2.73E-04	3.82E-03
MBOAT1	9.478	-0.615	0.169	3.638	2.75E-04	3.84E-03
TIAM2	6.423	0.822	0.226	-3.638	2.75E-04	3.84E-03
CTD-2555O16.4	1.400	1.432	0.394	-3.638	2.75E-04	3.84E-03
TCF25	476.747	0.387	0.106	-3.637	2.75E-04	3.84E-03
RPTOR	5.552	1.068	0.294	-3.637	2.76E-04	3.84E-03
QRSL1	112.450	0.605	0.166	-3.637	2.76E-04	3.85E-03
FOXC1	12.376	0.630	0.173	-3.636	2.77E-04	3.86E-03
ESRP2	10.991	0.747	0.206	-3.635	2.78E-04	3.86E-03
CLN8	33.622	0.676	0.186	-3.635	2.78E-04	3.86E-03
KCNJ14	16.896	0.705	0.194	-3.634	2.79E-04	3.88E-03
BRD7	111.538	0.360	0.099	-3.631	2.83E-04	3.92E-03
MKLN1	93.945	0.387	0.107	-3.630	2.83E-04	3.92E-03
IFT57	35.177	0.467	0.129	-3.630	2.83E-04	3.93E-03
FOXP2	2.426	-1.330	0.367	3.628	2.86E-04	3.96E-03
ASH2L	14.345	0.589	0.162	-3.628	2.86E-04	3.96E-03
LEFTY1	6.490	-0.949	0.262	3.626	2.88E-04	3.98E-03
SMCR8	2.512	1.348	0.372	-3.626	2.88E-04	3.98E-03
GTF2IRD2B	14.007	0.685	0.189	-3.624	2.90E-04	4.01E-03
RN7SL443P	1.354	1.350	0.373	-3.624	2.90E-04	4.01E-03
PIGM	92.623	0.455	0.126	-3.624	2.91E-04	4.01E-03
LARP7	52.643	0.358	0.099	-3.623	2.91E-04	4.02E-03
TMC6	308.718	0.467	0.129	-3.620	2.95E-04	4.06E-03
EHBP1L1	109.155	0.664	0.183	-3.619	2.96E-04	4.06E-03
SIDT2	71.517	0.396	0.110	-3.617	2.98E-04	4.09E-03
RIMS3	3.926	-0.873	0.241	3.615	3.00E-04	4.12E-03
HNRNPK	260.810	-0.313	0.087	3.615	3.00E-04	4.12E-03
ABCA7	31.251	0.845	0.234	-3.615	3.00E-04	4.12E-03
MCOLN2	30.066	0.548	0.152	-3.613	3.03E-04	4.15E-03
DKFZP779L1853	3.955	1.018	0.282	-3.612	3.04E-04	4.16E-03
FRMD3	3.046	-0.966	0.267	3.611	3.05E-04	4.17E-03
CTD-2083E4.4	11.918	1.183	0.328	-3.611	3.04E-04	4.17E-03
ZNF594	7.960	0.818	0.226	-3.611	3.05E-04	4.17E-03
LRRC2	9.792	0.707	0.196	-3.611	3.06E-04	4.17E-03
TPM4	342.960	0.291	0.081	-3.608	3.09E-04	4.21E-03
TBRG4	30.999	-0.391	0.108	3.607	3.10E-04	4.23E-03
KLHL5	24.121	0.624	0.173	-3.605	3.12E-04	4.25E-03
HMBOX1-IT1	19.153	0.767	0.213	-3.602	3.15E-04	4.29E-03
COX6B1	105.665	-0.295	0.082	3.600	3.18E-04	4.32E-03
NUGGC	5.214	1.111	0.309	-3.599	3.19E-04	4.33E-03
STAT6	89.443	0.363	0.101	-3.597	3.22E-04	4.37E-03
MED20	3.849	-0.918	0.255	3.596	3.23E-04	4.38E-03
KDM4B	96.861	0.498	0.139	-3.596	3.24E-04	4.38E-03
KLHDC4	32.760	0.512	0.142	-3.596	3.24E-04	4.38E-03
C15orf27	34.544	0.661	0.184	-3.595	3.24E-04	4.38E-03
FKBP9	13.413	-0.583	0.162	3.595	3.25E-04	4.39E-03
TMEM184C	7.968	0.665	0.185	-3.594	3.25E-04	4.40E-03
PGM1	15.430	-0.528	0.147	3.594	3.26E-04	4.40E-03
NFKBIL1	34.222	0.514	0.143	-3.594	3.26E-04	4.40E-03
TRAPPC3	36.482	-0.337	0.094	3.590	3.30E-04	4.45E-03
DCXR	65.268	0.486	0.135	-3.588	3.33E-04	4.48E-03
FOXP4	8.928	0.686	0.191	-3.587	3.34E-04	4.49E-03
EXOSC2	33.493	0.455	0.127	-3.587	3.34E-04	4.49E-03
SAMD8	7.243	0.686	0.191	-3.587	3.35E-04	4.49E-03
HCG11	17.151	0.633	0.176	-3.587	3.35E-04	4.49E-03
RNU1-4	1.393	1.283	0.358	-3.583	3.40E-04	4.55E-03
HIST1H2AL	1.453	1.356	0.379	-3.582	3.41E-04	4.56E-03
KLHL18	7.090	0.667	0.186	-3.582	3.41E-04	4.56E-03
IRF8	27.188	0.731	0.204	-3.582	3.42E-04	4.56E-03
ZNF227	6.881	0.665	0.186	-3.580	3.44E-04	4.59E-03
ACTN4	130.528	0.268	0.075	-3.577	3.47E-04	4.63E-03
NCAPH2	91.921	0.539	0.151	-3.576	3.49E-04	4.64E-03
SF3A2	42.917	0.565	0.158	-3.575	3.50E-04	4.66E-03
PAN3	46.141	-0.410	0.115	3.573	3.52E-04	4.69E-03

STAT5A	37.630	0.359	0.100	-3.573	3 53E-04	4.69E-03
LINC00969	65.326	0.477	0.133	-3.573	3 53E-04	4.69E-03
POLE	53.158	0.651	0.182	-3.572	3 54E-04	4.70E-03
SMARCB1	24.343	-0.404	0.113	3.572	3 54E-04	4.70E-03
ZC3HAV1	97.451	0.391	0.110	-3.570	3 57E-04	4.73E-03
COL4A2	110.695	0.358	0.100	-3.569	3 59E-04	4.75E-03
RN7SL32P	10.396	1.042	0.292	-3.568	3 59E-04	4.75E-03
ETV6	68.511	0.379	0.106	-3.568	3.60E-04	4.77E-03
DCUN1D2-AS	1.075	1.585	0.444	-3.567	3.61E-04	4.77E-03
RRP1	27.032	0.523	0.147	-3.566	3.63E-04	4.79E-03
ST13P15	6.155	0.769	0.216	-3.566	3.63E-04	4.79E-03
TMEM79	21.599	0.569	0.160	-3.561	3.69E-04	4.86E-03
CXCL9	11.313	0.714	0.201	-3.560	3.71E-04	4.89E-03
RC3H1	48.774	0.381	0.107	-3.558	3.73E-04	4.91E-03
VOPP1	67.438	-0.308	0.086	3.558	3.74E-04	4.91E-03
PLCH2	9.044	-0.786	0.221	3.556	3.77E-04	4.95E-03
RBBP4	41.732	0.510	0.144	-3.556	3.77E-04	4.95E-03
NMNAT1	9.062	0.666	0.187	-3.555	3.78E-04	4.95E-03
SNRPD2	128.924	-0.282	0.079	3.555	3.78E-04	4.96E-03
H6PD	99.254	0.492	0.138	-3.554	3 80E-04	4.98E-03
CALM3	34.363	0.515	0.145	-3.551	3 83E-04	5.02E-03
C7orf26	12.333	-0.520	0.146	3.550	3 85E-04	5.04E-03
LRRC26	5.049	1.021	0.288	-3.549	3 86E-04	5.04E-03
EIF5A	43.324	-0.469	0.132	3.547	3 89E-04	5.08E-03
TBX2	5.261	-0.861	0.243	3.546	3 91E-04	5.10E-03
FAM57B	0.554	1.706	0.481	-3.546	3 91E-04	5.10E-03
CCDC28A	24.505	0.462	0.130	-3.544	3 94E-04	5.12E-03
QSOX2	6.131	0.711	0.201	-3.544	3 93E-04	5.12E-03
HSF2BP	1.346	1.342	0.379	-3.545	3 93E-04	5.12E-03
MUS81	17.324	0.473	0.134	-3.543	3 96E-04	5.14E-03
HPS4	49.717	0.400	0.113	-3.540	4 00E-04	5.20E-03
EP400NL	11.584	0.797	0.225	-3.539	4 02E-04	5.21E-03
CDC73	34.543	0.458	0.129	-3.538	4 03E-04	5.22E-03
NOSIP	107.902	0.557	0.157	-3.538	4 03E-04	5.22E-03
GIGYF1	72.314	0.407	0.115	-3.534	4.10E-04	5.30E-03
HMCN2	4.649	-0.945	0.268	3.533	4.11E-04	5.31E-03
CFB	25.678	-0.473	0.134	3.533	4.11E-04	5.31E-03
CREBRF	72.577	0.486	0.137	-3.533	4.11E-04	5.31E-03
PTCHD2	1.091	1.351	0.383	-3.532	4.13E-04	5.33E-03
RSF1-IT1	5.571	0.946	0.268	-3.531	4.14E-04	5.34E-03
OR2A1-AS1	0.960	1.390	0.394	-3.530	4.15E-04	5.35E-03
STK4	273.444	0.355	0.101	-3.530	4.15E-04	5.35E-03
ZNF91	29.463	0.459	0.130	-3.530	4.16E-04	5.36E-03
GULP1	10.872	-0.697	0.198	3.528	4.18E-04	5.38E-03
DPP7	155.969	0.466	0.132	-3.527	4 21E-04	5.41E-03
TMBIM4	55.546	0.358	0.101	-3.526	4 22E-04	5.41E-03
LRMP	49.497	0.864	0.245	-3.524	4 25E-04	5.45E-03
INTS3	28.911	0.499	0.142	-3.524	4 26E-04	5.45E-03
TNIP1	133.127	0.400	0.114	-3.523	4 27E-04	5.47E-03
ZNF37BP	33.733	0.563	0.160	-3.523	4 27E-04	5.47E-03
SNORA42	2.797	1.137	0.323	-3.518	4 35E-04	5.56E-03
IGHA1	1811.509	0.711	0.202	-3.517	4 37E-04	5.58E-03
ZNF202	9.710	0.875	0.249	-3.515	4.40E-04	5.61E-03
PTAFR	16.987	0.586	0.167	-3.515	4.40E-04	5.61E-03
DZIP3	9.988	-0.575	0.164	3.514	4.42E-04	5.63E-03
GUSB	32.928	-0.362	0.103	3.513	4.43E-04	5.64E-03
N4BP2	27.133	0.659	0.188	-3.508	4 51E-04	5.74E-03
SCYL1	21.575	-0.451	0.129	3.507	4 54E-04	5.76E-03
IFI30	5.682	0.939	0.268	-3.506	4 54E-04	5.76E-03
TPT1	2097.379	-0.234	0.067	3.505	4 56E-04	5.79E-03
C15orf41	5.794	0.824	0.235	-3.504	4 58E-04	5.80E-03
C21orf58	20.079	0.753	0.215	-3.504	4 59E-04	5.82E-03
POLR2B	28.480	0.486	0.139	-3.502	4.62E-04	5.85E-03
CRYZL1	25.942	0.533	0.152	-3.502	4.62E-04	5.85E-03
IGF1R	31.684	0.453	0.129	-3.500	4.66E-04	5.89E-03
KRT18	9.500	0.822	0.235	-3.499	4.68E-04	5.91E-03
BAZ2A	146.650	0.322	0.092	-3.497	4.70E-04	5.94E-03
LACTB	16.635	0.484	0.139	-3.494	4.76E-04	6.01E-03
IZUMO4	30.007	0.721	0.206	-3.494	4.76E-04	6.01E-03
ECHDC2	62.128	0.357	0.102	-3.493	4.78E-04	6.03E-03
GCC2	177.346	0.326	0.093	-3.492	4.79E-04	6.03E-03
NAPEPLD	36.346	0.502	0.144	-3.492	4.79E-04	6.03E-03
GABPB1	8.467	0.658	0.188	-3.490	4 83E-04	6.07E-03
EEF1A1P25	0.828	1.567	0.449	-3.490	4 83E-04	6.07E-03
DSTYK	44.404	0.389	0.111	-3.488	4 86E-04	6.10E-03
PXN	8.935	0.634	0.182	-3.488	4 86E-04	6.10E-03
CTB-129O4.1	1.935	1.078	0.309	-3.487	4 89E-04	6.13E-03
KLHL21	19.988	-0.488	0.140	3.487	4 89E-04	6.13E-03
ZNF793	20.092	0.874	0.251	-3.485	4 92E-04	6.16E-03
LINC00476	9.757	0.700	0.201	-3.482	4 98E-04	6.24E-03
TUBBP1	1.517	1.023	0.294	-3.481	4 99E-04	6.24E-03
LINC00637	1.303	1.382	0.397	-3.481	4 99E-04	6.24E-03
CTC-575110.1	1.297	1.266	0.364	-3.478	5 06E-04	6.32E-03
HUS1	15.399	-0.505	0.145	3.477	5 08E-04	6.33E-03
SPATS2	37.440	0.442	0.127	-3.477	5 08E-04	6.33E-03
RNA18S5	7929.059	0.600	0.172	-3.477	5 07E-04	6.33E-03

PLN	1.822	-1.162	0.334	3.475	5.11E-04	6.37E-03
IGHV2-26	1.995	1.274	0.367	-3.475	5.12E-04	6.37E-03
PPP1R13L	15.903	0.537	0.155	-3.474	5.14E-04	6.39E-03
AP4M1	11.649	0.645	0.186	-3.472	5.17E-04	6.43E-03
CTD-2192J16.15	402.983	0.480	0.138	-3.471	5.19E-04	6.45E-03
MS4A8	2.454	-1.191	0.343	3.470	5.21E-04	6.47E-03
RPUSD2	8.525	-0.664	0.191	3.467	5.26E-04	6.52E-03
GTPBP8	24.771	0.454	0.131	-3.467	5.26E-04	6.52E-03
TMEM98	41.386	-0.556	0.160	3.466	5.28E-04	6.53E-03
GMIP	94.349	0.707	0.204	-3.466	5.28E-04	6.53E-03
HERC4	77.316	0.366	0.106	-3.465	5.31E-04	6.56E-03
ATAD2	47.446	0.385	0.111	-3.464	5.32E-04	6.57E-03
ZNF503	18.244	-0.523	0.151	3.463	5.34E-04	6.59E-03
YAF2	33.510	0.494	0.143	-3.463	5.34E-04	6.59E-03
SLC16A1-AS1	10.091	0.680	0.197	-3.459	5.41E-04	6.67E-03
PSMC1	81.186	0.498	0.144	-3.459	5.42E-04	6.68E-03
BCL6	85.641	0.514	0.149	-3.459	5.43E-04	6.68E-03
MBNL3	13.709	-0.515	0.149	3.458	5.43E-04	6.68E-03
CPT1A	92.408	-0.255	0.074	3.458	5.45E-04	6.70E-03
NCDN	21.453	0.721	0.209	-3.456	5.48E-04	6.74E-03
NT5DC4	0.494	1.839	0.532	-3.456	5.49E-04	6.74E-03
IRF3	83.260	0.479	0.139	-3.453	5.54E-04	6.78E-03
CCDC152	4.484	0.940	0.272	-3.453	5.54E-04	6.78E-03
MXRA5	28.184	-0.579	0.168	3.453	5.55E-04	6.80E-03
ASPSCR1	27.083	0.597	0.173	-3.452	5.56E-04	6.80E-03
RAF1	60.563	0.574	0.166	-3.450	5.60E-04	6.84E-03
TOMM7	254.851	-0.269	0.078	3.450	5.62E-04	6.86E-03
KIF2A	29.342	0.434	0.126	-3.449	5.63E-04	6.88E-03
GNAZ	1.908	-1.050	0.305	3.448	5.65E-04	6.90E-03
RSPH4A	1.759	-0.994	0.289	3.444	5.73E-04	6.98E-03
SNORA77	39.909	0.628	0.182	-3.444	5.73E-04	6.98E-03
PTBP1	292.107	0.333	0.097	-3.442	5.77E-04	7.02E-03
SAP18	100.091	-0.328	0.095	3.442	5.78E-04	7.03E-03
DDX49	25.855	0.507	0.147	-3.441	5.79E-04	7.03E-03
ZRANB1	31.695	0.474	0.138	-3.439	5.85E-04	7.10E-03
CEACAM19	4.089	0.821	0.239	-3.437	5.87E-04	7.13E-03
CTD-2228K2.7	8.020	-0.726	0.211	3.437	5.89E-04	7.14E-03
NTHL1	34.247	0.589	0.171	-3.437	5.89E-04	7.14E-03
OVOL1	3.035	1.096	0.319	-3.437	5.88E-04	7.14E-03
SRSF8	56.230	0.293	0.085	-3.435	5.92E-04	7.16E-03
SUMF2	93.331	0.447	0.130	-3.433	5.97E-04	7.22E-03
PRMT7	48.563	0.633	0.185	-3.433	5.98E-04	7.22E-03
ZNF773	15.372	0.970	0.283	-3.432	5.98E-04	7.22E-03
LIPG	3.866	0.915	0.267	-3.431	6.01E-04	7.26E-03
HAUS6	32.256	0.442	0.129	-3.429	6.05E-04	7.29E-03
RSPH3	21.988	0.491	0.143	-3.428	6.07E-04	7.32E-03
CTD-2267D19.2	0.837	-1.262	0.368	3.427	6.10E-04	7.35E-03
FAU	697.083	-0.185	0.054	3.427	6.10E-04	7.35E-03
PBRM1	82.522	0.279	0.081	-3.426	6.12E-04	7.36E-03
STK38	171.161	-0.254	0.074	3.426	6.13E-04	7.37E-03
SNORA76	576.788	0.465	0.136	-3.425	6.15E-04	7.38E-03
STAG3L2	38.075	0.768	0.225	-3.421	6.23E-04	7.48E-03
AES	5.440	0.765	0.224	-3.421	6.25E-04	7.50E-03
SHC1P2	1.290	1.358	0.397	-3.419	6.28E-04	7.53E-03
CNTRL	70.423	0.504	0.147	-3.418	6.30E-04	7.54E-03
CHST11	13.719	0.616	0.180	-3.418	6.30E-04	7.54E-03
RHOA-IT1	8.672	0.908	0.266	-3.417	6.32E-04	7.56E-03
TMEM2	43.847	0.371	0.109	-3.416	6.34E-04	7.58E-03
MFAP1	40.098	0.427	0.125	-3.414	6.39E-04	7.63E-03
STX12	26.732	-0.397	0.116	3.413	6.43E-04	7.67E-03
DXO	4.337	0.867	0.254	-3.412	6.46E-04	7.70E-03
GNB4	5.558	0.756	0.222	-3.410	6.49E-04	7.74E-03
NF1	90.048	0.343	0.100	-3.408	6.53E-04	7.78E-03
MFAP3	5.920	0.746	0.219	-3.409	6.53E-04	7.78E-03
THEMIS2	16.219	0.757	0.222	-3.408	6.55E-04	7.79E-03
OLA1	69.555	-0.286	0.084	3.407	6.56E-04	7.79E-03
ZNF611	29.767	0.580	0.170	-3.407	6.56E-04	7.79E-03
TRANK1	12.305	0.667	0.196	-3.407	6.57E-04	7.80E-03
GIGYF2	108.456	0.245	0.072	-3.406	6.59E-04	7.81E-03
ZNF514	20.623	0.483	0.142	-3.404	6.63E-04	7.86E-03
FIBP	47.667	0.524	0.154	-3.404	6.64E-04	7.86E-03
Z83844.1	2.018	1.147	0.337	-3.404	6.64E-04	7.86E-03
IQCA1	1.540	-1.177	0.346	3.403	6.66E-04	7.87E-03
PPP1R42	1.091	1.342	0.395	-3.399	6.77E-04	8.01E-03
TPR	111.849	0.330	0.097	-3.397	6.81E-04	8.05E-03
CCL11	2.758	-1.012	0.298	3.394	6.89E-04	8.13E-03
EPS8L3	34.513	-0.449	0.132	3.393	6.91E-04	8.15E-03
ATXN2L	109.389	0.454	0.134	-3.390	6.98E-04	8.23E-03
MRVI1	20.852	-0.536	0.158	3.388	7.05E-04	8.30E-03
HLA-B	912.651	0.229	0.068	-3.387	7.08E-04	8.33E-03
NGFR	5.152	0.784	0.231	-3.385	7.11E-04	8.35E-03
HCG25	12.358	0.852	0.252	-3.385	7.11E-04	8.35E-03
GPR21	0.666	1.575	0.465	-3.385	7.11E-04	8.36E-03
ADCY8	0.646	1.799	0.532	-3.385	7.12E-04	8.36E-03
TST	64.218	-0.316	0.093	3.384	7.15E-04	8.38E-03
ESCO2	10.124	0.620	0.183	-3.384	7.14E-04	8.38E-03

STX2	18.624	0.540	0.160	-3.383	7.16E-04	8.39E-03
RASSF4	89.172	0.647	0.191	-3.382	7.19E-04	8.40E-03
LYL1	16.330	1.084	0.321	-3.382	7.19E-04	8.40E-03
HOOK3	78.274	0.340	0.101	-3.381	7.22E-04	8.43E-03
NPAP1P4	2.640	1.012	0.299	-3.381	7.23E-04	8.44E-03
MIR222	0.895	1.942	0.575	-3.379	7.26E-04	8.48E-03
TNFSF13B	54.822	0.434	0.128	-3.378	7.30E-04	8.51E-03
SYNPO	35.760	0.522	0.155	-3.376	7.37E-04	8.58E-03
CTD-2020K17.1	53.316	0.884	0.262	-3.376	7.37E-04	8.58E-03
CCDC30	7.899	0.878	0.260	-3.374	7.40E-04	8.61E-03
UBE3C	85.488	0.372	0.110	-3.374	7.42E-04	8.62E-03
CASP9	3.460	0.901	0.267	-3.374	7.42E-04	8.62E-03
CTD-2026K11.2	1.191	1.404	0.416	-3.374	7.42E-04	8.62E-03
ARID4B	172.199	0.306	0.091	-3.373	7.43E-04	8.62E-03
SERF2	307.227	-0.238	0.071	3.370	7.51E-04	8.72E-03
NAT10	50.870	0.365	0.108	-3.370	7.52E-04	8.72E-03
NUDCD2	21.637	0.422	0.125	-3.370	7.52E-04	8.72E-03
IGFBP7	108.580	0.430	0.128	-3.369	7.54E-04	8.74E-03
ACTR1A	57.033	0.613	0.182	-3.368	7.57E-04	8.76E-03
C5orf45	6.421	0.789	0.234	-3.368	7.57E-04	8.76E-03
WDR60	43.189	0.390	0.116	-3.368	7.58E-04	8.76E-03
TNFAIP3	126.067	0.545	0.162	-3.366	7.63E-04	8.80E-03
MYOM2	3.376	-0.838	0.249	3.365	7.64E-04	8.82E-03
WDR59	24.086	-0.401	0.119	3.363	7.70E-04	8.88E-03
RN7SL587P	8.344	1.028	0.306	-3.362	7.74E-04	8.92E-03
PPIA	152.850	-0.483	0.144	3.361	7.77E-04	8.94E-03
NUP62	79.920	0.305	0.091	-3.361	7.77E-04	8.94E-03
KLHL29	6.461	0.776	0.231	-3.361	7.77E-04	8.94E-03
FBRS	47.446	0.418	0.124	-3.360	7.78E-04	8.94E-03
PGM5	3.016	-1.047	0.311	3.360	7.80E-04	8.95E-03
TTC39C	15.043	0.635	0.189	-3.360	7.80E-04	8.95E-03
ARFGAP1	123.298	0.425	0.127	-3.356	7.91E-04	9.07E-03
C11orf63	1.583	-1.062	0.316	3.356	7.91E-04	9.07E-03
PTMAP2	38.223	0.503	0.150	-3.355	7.94E-04	9.09E-03
CTD-2017C7.2	36.773	0.927	0.276	-3.354	7.96E-04	9.12E-03
MPST	64.336	-0.367	0.109	3.353	7.99E-04	9.13E-03
GSTO2	19.619	-0.436	0.130	3.353	8.01E-04	9.14E-03
SPATA4	0.854	1.536	0.458	-3.353	8.00E-04	9.14E-03
LMO4	59.389	0.442	0.132	-3.352	8.02E-04	9.15E-03
BLOC1S1	11.292	-0.569	0.170	3.352	8.02E-04	9.15E-03
UTRN	129.925	0.430	0.128	-3.352	8.03E-04	9.15E-03
DMTF1	62.354	0.489	0.146	-3.352	8.03E-04	9.15E-03
DOPEY2	33.014	0.451	0.134	-3.351	8.04E-04	9.15E-03
LCAT	19.954	0.733	0.219	-3.351	8.04E-04	9.15E-03
ZNF677	6.745	0.735	0.219	-3.349	8.11E-04	9.22E-03
CARD10	19.297	-0.452	0.135	3.349	8.12E-04	9.23E-03
CCSER2	14.760	0.474	0.142	-3.347	8.16E-04	9.27E-03
DARS	245.122	0.397	0.119	-3.347	8.18E-04	9.27E-03
VEGFC	0.900	1.422	0.425	-3.347	8.18E-04	9.27E-03
SETD5-AS1	64.561	0.472	0.141	-3.347	8.18E-04	9.27E-03
C19orf68	25.599	0.417	0.125	-3.345	8.24E-04	9.33E-03
DYNLRB1	45.063	-0.306	0.092	3.342	8.31E-04	9.40E-03
XAF1	62.449	0.440	0.132	-3.341	8.33E-04	9.42E-03
PDLIM1	99.294	-0.335	0.100	3.340	8.37E-04	9.46E-03
FAM214A	12.120	0.666	0.199	-3.340	8.37E-04	9.46E-03
DOCK2	20.987	0.738	0.221	-3.340	8.39E-04	9.47E-03
CTC-462L7.1	1.913	-1.041	0.312	3.337	8.47E-04	9.55E-03
GEMIN8	6.835	-0.634	0.190	3.336	8.50E-04	9.58E-03
KLHL32	0.903	-1.348	0.404	3.336	8.51E-04	9.58E-03
AIFM1	24.568	-0.423	0.127	3.336	8.51E-04	9.58E-03
SERPINE2	78.754	-0.351	0.105	3.334	8.56E-04	9.63E-03
ARPC1B	272.433	0.237	0.071	-3.333	8.59E-04	9.66E-03
PRKAA2	5.427	0.856	0.257	-3.333	8.59E-04	9.66E-03
CAV1	15.097	-0.481	0.144	3.333	8.60E-04	9.66E-03
CTD-2329K10.1	11.162	0.802	0.241	-3.333	8.60E-04	9.66E-03
SPATA5	17.974	0.514	0.154	-3.332	8.61E-04	9.66E-03
CARM1	132.561	0.251	0.075	-3.332	8.64E-04	9.69E-03
JKAMP	1.901	0.931	0.279	-3.330	8.68E-04	9.73E-03
C2orf68	54.403	0.478	0.144	-3.329	8.70E-04	9.75E-03
GPR137	20.026	0.782	0.235	-3.329	8.72E-04	9.77E-03
SYNE1	50.229	0.585	0.176	-3.328	8.74E-04	9.77E-03
ZBTB37	34.697	0.605	0.182	-3.328	8.74E-04	9.77E-03
STAR	10.705	0.915	0.275	-3.327	8.77E-04	9.80E-03
SPATA32	4.071	1.085	0.326	-3.327	8.78E-04	9.80E-03
ALG5	49.638	-0.348	0.105	3.327	8.79E-04	9.81E-03
AKAP6	2.455	-0.974	0.293	3.324	8.87E-04	9.90E-03
RNU7-41P	4.521	-0.790	0.238	3.324	8.88E-04	9.90E-03
DAB2IP	20.278	-0.573	0.172	3.323	8.90E-04	9.90E-03
KDM5A	198.726	0.325	0.098	-3.324	8.89E-04	9.90E-03
SLC20A2	8.855	0.515	0.155	-3.323	8.90E-04	9.90E-03
CPLX2	3.386	1.055	0.317	-3.322	8.93E-04	9.93E-03
GNL3L	61.930	0.493	0.148	-3.322	8.94E-04	9.93E-03
C4orf19	14.882	0.605	0.182	-3.320	8.99E-04	9.98E-03
CYP2D6	2.115	1.111	0.335	-3.320	9.01E-04	1.00E-02
TSR1	35.611	0.388	0.117	-3.319	9.04E-04	1.00E-02
FNDC3B	44.144	0.348	0.105	-3.318	9.07E-04	1.01E-02

MRFAP1L1	50.566	0.379	0.114	-3.318	9.08E-04	1.01E-02
TMEM158	25.638	-0.561	0.169	3.317	9.09E-04	1.01E-02
KAT6B	58.479	0.392	0.118	-3.317	9.09E-04	1.01E-02
LYVE1	9.953	1.158	0.349	-3.317	9.11E-04	1.01E-02
GMCL1	34.045	0.409	0.123	-3.316	9.13E-04	1.01E-02
EFNA4	8.234	-0.618	0.186	3.315	9.16E-04	1.01E-02
TMEM51	40.833	0.363	0.110	-3.314	9.20E-04	1.02E-02
CORO2A	46.537	0.413	0.125	-3.312	9.25E-04	1.02E-02
IER3	8.195	0.760	0.229	-3.312	9.26E-04	1.02E-02
GPATCH8	59.216	0.355	0.107	-3.311	9.31E-04	1.03E-02
FLT3LG	8.068	0.992	0.300	-3.309	9.35E-04	1.03E-02
SORBS1	10.718	-0.671	0.203	3.308	9.39E-04	1.03E-02
APMAP	43.101	-0.381	0.115	3.305	9.51E-04	1.04E-02
SHCBP1	15.399	0.623	0.189	-3.305	9.51E-04	1.04E-02
AKNA	120.278	0.520	0.157	-3.304	9.53E-04	1.05E-02
SLC39A13	17.335	-0.441	0.134	3.303	9.56E-04	1.05E-02
FOCAD	25.320	0.599	0.181	-3.302	9.60E-04	1.05E-02
SLC25A11	70.996	0.416	0.126	-3.302	9.61E-04	1.05E-02
SLC4A8	6.471	0.796	0.241	-3.301	9.62E-04	1.05E-02
CA11	4.999	0.822	0.249	-3.301	9.63E-04	1.05E-02
RAB3GAP2	84.419	0.324	0.098	-3.300	9.68E-04	1.06E-02
PAAF1	33.969	0.628	0.190	-3.299	9.71E-04	1.06E-02
TSSK4	2.319	1.080	0.327	-3.299	9.72E-04	1.06E-02
GTF2F1	45.130	0.445	0.135	-3.298	9.73E-04	1.06E-02
SMURF2	24.279	0.431	0.131	-3.298	9.75E-04	1.06E-02
NCAPG2	13.718	0.524	0.159	-3.296	9.81E-04	1.07E-02
PDE3B	22.120	0.635	0.193	-3.296	9.80E-04	1.07E-02
TPGS1	10.977	0.972	0.295	-3.296	9.81E-04	1.07E-02
CTD-2525I3.3	0.891	1.481	0.449	-3.296	9.82E-04	1.07E-02
MGRN1	47.481	0.360	0.109	-3.291	9.99E-04	1.09E-02
TNFRSF13B	21.731	0.768	0.233	-3.290	1.00E-03	1.09E-02
VWF	112.372	0.753	0.229	-3.289	1.01E-03	1.09E-02
PRPF4	42.955	0.387	0.118	-3.288	1.01E-03	1.09E-02
HDAC10	37.913	0.694	0.211	-3.287	1.01E-03	1.10E-02
ZDHHC20P1	1.009	1.402	0.427	-3.286	1.02E-03	1.10E-02
FAM215B	5.059	0.958	0.292	-3.285	1.02E-03	1.10E-02
NUBP2	52.298	0.553	0.169	-3.283	1.03E-03	1.11E-02
CYB5B	91.097	0.287	0.087	-3.283	1.03E-03	1.11E-02
ADIRF-AS1	3.545	0.838	0.255	-3.279	1.04E-03	1.12E-02
MEN1	68.939	-0.253	0.077	3.279	1.04E-03	1.12E-02
NDUFA13	19.500	0.724	0.221	-3.278	1.05E-03	1.13E-02
EXO5	2.063	1.021	0.311	-3.277	1.05E-03	1.13E-02
STOX2	1.842	-0.955	0.291	3.277	1.05E-03	1.13E-02
SH3RF3	9.126	0.790	0.241	-3.276	1.05E-03	1.13E-02
GABPB2	40.282	0.360	0.110	-3.275	1.05E-03	1.13E-02
ARID1A	77.395	0.336	0.103	-3.275	1.06E-03	1.13E-02
HOTTIP	11.647	-0.621	0.189	3.275	1.06E-03	1.13E-02
LNPEP	38.355	0.380	0.116	-3.274	1.06E-03	1.14E-02
ELAC2	36.480	0.354	0.108	-3.274	1.06E-03	1.14E-02
KB-1980E6.3	1.004	1.359	0.415	-3.273	1.07E-03	1.14E-02
HSPB1	267.265	-0.298	0.091	3.272	1.07E-03	1.14E-02
MSN	144.808	0.276	0.085	-3.272	1.07E-03	1.14E-02
DBP	9.521	0.879	0.269	-3.271	1.07E-03	1.15E-02
GLI1	6.679	0.713	0.218	-3.270	1.07E-03	1.15E-02
SPATA5L1	4.756	0.814	0.249	-3.269	1.08E-03	1.15E-02
RBM25	335.274	0.247	0.076	-3.267	1.09E-03	1.16E-02
IFIH1	8.683	0.588	0.180	-3.267	1.09E-03	1.16E-02
HIST1H3I	1.550	1.262	0.386	-3.267	1.09E-03	1.16E-02
LINC00106	63.425	0.684	0.210	-3.265	1.10E-03	1.17E-02
NFKB1	37.887	0.434	0.133	-3.263	1.10E-03	1.17E-02
PTPRE	49.542	0.409	0.125	-3.262	1.11E-03	1.18E-02
ZNF768	18.425	-0.373	0.114	3.261	1.11E-03	1.18E-02
SZT2	34.115	0.565	0.173	-3.260	1.11E-03	1.18E-02
SYDE2	4.696	-0.731	0.224	3.260	1.11E-03	1.18E-02
LONRF3	1.756	1.093	0.335	-3.259	1.12E-03	1.19E-02
FAM89B	9.602	-0.539	0.165	3.258	1.12E-03	1.19E-02
HAND2	3.370	-0.929	0.285	3.258	1.12E-03	1.19E-02
TAF8	28.716	0.479	0.147	-3.256	1.13E-03	1.20E-02
LCP2	37.788	0.590	0.181	-3.256	1.13E-03	1.20E-02
CEP63	31.757	0.465	0.143	-3.255	1.13E-03	1.20E-02
LRRC37A7P	1.157	1.198	0.368	-3.254	1.14E-03	1.21E-02
ALG13-AS1	10.244	0.617	0.190	-3.253	1.14E-03	1.21E-02
SMC2	33.607	-0.404	0.124	3.252	1.14E-03	1.21E-02
TGFBR1	10.939	-0.480	0.147	3.252	1.15E-03	1.21E-02
KRAS	39.019	0.388	0.119	-3.252	1.15E-03	1.21E-02
LRRC37A	0.533	1.770	0.545	-3.250	1.16E-03	1.22E-02
SIPA1	27.820	0.776	0.239	-3.249	1.16E-03	1.22E-02
STARD9	12.866	0.577	0.178	-3.247	1.17E-03	1.23E-02
SPP1	60.312	0.687	0.212	-3.246	1.17E-03	1.23E-02
HSD11B2	8.555	-0.633	0.195	3.244	1.18E-03	1.24E-02
SLC24A1	35.037	0.547	0.169	-3.244	1.18E-03	1.24E-02
MCMBP	17.066	0.479	0.148	-3.242	1.18E-03	1.24E-02
GBA2	25.416	0.618	0.191	-3.242	1.19E-03	1.25E-02
SAFB	36.751	0.383	0.118	-3.240	1.19E-03	1.25E-02
RBBP5	34.677	0.427	0.132	-3.241	1.19E-03	1.25E-02
FNDC3A	21.451	0.443	0.137	-3.240	1.20E-03	1.25E-02

ZSCAN29	33.405	0.462	0.143	-3.238	1 20E-03	1.26E-02
NMT2	46.270	0.483	0.149	-3.236	1 21E-03	1.27E-02
NUAK1	15.257	-0.553	0.171	3.235	1 22E-03	1.27E-02
PARP12	22.655	0.437	0.135	-3.235	1 22E-03	1.27E-02
RORA	47.811	0.480	0.148	-3.234	1 22E-03	1.27E-02
CHRM3	9.358	-0.699	0.216	3.234	1 22E-03	1.28E-02
PRRX1	26.765	0.521	0.161	-3.232	1 23E-03	1.28E-02
HELLS	58.859	0.317	0.098	-3.232	1 23E-03	1.28E-02
KIAA1919	27.283	0.508	0.157	-3.230	1 24E-03	1.29E-02
SLC41A3	11.972	0.569	0.176	-3.228	1 25E-03	1.30E-02
NT5C3A	38.462	0.300	0.093	-3.227	1 25E-03	1.30E-02
CREBZF	80.482	0.245	0.076	-3.223	1 27E-03	1.32E-02
CRIP1	16.395	0.480	0.149	-3.222	1 27E-03	1.32E-02
ENTPD6	7.212	0.680	0.211	-3.221	1 28E-03	1.32E-02
NUDT22	6.442	0.854	0.265	-3.220	1 28E-03	1.33E-02
CLEC7A	4.127	0.827	0.257	-3.220	1 28E-03	1.33E-02
DPEP1	51.938	0.517	0.161	-3.220	1 28E-03	1.33E-02
CLIP3	17.218	0.810	0.252	-3.218	1 29E-03	1.34E-02
RRNAD1	6.440	-0.635	0.198	3.217	1 30E-03	1.34E-02
MED4	67.648	-0.346	0.108	3.215	1 30E-03	1.35E-02
MSNP1	11.818	0.754	0.235	-3.216	1 30E-03	1.35E-02
ZSWIM8	17.151	0.610	0.190	-3.215	1 30E-03	1.35E-02
MGME1	12.460	-0.466	0.145	3.214	1 31E-03	1.35E-02
CCNDBP1	34.611	0.418	0.130	-3.214	1 31E-03	1.35E-02
ARL1	42.839	-0.292	0.091	3.214	1 31E-03	1.35E-02
IMP4	41.097	0.457	0.142	-3.212	1 32E-03	1.36E-02
RGS16	37.121	0.560	0.174	-3.212	1 32E-03	1.36E-02
U62631.5	7.143	0.932	0.290	-3.211	1 32E-03	1.36E-02
GSPT1	136.108	-0.175	0.055	3.209	1 33E-03	1.37E-02
BRD9	136.258	0.402	0.125	-3.206	1 35E-03	1.38E-02
FLAD1	30.724	0.527	0.164	-3.206	1 35E-03	1.38E-02
U2AF1	205.042	0.207	0.065	-3.205	1 35E-03	1.38E-02
RAD9A	20.125	0.703	0.219	-3.205	1 35E-03	1.38E-02
EGLN3	11.227	0.620	0.194	-3.204	1 35E-03	1.39E-02
KCTD20	41.607	0.389	0.121	-3.203	1 36E-03	1.39E-02
GLYR1	6.319	0.796	0.248	-3.203	1 36E-03	1.39E-02
USP50	8.488	0.893	0.279	-3.202	1 36E-03	1.39E-02
EXOG	39.707	0.538	0.168	-3.199	1 38E-03	1.41E-02
TRAPPC12	43.947	0.457	0.143	-3.199	1 38E-03	1.41E-02
MEGF9	7.247	0.646	0.202	-3.199	1 38E-03	1.41E-02
PPM1L	5.144	-0.719	0.225	3.198	1 38E-03	1.41E-02
RAD51-AS1	18.968	0.648	0.203	-3.198	1 38E-03	1.41E-02
STRBP	98.210	0.364	0.114	-3.197	1 39E-03	1.41E-02
MIR30E	0.542	1.920	0.601	-3.197	1 39E-03	1.41E-02
ZNF506	35.745	0.516	0.161	-3.196	1 39E-03	1.42E-02
ZCCHC2	30.545	0.435	0.136	-3.195	1.40E-03	1.42E-02
DYNLT3	11.574	-0.493	0.154	3.194	1.40E-03	1.42E-02
PROSER1	15.012	-0.504	0.158	3.194	1.41E-03	1.43E-02
POLK	27.727	0.487	0.152	-3.194	1.40E-03	1.43E-02
MFSD10	69.469	-0.275	0.086	3.193	1.41E-03	1.43E-02
NPTX2	2.997	1.106	0.346	-3.192	1.41E-03	1.43E-02
FRMD4A	44.611	0.521	0.163	-3.189	1.43E-03	1.45E-02
CSTF2T	18.330	0.472	0.148	-3.186	1.44E-03	1.46E-02
RBP5	13.022	0.746	0.234	-3.185	1.45E-03	1.46E-02
F10	1.247	1.178	0.370	-3.183	1.46E-03	1.47E-02
IL411	13.879	0.710	0.223	-3.182	1.47E-03	1.48E-02
TRAJ20	12.098	1.061	0.333	-3.181	1.47E-03	1.48E-02
KLHL6	37.620	0.716	0.225	-3.181	1.47E-03	1.48E-02
MILR1	10.960	0.778	0.245	-3.179	1.48E-03	1.49E-02
CES4A	14.733	0.865	0.272	-3.179	1.48E-03	1.49E-02
THOC7-AS1	4.438	0.891	0.280	-3.179	1.48E-03	1.49E-02
SRF	54.926	-0.477	0.150	3.176	1.49E-03	1.50E-02
G3BP2	43.919	0.333	0.105	-3.176	1.49E-03	1.50E-02
SOCS3	6.440	0.606	0.191	-3.176	1.49E-03	1.50E-02
ZNF397	50.356	0.671	0.211	-3.175	1 50E-03	1.51E-02
PARVB	34.048	0.541	0.171	-3.174	1 50E-03	1.51E-02
KCNQ1-AS1	1.272	-1.123	0.354	3.173	1 51E-03	1.52E-02
DCLRE1C	76.307	0.497	0.157	-3.172	1 51E-03	1.52E-02
MIR126	0.468	1.777	0.560	-3.171	1 52E-03	1.52E-02
PGM2	8.264	-0.627	0.198	3.170	1 53E-03	1.53E-02
ZNF271	44.577	0.408	0.129	-3.170	1 53E-03	1.53E-02
GABRR2	0.731	1.367	0.431	-3.169	1 53E-03	1.53E-02
ATP7A	6.559	0.635	0.201	-3.169	1 53E-03	1.53E-02
MAPK13	39.511	-0.399	0.126	3.166	1 55E-03	1.55E-02
SNORA84	0.733	1.418	0.448	-3.166	1 55E-03	1.55E-02
ZAP70	79.754	0.931	0.294	-3.165	1 55E-03	1.55E-02
IL1R2	16.845	0.575	0.182	-3.165	1 55E-03	1.55E-02
C1orf222	4.011	1.033	0.327	-3.164	1 56E-03	1.55E-02
DERL1	23.126	0.413	0.131	-3.163	1 56E-03	1.56E-02
COL6A1	319.888	-0.351	0.111	3.162	1 57E-03	1.56E-02
RNU4-38P	1.599	-1.137	0.360	3.161	1 57E-03	1.57E-02
DHX9	50.243	0.337	0.107	-3.158	1 59E-03	1.58E-02
JAG1	24.908	0.408	0.129	-3.158	1 59E-03	1.58E-02
ARHGEF9-IT1	0.564	1.936	0.613	-3.158	1 59E-03	1.58E-02
ACSF3	33.719	0.524	0.166	-3.157	1 59E-03	1.58E-02
FAM3D	47.307	-0.542	0.172	3.157	1.60E-03	1.58E-02

PPP1R9B	34.454	0.432	0.137	-3.157	1.59E-03	1.58E-02
ATPAF2	18.300	0.557	0.176	-3.157	1.60E-03	1.58E-02
RNF125	15.209	0.708	0.224	-3.155	1.61E-03	1.59E-02
SLX4	6.284	0.767	0.243	-3.154	1.61E-03	1.59E-02
OPTN	111.916	0.280	0.089	-3.154	1.61E-03	1.59E-02
OAZ2	96.466	0.288	0.091	-3.153	1.62E-03	1.60E-02
SEP9	182.392	-0.249	0.079	3.152	1.62E-03	1.60E-02
HARS	7.076	0.707	0.224	-3.151	1.63E-03	1.61E-02
USP18	8.689	0.712	0.226	-3.151	1.63E-03	1.61E-02
ZNF619	3.855	0.799	0.254	-3.149	1.64E-03	1.61E-02
WDR45B	66.193	-0.380	0.121	3.148	1.64E-03	1.62E-02
HIVEP2	24.693	0.443	0.141	-3.149	1.64E-03	1.62E-02
CTB-181H17.1	12.776	0.592	0.188	-3.148	1.64E-03	1.62E-02
CKMT2-AS1	12.645	0.751	0.239	-3.148	1.64E-03	1.62E-02
TM9SF1	27.311	0.565	0.179	-3.148	1.65E-03	1.62E-02
NREP	7.661	-0.675	0.214	3.147	1.65E-03	1.62E-02
POLA1	32.560	0.425	0.135	-3.147	1.65E-03	1.62E-02
SOX11	1.037	-1.167	0.371	3.144	1.67E-03	1.64E-02
ZC3H7A	43.082	0.412	0.131	-3.143	1.67E-03	1.64E-02
SLC16A4	15.500	0.537	0.171	-3.142	1.68E-03	1.64E-02
ZNF682	8.831	0.744	0.237	-3.142	1.68E-03	1.64E-02
CTA-292E10.6	28.935	0.680	0.217	-3.141	1.68E-03	1.65E-02
TAS2R5	0.426	1.974	0.629	-3.141	1.68E-03	1.65E-02
CDHR1	14.089	0.687	0.219	-3.140	1.69E-03	1.65E-02
ARHGAP39	4.682	0.789	0.251	-3.140	1.69E-03	1.65E-02
CTA-360L10.1	6.435	1.055	0.336	-3.140	1.69E-03	1.65E-02
ZNF644	26.825	0.392	0.125	-3.139	1.70E-03	1.66E-02
ABCG1	26.001	0.497	0.158	-3.139	1.70E-03	1.66E-02
MLXIP	186.403	0.301	0.096	-3.138	1.70E-03	1.66E-02
DHRS4-AS1	6.445	0.612	0.195	-3.138	1.70E-03	1.66E-02
RET	0.749	-1.557	0.497	3.134	1.72E-03	1.68E-02
CYFIP1	24.574	0.418	0.134	-3.133	1.73E-03	1.69E-02
GSTZ1	26.227	0.457	0.146	-3.132	1.73E-03	1.69E-02
EIF6	217.322	0.302	0.096	-3.132	1.74E-03	1.69E-02
ATP2A3	183.061	0.557	0.178	-3.131	1.74E-03	1.69E-02
SLC25A6P5	0.569	1.695	0.541	-3.131	1.74E-03	1.69E-02
C12orf65	25.351	0.377	0.120	-3.131	1.74E-03	1.69E-02
AP2M1	135.494	0.230	0.073	-3.130	1.75E-03	1.70E-02
GRK6	64.664	0.508	0.162	-3.130	1.75E-03	1.70E-02
ABCG5	0.847	-1.259	0.402	3.129	1.75E-03	1.70E-02
GPXKOW	7.123	-0.613	0.196	3.129	1.75E-03	1.70E-02
IGSF9B	13.905	0.732	0.234	-3.129	1.76E-03	1.70E-02
SCAI	23.322	0.631	0.202	-3.128	1.76E-03	1.71E-02
DTYMK	16.549	0.521	0.167	-3.127	1.77E-03	1.71E-02
TIMM17B	18.488	-0.473	0.151	3.126	1.77E-03	1.71E-02
CEBPZ-AS1	57.926	0.333	0.106	-3.126	1.77E-03	1.71E-02
PPIL2	8.329	0.758	0.243	-3.125	1.78E-03	1.72E-02
PLXNA3	55.691	0.384	0.123	-3.124	1.78E-03	1.72E-02
ZNF703	150.991	-0.318	0.102	3.123	1.79E-03	1.73E-02
PPFIA1	50.204	0.308	0.099	-3.123	1.79E-03	1.73E-02
THBS4	3.975	-0.856	0.274	3.122	1.79E-03	1.73E-02
SMAP1	7.921	-0.587	0.188	3.122	1.80E-03	1.73E-02
UBXN7	61.192	0.232	0.074	-3.121	1.80E-03	1.74E-02
VPS37A	6.348	-0.595	0.191	3.121	1.80E-03	1.74E-02
ECHS1	207.442	0.236	0.076	-3.119	1.82E-03	1.75E-02
NRG2	1.933	0.974	0.313	-3.117	1.83E-03	1.75E-02
FMO5	7.746	0.655	0.210	-3.117	1.83E-03	1.76E-02
TNRC18	40.310	0.461	0.148	-3.116	1.84E-03	1.76E-02
RRN3P3	27.992	0.475	0.153	-3.114	1.84E-03	1.77E-02
RNF144B	11.266	0.736	0.236	-3.114	1.85E-03	1.77E-02
CD84	59.269	0.586	0.188	-3.114	1.85E-03	1.77E-02
TNRC6B	180.179	0.312	0.100	-3.113	1.85E-03	1.78E-02
PRDX3	22.846	0.511	0.164	-3.111	1.86E-03	1.78E-02
WDFY4	58.158	0.753	0.242	-3.111	1.86E-03	1.78E-02
GPC1	8.705	-0.668	0.215	3.111	1.86E-03	1.78E-02
CLIP1	86.634	0.281	0.090	-3.110	1.87E-03	1.79E-02
RNH1	42.730	-0.289	0.093	3.109	1.88E-03	1.79E-02
HEATR5A	13.550	0.532	0.171	-3.109	1.88E-03	1.79E-02
EVL	492.000	0.693	0.223	-3.108	1.88E-03	1.80E-02
S100PBP	73.224	0.387	0.124	-3.108	1.88E-03	1.80E-02
APOE	175.706	0.523	0.168	-3.107	1.89E-03	1.80E-02
MPI	26.028	0.595	0.192	-3.107	1.89E-03	1.80E-02
OR10A5	0.867	1.558	0.501	-3.107	1.89E-03	1.80E-02
HP1BP3	53.503	-0.281	0.091	3.107	1.89E-03	1.80E-02
PIEZO1	14.906	0.495	0.159	-3.106	1.89E-03	1.80E-02
ZBTB32	10.876	1.009	0.325	-3.107	1.89E-03	1.80E-02
MUC20	44.601	0.449	0.145	-3.104	1.91E-03	1.81E-02
MAML2	40.523	0.353	0.114	-3.102	1.92E-03	1.83E-02
KLHL36	67.387	0.328	0.106	-3.101	1.93E-03	1.83E-02
RARS2	24.327	-0.352	0.114	3.101	1.93E-03	1.83E-02
TMA7	98.224	-0.362	0.117	3.100	1.94E-03	1.83E-02
IKBK	1.096	1.659	0.535	-3.100	1.94E-03	1.83E-02
VRK2	12.738	0.568	0.183	-3.100	1.94E-03	1.83E-02
EMC3-AS1	21.319	0.673	0.217	-3.099	1.94E-03	1.84E-02
HIST1H3A	1.955	1.102	0.356	-3.099	1.94E-03	1.84E-02
SLC6A14	1.308	-1.119	0.361	3.098	1.95E-03	1.84E-02

BRD7P2	3.243	0.689	0.222	-3.098	1.95E-03	1.84E-02
GOLGA6L5P	5.010	0.716	0.231	-3.098	1.95E-03	1.84E-02
GPR56	29.637	-0.495	0.160	3.097	1.95E-03	1.84E-02
RGS1	21.392	0.444	0.143	-3.097	1.95E-03	1.84E-02
CABP1	1.351	1.104	0.357	-3.097	1.95E-03	1.84E-02
PARP1	104.787	0.348	0.112	-3.097	1.96E-03	1.85E-02
CTD-2085J24.4	2.479	-0.811	0.262	3.096	1.96E-03	1.85E-02
PRPS1	22.871	-0.415	0.134	3.096	1.96E-03	1.85E-02
TUBD1	14.241	0.560	0.181	-3.096	1.96E-03	1.85E-02
CTD-2196E14.6	6.614	0.616	0.199	-3.096	1.96E-03	1.85E-02
SGSM1	4.044	0.767	0.248	-3.096	1.96E-03	1.85E-02
NDUFV2	2.067	0.938	0.303	-3.095	1.97E-03	1.85E-02
RNY4	1.107	-1.111	0.359	3.095	1.97E-03	1.85E-02
GTF2E2	19.885	-0.429	0.139	3.093	1.98E-03	1.86E-02
PWWP2B	18.228	-0.404	0.131	3.093	1.98E-03	1.86E-02
SHARPIN	23.682	0.501	0.162	-3.092	1.99E-03	1.86E-02
SMPD3	14.581	0.709	0.229	-3.091	1.99E-03	1.87E-02
UBR2	62.064	0.450	0.145	-3.091	2.00E-03	1.87E-02
SCNN1A	47.473	0.403	0.130	-3.089	2.01E-03	1.88E-02
RAMP2-AS1	7.022	1.034	0.335	-3.089	2.01E-03	1.88E-02
PDDC1	95.296	0.277	0.090	-3.088	2.02E-03	1.88E-02
DNAJC15	27.318	-0.491	0.159	3.086	2.03E-03	1.89E-02
SAMHD1	111.060	0.357	0.116	-3.087	2.03E-03	1.89E-02
GRIA4	6.131	0.963	0.312	-3.086	2.03E-03	1.89E-02
IP6K2	41.074	-0.351	0.114	3.086	2.03E-03	1.89E-02
PSMB8	49.389	-0.315	0.102	3.083	2.05E-03	1.91E-02
RASGEF1B	11.529	0.626	0.203	-3.083	2.05E-03	1.91E-02
GGA1	112.237	0.356	0.116	-3.082	2.06E-03	1.91E-02
RSAD2	5.025	0.723	0.235	-3.081	2.06E-03	1.92E-02
MLYCD	23.971	0.443	0.144	-3.081	2.06E-03	1.92E-02
FAM106A	1.812	1.172	0.380	-3.081	2.06E-03	1.92E-02
KCNQ1OT1	93.681	0.304	0.099	-3.081	2.07E-03	1.92E-02
SCARNA20	0.658	1.472	0.478	-3.080	2.07E-03	1.92E-02
TYROBP	70.452	0.460	0.149	-3.079	2.08E-03	1.93E-02
MIR4459	5.057	0.861	0.280	-3.079	2.08E-03	1.93E-02
SLC26A3	9.298	-0.903	0.293	3.079	2.08E-03	1.93E-02
RNF34	28.863	0.311	0.101	-3.079	2.08E-03	1.93E-02
LPL	13.816	0.560	0.182	-3.079	2.08E-03	1.93E-02
SEPT10P1	0.525	1.364	0.443	-3.078	2.08E-03	1.93E-02
CTB-102L5.8	4.899	0.989	0.321	-3.078	2.08E-03	1.93E-02
FBXO25	19.291	-0.423	0.138	3.076	2.10E-03	1.94E-02
RNF214	19.593	0.406	0.132	-3.075	2.10E-03	1.94E-02
HAT1	21.512	0.444	0.144	-3.075	2.11E-03	1.95E-02
LINC00924	0.753	1.458	0.475	-3.072	2.12E-03	1.96E-02
PHKG2	41.755	0.500	0.163	-3.072	2.13E-03	1.96E-02
CSTF3	27.619	0.327	0.106	-3.070	2.14E-03	1.97E-02
TVP23B	6.248	0.729	0.237	-3.070	2.14E-03	1.97E-02
HOXC6	14.901	0.599	0.195	-3.070	2.14E-03	1.97E-02
PDP2	2.528	0.950	0.310	-3.069	2.15E-03	1.98E-02
NAALAD2	1.352	-0.960	0.313	3.068	2.16E-03	1.98E-02
TMEM192	53.332	0.429	0.140	-3.067	2.16E-03	1.99E-02
SNORD46	22.439	0.756	0.247	-3.066	2.17E-03	1.99E-02
GK-IT1	2.938	0.873	0.285	-3.066	2.17E-03	1.99E-02
HIST1H1E	15.925	0.670	0.219	-3.065	2.18E-03	2.00E-02
SKP1	135.490	-0.191	0.062	3.064	2.18E-03	2.00E-02
RUNDC1	38.639	0.385	0.126	-3.064	2.18E-03	2.00E-02
MCAT	4.559	-0.790	0.258	3.062	2.20E-03	2.01E-02
TMEM81	1.554	-0.997	0.326	3.061	2.21E-03	2.02E-02
SS18L2	16.785	-0.382	0.125	3.060	2.21E-03	2.03E-02
KRT18P5	0.564	1.418	0.464	-3.060	2.22E-03	2.03E-02
MBTD1	6.127	0.571	0.187	-3.058	2.23E-03	2.04E-02
TUSC1	3.812	-0.750	0.245	3.057	2.23E-03	2.04E-02
RASGRP4	1.036	1.063	0.348	-3.057	2.24E-03	2.04E-02
GTF2IRD1	19.280	-0.477	0.156	3.056	2.24E-03	2.05E-02
LMO3	3.330	-0.902	0.295	3.056	2.25E-03	2.05E-02
WBP1L	35.668	0.389	0.127	-3.055	2.25E-03	2.05E-02
ZNF44	22.813	0.499	0.163	-3.055	2.25E-03	2.05E-02
ZNF276	25.175	0.603	0.197	-3.053	2.27E-03	2.06E-02
PKM	291.879	0.290	0.095	-3.052	2.27E-03	2.07E-02
ATP5S	22.318	0.493	0.162	-3.052	2.27E-03	2.07E-02
XXYL1-AS2	0.874	-1.164	0.381	3.051	2.28E-03	2.07E-02
CPED1	5.669	-0.634	0.208	3.051	2.28E-03	2.07E-02
HDC	6.969	0.927	0.304	-3.050	2.29E-03	2.08E-02
SUPT6H	62.697	0.408	0.134	-3.050	2.29E-03	2.08E-02
NAT9	15.659	0.566	0.185	-3.050	2.29E-03	2.08E-02
PHKB	52.502	0.340	0.112	-3.049	2.29E-03	2.08E-02
SAMD3	3.511	0.858	0.282	-3.047	2.31E-03	2.10E-02
ESPN	10.142	0.705	0.232	-3.047	2.31E-03	2.10E-02
SLC2A1	40.167	0.487	0.160	-3.046	2.32E-03	2.10E-02
RREB1	83.184	0.349	0.115	-3.046	2.32E-03	2.10E-02
CD28	4.965	0.830	0.272	-3.046	2.32E-03	2.10E-02
CTD-2017C7.1	10.485	0.834	0.274	-3.045	2.32E-03	2.10E-02
TECPR1	80.568	0.524	0.172	-3.045	2.33E-03	2.10E-02
EEF1D	68.920	0.338	0.111	-3.044	2.33E-03	2.11E-02
SPOP	253.402	0.366	0.120	-3.044	2.33E-03	2.11E-02
KIAA1429	37.674	0.377	0.124	-3.044	2.33E-03	2.11E-02

UQCC2	34.211	0.396	0.130	-3.044	2.34E-03	2.11E-02
C9orf37	24.762	0.500	0.164	-3.043	2.34E-03	2.11E-02
BCAS2P2	0.708	1.502	0.494	-3.043	2.34E-03	2.11E-02
PPM1K	90.848	0.524	0.172	-3.043	2.34E-03	2.11E-02
SPNS1	16.859	0.667	0.219	-3.042	2.35E-03	2.11E-02
UBE2S	3.440	-0.853	0.281	3.042	2.35E-03	2.11E-02
HIST1H3C	2.446	1.071	0.352	-3.042	2.35E-03	2.11E-02
PEBP1	368.839	-0.208	0.068	3.042	2.35E-03	2.12E-02
DHX8	19.982	0.431	0.142	-3.038	2.38E-03	2.14E-02
MOB2	30.243	0.484	0.159	-3.038	2.38E-03	2.14E-02
SOX4	370.819	0.243	0.080	-3.037	2.39E-03	2.14E-02
RBMS3	18.157	0.505	0.166	-3.036	2.39E-03	2.15E-02
CPSF1	30.074	0.512	0.169	-3.036	2.39E-03	2.15E-02
TMEM47	7.910	-0.540	0.178	3.035	2.41E-03	2.16E-02
CTD-2118P12.1	0.531	1.493	0.492	-3.034	2.41E-03	2.16E-02
PLEKHB1	5.914	-0.686	0.226	3.034	2.41E-03	2.16E-02
LRBA	94.052	0.321	0.106	-3.032	2.43E-03	2.17E-02
MCM5	95.514	0.526	0.173	-3.032	2.43E-03	2.17E-02
SKIV2L	38.976	0.547	0.180	-3.032	2.43E-03	2.17E-02
FER	17.943	0.535	0.176	-3.031	2.43E-03	2.17E-02
TEK	2.047	-0.929	0.306	3.031	2.44E-03	2.18E-02
EMC10	6.787	0.568	0.188	-3.030	2.44E-03	2.18E-02
ULK1	39.593	-0.326	0.108	3.030	2.45E-03	2.18E-02
HYAL1	5.102	0.879	0.290	-3.030	2.45E-03	2.18E-02
ELMSAN1	42.444	0.379	0.125	-3.030	2.45E-03	2.18E-02
NDUFS6	180.389	0.350	0.115	-3.029	2.45E-03	2.18E-02
ROCK1	111.293	0.227	0.075	-3.029	2.45E-03	2.18E-02
FOXS1	5.962	0.671	0.222	-3.028	2.46E-03	2.19E-02
GBP4	27.775	0.507	0.168	-3.027	2.47E-03	2.19E-02
CCNK	11.341	0.567	0.187	-3.027	2.47E-03	2.19E-02
NRXN2	3.949	0.779	0.257	-3.027	2.47E-03	2.20E-02
BDH1	22.276	-0.402	0.133	3.026	2.48E-03	2.20E-02
ADCK4	22.527	0.576	0.190	-3.025	2.49E-03	2.20E-02
MYOCD	1.590	-1.209	0.400	3.025	2.49E-03	2.21E-02
SCIMP	80.275	0.769	0.254	-3.024	2.49E-03	2.21E-02
PRKCSH	32.182	0.390	0.129	-3.023	2.50E-03	2.22E-02
CHD3	269.011	0.415	0.137	-3.023	2.51E-03	2.22E-02
DKK4	1.507	1.234	0.408	-3.023	2.50E-03	2.22E-02
EIF3D	169.440	-0.191	0.063	3.020	2.53E-03	2.24E-02
EIF3KP3	1.035	1.171	0.388	-3.018	2.55E-03	2.25E-02
SEP7	130.606	0.198	0.066	-3.017	2.55E-03	2.25E-02
NPHP1	3.101	0.796	0.264	-3.017	2.55E-03	2.25E-02
ADAMTSL5	1.450	-0.935	0.310	3.014	2.58E-03	2.27E-02
GUCD1	58.275	-0.291	0.096	3.014	2.58E-03	2.27E-02
ZDHHC13	6.288	0.574	0.190	-3.014	2.58E-03	2.27E-02
GRM8	1.381	-1.358	0.451	3.013	2.59E-03	2.28E-02
RRAGD	9.039	0.573	0.190	-3.012	2.60E-03	2.29E-02
ULK3	30.086	-0.341	0.113	3.011	2.60E-03	2.29E-02
ZNF530	6.218	0.862	0.286	-3.011	2.60E-03	2.29E-02
DSEL	2.526	0.914	0.303	-3.011	2.60E-03	2.29E-02
CNTNAP3B	2.927	0.940	0.312	-3.011	2.60E-03	2.29E-02
VIM	104.022	0.310	0.103	-3.011	2.60E-03	2.29E-02
SPON2	187.059	-0.326	0.108	3.011	2.60E-03	2.29E-02
DIMT1	30.195	-0.312	0.104	3.010	2.61E-03	2.29E-02
PCNX	76.482	0.447	0.148	-3.010	2.61E-03	2.29E-02
OSTM1	26.139	0.539	0.179	-3.010	2.61E-03	2.29E-02
GTF2H2B	9.302	0.710	0.236	-3.010	2.61E-03	2.29E-02
GOLGA6L4	0.873	1.086	0.361	-3.010	2.61E-03	2.29E-02
SLC43A3	9.412	0.729	0.242	-3.009	2.62E-03	2.30E-02
DHRS7	23.956	0.424	0.141	-3.008	2.63E-03	2.30E-02
SNORA2A	3.998	1.026	0.341	-3.008	2.63E-03	2.30E-02
CISH	10.471	0.562	0.187	-3.007	2.63E-03	2.30E-02
ZNRD1	9.231	-0.483	0.161	3.007	2.64E-03	2.31E-02
ZNF233	14.426	0.751	0.250	-3.007	2.64E-03	2.31E-02
PATL1	21.427	0.402	0.134	-3.007	2.64E-03	2.31E-02
RFC5	9.898	-0.537	0.179	3.006	2.65E-03	2.31E-02
TMED8	58.681	0.343	0.114	-3.005	2.65E-03	2.31E-02
DDX3X	98.117	0.234	0.078	-3.004	2.66E-03	2.32E-02
YPEL4	0.596	1.639	0.546	-3.003	2.67E-03	2.33E-02
TSNARE1	0.717	-1.392	0.463	3.002	2.68E-03	2.33E-02
WDR5	25.420	-0.455	0.152	3.003	2.68E-03	2.33E-02
YWHAB	1517.264	0.216	0.072	-3.003	2.68E-03	2.33E-02
CTA-221G9.11	14.132	0.723	0.241	-3.002	2.68E-03	2.33E-02
STAG3L1	2.839	0.851	0.283	-3.003	2.68E-03	2.33E-02
OXTR	2.295	0.855	0.285	-3.003	2.68E-03	2.33E-02
KIAA2013	61.122	-0.333	0.111	3.002	2.68E-03	2.33E-02
PSMD8	179.598	0.169	0.056	-3.002	2.69E-03	2.33E-02
VMA21	63.310	-0.316	0.105	3.001	2.69E-03	2.33E-02
ANKLE2	60.135	0.247	0.082	-2.998	2.72E-03	2.36E-02
LZTS1	9.057	0.710	0.237	-2.998	2.72E-03	2.36E-02
COPS5	13.890	0.501	0.167	-2.997	2.73E-03	2.36E-02
PPP1R16B	5.764	0.857	0.286	-2.996	2.73E-03	2.37E-02
PER2	13.946	0.467	0.156	-2.996	2.73E-03	2.37E-02
CYLD	114.871	0.414	0.138	-2.995	2.74E-03	2.37E-02
BRS3	1.170	-1.145	0.382	2.994	2.75E-03	2.38E-02
NEO1	36.068	-0.456	0.152	2.994	2.75E-03	2.38E-02

DEK	193.717	0.189	0.063	-2.993	2.76E-03	2.39E-02
ATRIP	1.609	-1.025	0.342	2.992	2.77E-03	2.39E-02
JPH2	5.274	-0.754	0.252	2.992	2.77E-03	2.39E-02
HES6	22.128	-0.591	0.197	2.992	2.77E-03	2.39E-02
HADHA	130.945	0.256	0.086	-2.992	2.77E-03	2.39E-02
NF2	10.611	0.560	0.187	-2.991	2.78E-03	2.40E-02
CCDC88B	74.275	0.527	0.176	-2.990	2.79E-03	2.41E-02
TTC26	13.810	0.557	0.186	-2.989	2.80E-03	2.41E-02
FAM211B	3.490	0.810	0.271	-2.989	2.80E-03	2.41E-02
HOXB9	71.589	-0.543	0.182	2.988	2.81E-03	2.42E-02
KIAA0196-AS1	0.526	2.000	0.669	-2.988	2.81E-03	2.42E-02
SUZ12P	33.189	0.472	0.158	-2.987	2.82E-03	2.42E-02
CD163	20.055	0.500	0.167	-2.987	2.82E-03	2.42E-02
HMGXB3	5.534	0.728	0.244	-2.985	2.83E-03	2.43E-02
MAP3K8	11.462	0.523	0.175	-2.985	2.84E-03	2.44E-02
ALG10	1.277	1.150	0.385	-2.985	2.84E-03	2.44E-02
LEMD2	23.565	-0.450	0.151	2.984	2.84E-03	2.44E-02
AKR1A1	79.203	0.377	0.126	-2.984	2.84E-03	2.44E-02
SLC16A5	2.779	0.805	0.270	-2.984	2.84E-03	2.44E-02
JAK2	25.240	0.393	0.132	-2.983	2.85E-03	2.44E-02
TMEM171	2.593	0.908	0.304	-2.983	2.85E-03	2.44E-02
SHMT2	29.922	-0.347	0.116	2.983	2.85E-03	2.44E-02
MYH14	33.711	0.547	0.184	-2.982	2.86E-03	2.45E-02
TBC1D19	1.625	-1.101	0.369	2.982	2.86E-03	2.45E-02
ALDOA	461.366	0.199	0.067	-2.982	2.87E-03	2.45E-02
KNTC1	12.248	0.451	0.151	-2.981	2.87E-03	2.45E-02
DIP2B	30.870	0.318	0.107	-2.981	2.88E-03	2.45E-02
EIF3G	8.152	0.640	0.215	-2.981	2.88E-03	2.46E-02
CLDN11	2.522	1.066	0.358	-2.980	2.88E-03	2.46E-02
TNFAIP2	45.381	0.398	0.134	-2.980	2.88E-03	2.46E-02
ARHGEF9	11.502	0.588	0.197	-2.978	2.90E-03	2.47E-02
PTTG1IP	116.347	0.218	0.073	-2.978	2.91E-03	2.47E-02
ABCB4	32.854	0.898	0.302	-2.977	2.91E-03	2.47E-02
MYOF	20.735	0.489	0.164	-2.977	2.91E-03	2.48E-02
SOCS7	1.021	1.211	0.407	-2.975	2.93E-03	2.49E-02
PCF11	45.284	0.400	0.135	-2.974	2.94E-03	2.50E-02
AICDA	18.353	0.885	0.298	-2.974	2.94E-03	2.50E-02
HELZ2	48.189	0.326	0.110	-2.973	2.95E-03	2.51E-02
RNU4-82P	0.784	1.166	0.392	-2.972	2.95E-03	2.51E-02
UHMK1	31.648	0.315	0.106	-2.972	2.96E-03	2.51E-02
KLF7	27.263	-0.364	0.123	2.971	2.97E-03	2.51E-02
ILF3-AS1	68.344	0.465	0.157	-2.970	2.97E-03	2.52E-02
FAXC	1.841	-1.019	0.343	2.969	2.98E-03	2.53E-02
HIST2H2AC	14.277	0.605	0.204	-2.969	2.98E-03	2.53E-02
TDP2	8.256	0.535	0.180	-2.969	2.99E-03	2.53E-02
RAD54L2	66.119	0.355	0.120	-2.964	3.04E-03	2.57E-02
SLC5A6	60.390	-0.302	0.102	2.963	3.05E-03	2.57E-02
ZNF526	20.904	0.464	0.157	-2.962	3.06E-03	2.58E-02
VDAC1P6	2.428	0.893	0.302	-2.962	3.05E-03	2.58E-02
TRAC	397.627	0.653	0.220	-2.962	3.06E-03	2.58E-02
SDHAF2	25.582	-0.328	0.111	2.960	3.07E-03	2.59E-02
NPM1P27	7.978	0.561	0.189	-2.960	3.07E-03	2.59E-02
THAP9-AS1	8.690	-0.562	0.190	2.959	3.09E-03	2.60E-02
SCFD1	40.141	0.328	0.111	-2.957	3.11E-03	2.61E-02
SNRNP25	17.584	0.698	0.236	-2.956	3.11E-03	2.62E-02
NAB2	1.839	1.131	0.383	-2.956	3.12E-03	2.62E-02
TBC1D22A	51.678	0.397	0.135	-2.955	3.13E-03	2.63E-02
C1orf109	3.810	0.688	0.233	-2.954	3.13E-03	2.63E-02
LYRM2	36.669	0.317	0.107	-2.954	3.13E-03	2.63E-02
SF3B4	9.409	-0.590	0.200	2.953	3.15E-03	2.64E-02
ZNF211	2.729	-0.821	0.278	2.951	3.17E-03	2.65E-02
RAB11FIP5	29.850	-0.378	0.128	2.951	3.17E-03	2.66E-02
ABHD4	3.906	-0.749	0.254	2.949	3.19E-03	2.66E-02
STMN2	1.821	1.216	0.412	-2.949	3.19E-03	2.67E-02
TDP1	12.160	-0.454	0.154	2.948	3.20E-03	2.67E-02
ADCY7	6.362	0.741	0.251	-2.948	3.20E-03	2.67E-02
PSORS1C1	2.773	-0.894	0.304	2.945	3.23E-03	2.69E-02
ICOSLG	26.980	0.482	0.164	-2.945	3.23E-03	2.69E-02
NMUR1	1.619	-0.983	0.334	2.943	3.25E-03	2.71E-02
RNF216	59.307	0.389	0.132	-2.943	3.26E-03	2.71E-02
PRKAG2	23.146	0.528	0.180	-2.942	3.26E-03	2.72E-02
IL13RA1	51.016	-0.336	0.114	2.941	3.27E-03	2.72E-02
RIF1	100.056	0.210	0.071	-2.941	3.27E-03	2.72E-02
HLA-E	686.770	0.373	0.127	-2.940	3.28E-03	2.73E-02
MIR98	1.314	1.069	0.364	-2.939	3.29E-03	2.74E-02
GNA11	26.347	0.324	0.110	-2.938	3.30E-03	2.74E-02
ABI2	40.578	0.323	0.110	-2.938	3.30E-03	2.74E-02
SMC1A	140.170	0.262	0.089	-2.936	3.32E-03	2.76E-02
SEC24B	11.350	0.482	0.164	-2.936	3.33E-03	2.76E-02
PFDN5	710.023	0.312	0.106	-2.935	3.33E-03	2.76E-02
AGO3	62.327	0.359	0.122	-2.935	3.34E-03	2.76E-02
ADC	5.206	0.774	0.264	-2.935	3.33E-03	2.76E-02
ACO2	4.863	0.665	0.227	-2.934	3.34E-03	2.77E-02
ATG10	23.360	0.520	0.177	-2.934	3.35E-03	2.77E-02
SLC35D3	1.718	-1.146	0.391	2.934	3.35E-03	2.77E-02
TOR1AIP1	17.870	0.420	0.143	-2.934	3.35E-03	2.77E-02

BTNL9	28.480	0.442	0.151	-2.933	3.35E-03	2.77E-02
CDK10	43.741	0.397	0.135	-2.931	3.38E-03	2.79E-02
KLHDC1	3.013	0.792	0.270	-2.931	3.38E-03	2.79E-02
PSMG3	54.752	-0.299	0.102	2.930	3.39E-03	2.79E-02
ENTPD4	64.318	0.431	0.147	-2.930	3.39E-03	2.80E-02
QTRT1	34.610	-0.371	0.127	2.929	3.40E-03	2.80E-02
FBXL17	30.573	-0.335	0.114	2.929	3.40E-03	2.80E-02
HNRNPA1	202.126	-0.259	0.088	2.928	3.41E-03	2.81E-02
APLP1	2.550	0.878	0.300	-2.928	3.41E-03	2.81E-02
TRIM65	29.190	0.359	0.123	-2.928	3.41E-03	2.81E-02
ONECUT3	6.507	0.970	0.332	-2.926	3.43E-03	2.82E-02
SRGAP3	17.280	0.576	0.197	-2.926	3.44E-03	2.82E-02
ZNF626	5.846	0.798	0.273	-2.925	3.44E-03	2.83E-02
LGR4	10.059	-0.620	0.212	2.924	3.45E-03	2.83E-02
PHLDB1	8.246	0.586	0.200	-2.924	3.45E-03	2.83E-02
NEMF	103.105	0.261	0.089	-2.924	3.46E-03	2.83E-02
PDPR	42.195	0.401	0.137	-2.924	3.46E-03	2.84E-02
ZNF251	6.326	0.617	0.211	-2.923	3.47E-03	2.84E-02
CTD-2313N18.5	1.136	1.186	0.406	-2.923	3.47E-03	2.84E-02
BAZ1A	58.131	0.274	0.094	-2.922	3.47E-03	2.84E-02
SQSTM1	58.022	0.378	0.129	-2.922	3.48E-03	2.85E-02
CCDC142	34.465	0.513	0.175	-2.922	3.48E-03	2.85E-02
CDK4	1.645	1.038	0.355	-2.922	3.48E-03	2.85E-02
CTA-85E5.10	8.103	0.751	0.257	-2.921	3.49E-03	2.85E-02
TSC22D3	410.881	0.335	0.115	-2.920	3.50E-03	2.86E-02
RNF126	27.682	-0.364	0.125	2.920	3.50E-03	2.86E-02
PDE10A	3.456	0.820	0.281	-2.920	3.51E-03	2.86E-02
SMIM11	23.461	0.428	0.146	-2.919	3.51E-03	2.86E-02
ZNF417	18.776	0.642	0.220	-2.919	3.51E-03	2.86E-02
LINC00243	6.240	0.958	0.328	-2.919	3.52E-03	2.87E-02
ZNF787	45.769	0.425	0.146	-2.917	3.53E-03	2.87E-02
PPP1R7	34.109	0.409	0.140	-2.917	3.53E-03	2.88E-02
GPR82	6.308	0.651	0.223	-2.917	3.53E-03	2.88E-02
CAPG	102.996	0.291	0.100	-2.916	3.54E-03	2.88E-02
C9orf169	0.420	1.740	0.597	-2.916	3.55E-03	2.88E-02
CMTM8	10.348	-0.458	0.157	2.915	3.56E-03	2.89E-02
DPYSL4	1.443	1.099	0.377	-2.915	3.56E-03	2.89E-02
CALML4	6.262	-0.575	0.198	2.913	3.58E-03	2.91E-02
BIRC2	54.361	-0.275	0.095	2.912	3.60E-03	2.91E-02
OSGEPL1	38.022	0.552	0.190	-2.911	3.60E-03	2.92E-02
LINC00565	0.881	-1.091	0.375	2.911	3.60E-03	2.92E-02
RCL1	16.558	0.426	0.147	-2.909	3.62E-03	2.93E-02
COLEC11	1.666	1.011	0.348	-2.909	3.63E-03	2.93E-02
ZNF566	25.487	0.464	0.159	-2.909	3.63E-03	2.94E-02
USP39	35.791	0.312	0.107	-2.908	3.64E-03	2.94E-02
CTD-2006C1.12	6.674	0.888	0.305	-2.908	3.64E-03	2.94E-02
ATP6VOD1	57.515	0.346	0.119	-2.908	3.64E-03	2.94E-02
XXYLT1	29.583	0.383	0.132	-2.907	3.65E-03	2.95E-02
ZNF350	13.901	0.689	0.237	-2.907	3.65E-03	2.95E-02
TLR10	18.634	0.744	0.256	-2.904	3.68E-03	2.97E-02
PRPF8	64.902	0.285	0.098	-2.904	3.68E-03	2.97E-02
SRSF12	1.045	-1.148	0.395	2.903	3.70E-03	2.98E-02
HSP90AA5P	0.858	-1.109	0.382	2.903	3.70E-03	2.98E-02
PDZD2	4.681	0.745	0.257	-2.901	3.71E-03	2.99E-02
LINC01004	9.212	0.662	0.228	-2.899	3.75E-03	3.02E-02
PORCN	9.198	-0.491	0.169	2.898	3.76E-03	3.02E-02
GRHPR	38.488	0.334	0.115	-2.898	3.76E-03	3.02E-02
KAT8	26.701	0.460	0.159	-2.898	3.76E-03	3.02E-02
FAM131A	8.225	0.619	0.214	-2.898	3.75E-03	3.02E-02
FAM96B	54.156	-0.341	0.118	2.897	3.77E-03	3.03E-02
ATG4A	10.023	-0.461	0.159	2.897	3.77E-03	3.03E-02
LRR6	4.192	-0.832	0.287	2.894	3.81E-03	3.05E-02
EPS15L1	16.312	0.428	0.148	-2.894	3.81E-03	3.05E-02
IL21R	10.522	0.822	0.284	-2.894	3.81E-03	3.05E-02
ILKAP	42.333	-0.256	0.088	2.893	3.81E-03	3.05E-02
WASF2	315.782	0.352	0.122	-2.893	3.81E-03	3.05E-02
NECAP2	17.406	0.412	0.143	-2.892	3.83E-03	3.07E-02
RXRA	52.507	0.283	0.098	-2.890	3.85E-03	3.07E-02
SYMPK	124.365	0.286	0.099	-2.890	3.85E-03	3.07E-02
PRSS3	31.136	0.447	0.155	-2.890	3.85E-03	3.07E-02
BAX	63.350	-0.303	0.105	2.890	3.86E-03	3.08E-02
NUDT3	33.371	0.351	0.122	-2.889	3.86E-03	3.08E-02
ZNF589	35.514	0.546	0.189	-2.885	3.91E-03	3.11E-02
FAM73A	20.235	0.427	0.148	-2.885	3.92E-03	3.12E-02
TP53RK	10.837	-0.489	0.170	2.885	3.92E-03	3.12E-02
VAC14	12.054	0.507	0.176	-2.884	3.93E-03	3.12E-02
SLC16A10	11.512	0.513	0.178	-2.884	3.93E-03	3.12E-02
APBB1P	26.667	0.591	0.205	-2.884	3.93E-03	3.12E-02
SNX29P2	24.343	0.607	0.210	-2.883	3.94E-03	3.13E-02
CTC-308K20.1	2.134	0.949	0.329	-2.883	3.94E-03	3.13E-02
FAM46C	13.821	0.503	0.175	-2.882	3.95E-03	3.13E-02
UQCRH	261.273	-0.220	0.076	2.882	3.95E-03	3.14E-02
LOXL2	9.091	-0.550	0.191	2.880	3.97E-03	3.14E-02
CENPL	7.570	0.574	0.199	-2.880	3.97E-03	3.14E-02
RUVBL1	44.465	0.299	0.104	-2.880	3.98E-03	3.15E-02
ZNF416	1.972	-0.958	0.333	2.879	3.99E-03	3.15E-02

SRM	171.546	0.362	0.126	-2.879	3.99E-03	3.15E-02
SLC43A2	33.660	0.352	0.122	-2.878	4.00E-03	3.16E-02
SOCS4	7.252	-0.543	0.189	2.875	4.03E-03	3.18E-02
FAM180A	0.555	-1.559	0.542	2.875	4.04E-03	3.19E-02
PRSS22	5.555	-0.801	0.279	2.875	4.04E-03	3.19E-02
MARCH2	4.005	-0.727	0.253	2.874	4.05E-03	3.19E-02
B3GNT1	36.334	0.410	0.143	-2.874	4.05E-03	3.19E-02
SCNM1	37.483	0.443	0.154	-2.874	4.05E-03	3.19E-02
LRSAM1	19.199	0.598	0.208	-2.874	4.05E-03	3.19E-02
RAD1	43.227	0.379	0.132	-2.873	4.07E-03	3.20E-02
RN7SL9P	2.708	0.914	0.318	-2.873	4.07E-03	3.20E-02
PRRC2B	117.110	0.254	0.088	-2.871	4.09E-03	3.22E-02
NDUFA1	60.202	-0.358	0.125	2.870	4.10E-03	3.22E-02
EMR2	7.216	0.659	0.230	-2.870	4.10E-03	3.22E-02
CTC-429L19.3	6.221	0.899	0.313	-2.870	4.11E-03	3.23E-02
RN7SL349P	4.160	0.951	0.331	-2.869	4.12E-03	3.24E-02
SPINK4	7.242	-0.789	0.275	2.868	4.12E-03	3.24E-02
CDO1	0.656	1.309	0.456	-2.868	4.14E-03	3.24E-02
HS3ST3B1	5.231	0.807	0.281	-2.867	4.14E-03	3.25E-02
ANAPC11	76.982	-0.228	0.079	2.867	4.15E-03	3.25E-02
ERAP2	61.106	0.527	0.184	-2.866	4.16E-03	3.26E-02
TOB2	42.698	0.316	0.110	-2.866	4.16E-03	3.26E-02
CTB-193M12.3	23.220	0.566	0.198	-2.864	4.18E-03	3.27E-02
KIFC3	10.429	-0.487	0.170	2.864	4.19E-03	3.27E-02
PPP1R26	15.458	-0.389	0.136	2.864	4.19E-03	3.27E-02
SLC2A4RG	69.836	-0.252	0.088	2.863	4.19E-03	3.28E-02
ATHL1	88.593	0.544	0.190	-2.863	4.20E-03	3.28E-02
AK4P1	0.709	1.327	0.464	-2.862	4.20E-03	3.28E-02
INHBE	1.657	0.938	0.328	-2.862	4.21E-03	3.29E-02
SF3A1	63.184	0.227	0.079	-2.861	4.22E-03	3.29E-02
RIOK3	34.629	0.304	0.106	-2.862	4.22E-03	3.29E-02
SFXN5	20.738	0.439	0.153	-2.860	4.23E-03	3.30E-02
FRY-AS1	1.578	0.982	0.343	-2.860	4.23E-03	3.30E-02
SRSF7	59.420	0.268	0.094	-2.860	4.24E-03	3.30E-02
GSG2	2.393	0.862	0.302	-2.860	4.24E-03	3.30E-02
SMIM7	106.661	0.285	0.100	-2.859	4.25E-03	3.31E-02
GSTT1	4.090	0.660	0.231	-2.858	4.26E-03	3.31E-02
FBXO45	14.496	-0.495	0.173	2.858	4.26E-03	3.31E-02
GCA1	6.967	0.649	0.227	-2.858	4.27E-03	3.31E-02
YPEL1	3.432	0.751	0.263	-2.858	4.26E-03	3.31E-02
ATP6V0E1	93.521	-0.215	0.075	2.858	4.27E-03	3.31E-02
SLC16A7	10.075	0.771	0.270	-2.858	4.27E-03	3.31E-02
CEP152	38.070	0.293	0.102	-2.856	4.29E-03	3.33E-02
RNU6-661P	0.822	-1.220	0.427	2.855	4.30E-03	3.33E-02
W12-89927D4.1	1.038	-1.159	0.406	2.855	4.30E-03	3.33E-02
AP1S1	42.177	-0.367	0.128	2.855	4.30E-03	3.33E-02
BTBD2	17.817	-0.358	0.125	2.855	4.30E-03	3.33E-02
SBNQ2	72.945	0.274	0.096	-2.855	4.30E-03	3.33E-02
GALNT2	31.754	0.385	0.135	-2.855	4.30E-03	3.33E-02
LINC00954	6.031	0.792	0.277	-2.855	4.31E-03	3.33E-02
HIST1H2BJ	2.664	0.781	0.274	-2.854	4.32E-03	3.34E-02
FAM196A	8.272	-0.510	0.179	2.853	4.33E-03	3.35E-02
ABTB2	9.672	-0.520	0.182	2.851	4.36E-03	3.36E-02
KCNAB2	42.090	0.686	0.240	-2.851	4.36E-03	3.36E-02
GCLM	17.976	0.453	0.159	-2.851	4.36E-03	3.36E-02
LG14	11.789	0.578	0.203	-2.850	4.37E-03	3.37E-02
TTC5	7.375	-0.524	0.184	2.849	4.38E-03	3.37E-02
IARS	27.665	0.339	0.119	-2.849	4.39E-03	3.38E-02
FAM63B	46.368	0.318	0.112	-2.848	4.40E-03	3.38E-02
EDRF1-AS1	11.104	0.719	0.252	-2.848	4.40E-03	3.39E-02
AUP1	238.008	0.185	0.065	-2.847	4.41E-03	3.39E-02
PCED1A	184.306	0.382	0.134	-2.847	4.42E-03	3.39E-02
NUB1	116.250	0.208	0.073	-2.845	4.44E-03	3.41E-02
C14orf79	2.026	0.814	0.286	-2.845	4.45E-03	3.42E-02
MKL1	32.067	0.372	0.131	-2.843	4.47E-03	3.43E-02
ZXDB	2.950	-0.673	0.237	2.843	4.47E-03	3.43E-02
SLC2A3	101.747	0.311	0.110	-2.842	4.48E-03	3.44E-02
SULT1B1	3.449	-0.675	0.238	2.840	4.51E-03	3.45E-02
ZC3HC1	8.093	-0.549	0.193	2.840	4.51E-03	3.45E-02
FAM84A	149.727	-0.413	0.145	2.840	4.50E-03	3.45E-02
ARHGEF38	16.349	0.488	0.172	-2.840	4.52E-03	3.46E-02
RGL1	30.392	0.529	0.186	-2.840	4.51E-03	3.46E-02
ITPKB	34.929	0.601	0.212	-2.839	4.52E-03	3.46E-02
MAFG	24.199	-0.409	0.144	2.838	4.54E-03	3.47E-02
MICALL1	15.748	0.493	0.174	-2.838	4.54E-03	3.47E-02
NDUFA4	40.965	-0.422	0.149	2.837	4.55E-03	3.47E-02
FAM105B	21.387	0.339	0.120	-2.838	4.55E-03	3.47E-02
ERI3	20.657	0.485	0.171	-2.837	4.55E-03	3.47E-02
AP5Z1	75.682	-0.258	0.091	2.837	4.55E-03	3.47E-02
AHCTF1P1	1.173	1.096	0.386	-2.836	4.56E-03	3.48E-02
ZFAS1	71.493	-0.320	0.113	2.835	4.58E-03	3.49E-02
ELMOD3	19.697	0.524	0.185	-2.835	4.58E-03	3.49E-02
CTSS	102.201	0.271	0.096	-2.835	4.59E-03	3.49E-02
CXCL1	13.349	-0.566	0.200	2.834	4.60E-03	3.50E-02
HIST1H4C	196.760	0.355	0.125	-2.834	4.60E-03	3.50E-02
LILRB5	2.450	0.890	0.314	-2.833	4.62E-03	3.51E-02

KLC1	49.995	-0.288	0.102	2.832	4.63E-03	3.51E-02
ZNF333	12.513	0.595	0.210	-2.832	4.62E-03	3.51E-02
HOXC9	2.295	1.024	0.362	-2.832	4.63E-03	3.51E-02
TMEM126B	12.577	-0.407	0.144	2.831	4.64E-03	3.52E-02
RGMB	9.431	-0.535	0.189	2.831	4.65E-03	3.53E-02
KCTD19	0.586	1.180	0.417	-2.830	4.65E-03	3.53E-02
MAP1LC3A	28.327	0.558	0.197	-2.830	4.65E-03	3.53E-02
TTC8	5.231	-0.883	0.312	2.828	4.68E-03	3.54E-02
NABP2	29.977	0.357	0.126	-2.828	4.68E-03	3.54E-02
FTH1P2	3.674	-0.648	0.229	2.828	4.68E-03	3.54E-02
EPB41	21.368	0.488	0.173	-2.827	4.70E-03	3.56E-02
CALM1	390.645	-0.137	0.048	2.826	4.72E-03	3.57E-02
CTB-47B11.3	9.134	0.729	0.258	-2.825	4.73E-03	3.58E-02
SNX2	21.692	0.346	0.123	-2.824	4.75E-03	3.59E-02
TMEM116	1.752	0.937	0.332	-2.823	4.76E-03	3.59E-02
HMOX1	119.459	0.317	0.112	-2.822	4.77E-03	3.60E-02
SHPRH	34.712	0.342	0.121	-2.822	4.77E-03	3.60E-02
PLCXD1	32.134	0.442	0.157	-2.823	4.76E-03	3.60E-02
SHE	6.781	0.653	0.231	-2.822	4.77E-03	3.60E-02
CTD-3203P2.2	2.815	-0.809	0.287	2.822	4.77E-03	3.60E-02
HARS2	9.459	-0.476	0.169	2.821	4.78E-03	3.60E-02
PEX2	7.995	-0.465	0.165	2.821	4.78E-03	3.60E-02
NOM1	50.632	0.336	0.119	-2.822	4.78E-03	3.60E-02
TAPT1	34.883	0.405	0.144	-2.821	4.79E-03	3.61E-02
PLEKHG6	8.442	0.571	0.202	-2.820	4.80E-03	3.61E-02
WIPF1	74.081	0.459	0.163	-2.820	4.80E-03	3.61E-02
APOL3	45.644	0.517	0.184	-2.819	4.82E-03	3.63E-02
UBAC2	36.190	-0.319	0.113	2.818	4.84E-03	3.64E-02
STAG1	28.025	0.398	0.141	-2.817	4.84E-03	3.64E-02
TGOLN2	259.721	0.191	0.068	-2.817	4.85E-03	3.64E-02
FKBP14	17.354	0.376	0.133	-2.816	4.86E-03	3.65E-02
TIMM44	15.795	0.466	0.165	-2.816	4.86E-03	3.65E-02
MXD1	27.680	0.382	0.136	-2.816	4.87E-03	3.65E-02
MSH6	23.950	0.455	0.162	-2.816	4.87E-03	3.65E-02
FAM155B	0.828	-1.201	0.427	2.814	4.89E-03	3.66E-02
TMEM218	4.277	-0.629	0.224	2.814	4.90E-03	3.67E-02
PRUNE2	10.226	-0.518	0.184	2.814	4.90E-03	3.67E-02
CDK19	41.330	0.320	0.114	-2.813	4.90E-03	3.67E-02
GABRB1	2.371	-0.913	0.324	2.813	4.91E-03	3.67E-02
BBS1	1.392	0.993	0.353	-2.812	4.93E-03	3.68E-02
PSME4	87.540	0.172	0.061	-2.811	4.93E-03	3.68E-02
DGCR14	6.639	0.640	0.228	-2.811	4.93E-03	3.68E-02
CHD4	278.350	0.196	0.070	-2.811	4.94E-03	3.69E-02
UBE2Z	113.951	0.218	0.077	-2.810	4.95E-03	3.69E-02
SRGAP2C	1.217	1.145	0.408	-2.810	4.95E-03	3.69E-02
BABAM1	40.671	-0.268	0.095	2.810	4.96E-03	3.70E-02
TUBE1	7.563	0.580	0.206	-2.809	4.97E-03	3.70E-02
FOXA2	32.526	-0.422	0.150	2.809	4.98E-03	3.70E-02
AHR	59.268	0.300	0.107	-2.809	4.97E-03	3.70E-02
RNF4	11.478	0.455	0.162	-2.807	5.00E-03	3.72E-02
NAP1L3	0.979	-1.073	0.382	2.807	5.00E-03	3.72E-02
C14orf166	76.147	-0.266	0.095	2.807	5.00E-03	3.72E-02
FAM173A	10.968	-0.458	0.163	2.806	5.01E-03	3.72E-02
CTC-228N24.3	39.106	0.312	0.111	-2.806	5.02E-03	3.73E-02
PRKACA	90.111	-0.244	0.087	2.805	5.03E-03	3.73E-02
SLTM	100.147	0.235	0.084	-2.804	5.05E-03	3.74E-02
BBIP1	62.447	0.512	0.183	-2.803	5.06E-03	3.75E-02
ARF1P2	2.088	0.987	0.352	-2.803	5.06E-03	3.75E-02
ZNF572	3.167	-0.690	0.246	2.802	5.08E-03	3.76E-02
AGBL3	2.155	0.854	0.305	-2.801	5.09E-03	3.77E-02
LRRC8B	23.234	0.395	0.141	-2.801	5.09E-03	3.77E-02
PIK3C3	28.728	0.319	0.114	-2.801	5.10E-03	3.77E-02
PTPLAD1	42.830	-0.333	0.119	2.800	5.11E-03	3.78E-02
EP400	42.353	0.310	0.111	-2.800	5.12E-03	3.78E-02
COX6A1	63.607	-0.264	0.094	2.799	5.13E-03	3.79E-02
GPR126	5.655	-0.780	0.279	2.798	5.13E-03	3.79E-02
TBCB	81.552	-0.283	0.101	2.798	5.14E-03	3.79E-02
YPEL3	121.635	0.483	0.173	-2.798	5.14E-03	3.79E-02
C9orf64	8.529	0.505	0.181	-2.798	5.14E-03	3.79E-02
ERVH48-1	0.792	1.336	0.477	-2.798	5.14E-03	3.79E-02
BRK1	107.605	-0.214	0.077	2.797	5.15E-03	3.80E-02
THOC7	11.370	-0.523	0.187	2.797	5.16E-03	3.80E-02
ZNF318	73.381	0.325	0.116	-2.797	5.17E-03	3.80E-02
APBA1	4.183	-0.711	0.254	2.796	5.17E-03	3.80E-02
CNOT4	41.902	0.346	0.124	-2.795	5.18E-03	3.81E-02
KSR1	6.380	0.658	0.235	-2.795	5.18E-03	3.81E-02
HK2	19.844	0.476	0.170	-2.795	5.20E-03	3.82E-02
TMEM164	21.214	-0.354	0.127	2.794	5.20E-03	3.82E-02
DCAF16	86.975	0.299	0.107	-2.794	5.20E-03	3.82E-02
ZNF717	1.639	0.903	0.323	-2.794	5.21E-03	3.82E-02
SLC25A14	11.236	-0.453	0.162	2.793	5.22E-03	3.83E-02
PTPRN2	5.059	0.714	0.256	-2.793	5.22E-03	3.83E-02
SLC7A2	3.560	0.724	0.259	-2.793	5.23E-03	3.83E-02
NFYC	13.200	-0.410	0.147	2.792	5.25E-03	3.84E-02
MBLAC1	0.659	-1.245	0.446	2.791	5.26E-03	3.85E-02
TIMP1	243.188	0.258	0.092	-2.791	5.26E-03	3.85E-02

AIG1	14.031	-0.372	0.133	2.790	5 28E-03	3.86E-02
SMARCE1P6	1.671	0.874	0.313	-2.790	5 28E-03	3.86E-02
NEDD1	26.317	-0.315	0.113	2.788	5 30E-03	3.87E-02
HDAC3	61.663	-0.262	0.094	2.788	5 30E-03	3.87E-02
IGLL5	14.653	0.689	0.247	-2.787	5 31E-03	3.88E-02
KLHDC10	25.048	-0.351	0.126	2.786	5 33E-03	3.89E-02
PRDM6	4.480	0.554	0.199	-2.786	5 33E-03	3.89E-02
TOP3B	1.106	1.138	0.408	-2.787	5 33E-03	3.89E-02
MESDC2	97.793	0.249	0.089	-2.786	5 33E-03	3.89E-02
FAM227B	2.097	-0.815	0.293	2.784	5 36E-03	3.91E-02
CTC-458I2.2	2.351	0.810	0.291	-2.783	5 38E-03	3.92E-02
TGFB111	67.702	0.422	0.152	-2.782	5.40E-03	3.93E-02
CAMK2D	30.082	0.355	0.128	-2.782	5.40E-03	3.93E-02
SNHG3	36.700	-0.346	0.124	2.781	5.41E-03	3.94E-02
TPM3	160.884	0.249	0.089	-2.781	5.41E-03	3.94E-02
DRG1	76.749	0.397	0.143	-2.781	5.42E-03	3.94E-02
BASP1	190.317	0.483	0.174	-2.781	5.41E-03	3.94E-02
WSCD1	8.039	0.677	0.243	-2.781	5.42E-03	3.94E-02
TSPAN33	17.842	0.782	0.281	-2.781	5.42E-03	3.94E-02
ZDHHC15	7.138	0.932	0.335	-2.780	5.44E-03	3.95E-02
S1PR1	3.864	0.681	0.245	-2.779	5.45E-03	3.95E-02
AKAP12	80.499	0.358	0.129	-2.778	5.46E-03	3.96E-02
ATP6V0E2	12.117	0.564	0.203	-2.778	5.47E-03	3.97E-02
TMEM8B	29.271	0.453	0.163	-2.777	5.48E-03	3.97E-02
SPI1	8.966	0.626	0.225	-2.777	5.49E-03	3.97E-02
APC	20.904	0.403	0.145	-2.776	5 50E-03	3.98E-02
APOLD1	50.125	0.445	0.160	-2.775	5 52E-03	3.99E-02
CABP4	6.088	0.778	0.280	-2.775	5 52E-03	3.99E-02
NSUN2	29.769	-0.434	0.156	2.774	5 53E-03	3.99E-02
KLC2	16.753	0.573	0.206	-2.774	5 53E-03	3.99E-02
S100A4	96.627	-0.311	0.112	2.774	5 54E-03	4.00E-02
REL	43.245	0.426	0.154	-2.772	5 56E-03	4.02E-02
UBAP2L	47.442	0.231	0.083	-2.772	5 57E-03	4.02E-02
MAPK8	11.133	-0.431	0.156	2.771	5 59E-03	4.03E-02
SSR4	266.312	0.351	0.127	-2.771	5 59E-03	4.03E-02
FAM73B	24.103	0.516	0.186	-2.771	5.60E-03	4.03E-02
LDHAL6A	2.812	0.803	0.290	-2.770	5.61E-03	4.04E-02
ARHGAP4	373.008	0.419	0.151	-2.769	5.62E-03	4.04E-02
SLC25A6	589.015	-0.177	0.064	2.769	5.62E-03	4.04E-02
ARID3A	72.702	0.336	0.121	-2.768	5.63E-03	4.05E-02
CTD-2227E11.1	5.389	-0.612	0.221	2.768	5.64E-03	4.05E-02
ATP1B1	82.799	-0.375	0.135	2.768	5.64E-03	4.05E-02
DCK	2.820	-0.722	0.261	2.767	5.66E-03	4.06E-02
PSMB10	0.872	1.066	0.385	-2.765	5.69E-03	4.08E-02
RN7SL842P	0.479	1.899	0.687	-2.765	5.69E-03	4.08E-02
ANTXR1	18.994	-0.502	0.182	2.765	5.70E-03	4.09E-02
TMBIM6	323.741	-0.138	0.050	2.765	5.70E-03	4.09E-02
TOM1	33.507	0.440	0.159	-2.764	5.70E-03	4.09E-02
TTL	40.147	0.304	0.110	-2.764	5.71E-03	4.09E-02
ZMIZ2	109.959	0.307	0.111	-2.764	5.71E-03	4.09E-02
CHRNA3	0.749	-1.190	0.431	2.763	5.72E-03	4.09E-02
C6orf48	79.672	-0.378	0.137	2.763	5.72E-03	4.09E-02
MCOLN3	4.182	-0.676	0.245	2.762	5.75E-03	4.11E-02
THEM6	13.684	-0.451	0.163	2.761	5.75E-03	4.11E-02
TMEM255B	2.710	-0.817	0.296	2.761	5.76E-03	4.12E-02
RLIMP1	1.150	-0.978	0.354	2.761	5.77E-03	4.12E-02
CAMK2G	16.610	0.360	0.130	-2.760	5.77E-03	4.12E-02
MEIS3	4.013	-0.781	0.283	2.760	5.78E-03	4.12E-02
YTHDC2	15.404	0.404	0.147	-2.759	5.79E-03	4.13E-02
ZNF189	7.983	0.516	0.187	-2.759	5 80E-03	4.13E-02
TBC1D3F	1.728	-0.879	0.319	2.758	5 81E-03	4.14E-02
COMMD6	46.710	-0.418	0.151	2.759	5 81E-03	4.14E-02
RN7SL230P	2.796	0.917	0.333	-2.758	5 81E-03	4.14E-02
MKKS	11.783	-0.477	0.173	2.758	5 82E-03	4.14E-02
ADSL	30.971	0.347	0.126	-2.758	5 82E-03	4.14E-02
KIAA1644	0.591	-1.280	0.464	2.757	5 83E-03	4.15E-02
EDIL3	3.694	-0.758	0.275	2.756	5 85E-03	4.15E-02
AKR1C3	13.453	-0.559	0.203	2.756	5 85E-03	4.15E-02
WTAP	40.153	-0.241	0.087	2.756	5 84E-03	4.15E-02
KCTD12	26.558	0.428	0.155	-2.756	5 85E-03	4.15E-02
MFAP1P1	1.602	1.057	0.384	-2.756	5 85E-03	4.15E-02
KPNA4	55.480	0.251	0.091	-2.754	5 89E-03	4.18E-02
HNRNPA1P40	0.770	1.217	0.442	-2.753	5 90E-03	4.18E-02
UHRF1BP1L	10.435	0.500	0.182	-2.753	5 91E-03	4.19E-02
PEA15	143.363	0.289	0.105	-2.751	5 94E-03	4.21E-02
POLR2E	97.205	-0.199	0.072	2.750	5 95E-03	4.21E-02
HAND1	0.961	-1.176	0.428	2.750	5 96E-03	4.22E-02
XPO6	35.262	-0.331	0.120	2.750	5 96E-03	4.22E-02
BDKRB2	2.891	-0.760	0.276	2.749	5 97E-03	4.22E-02
HIST2H2AB	3.858	0.780	0.284	-2.749	5 98E-03	4.23E-02
LG12	2.774	-0.832	0.303	2.748	5 99E-03	4.24E-02
DYRK4	4.553	-0.598	0.218	2.747	6 01E-03	4.24E-02
PLAGL2	17.779	-0.455	0.165	2.747	6 01E-03	4.25E-02
ARHGEF2	46.759	0.394	0.144	-2.747	6 01E-03	4.25E-02
MBOAT7	49.260	0.332	0.121	-2.747	6 02E-03	4.25E-02
AMZ2	18.528	0.418	0.152	-2.746	6 03E-03	4.25E-02

CTD-3131K8.3	5.781	0.726	0.264	-2.746	6 03E-03	4.25E-02
RBPMS	26.958	-0.363	0.132	2.745	6 05E-03	4.26E-02
ADAMTS14	3.107	-0.727	0.265	2.744	6 06E-03	4.27E-02
NIT2	15.463	-0.401	0.146	2.744	6 07E-03	4.28E-02
CEP128	7.984	0.495	0.181	-2.744	6 07E-03	4.28E-02
CAPN5	16.157	-0.448	0.163	2.744	6 08E-03	4.28E-02
ACTC1	0.792	-1.054	0.384	2.743	6 08E-03	4.28E-02
CELF2-AS2	1.707	-0.877	0.320	2.743	6 09E-03	4.28E-02
FAM161B	2.242	0.841	0.307	-2.743	6.10E-03	4.29E-02
HTT	49.034	0.277	0.101	-2.742	6.11E-03	4.29E-02
SP140L	61.492	0.436	0.159	-2.742	6.10E-03	4.29E-02
YEATS2-AS1	0.762	1.185	0.432	-2.742	6.11E-03	4.29E-02
HOMER3	28.538	-0.331	0.121	2.741	6.12E-03	4.30E-02
MYCBPAP	8.871	0.961	0.351	-2.741	6.12E-03	4.30E-02
FADS1	31.860	0.428	0.156	-2.740	6.14E-03	4.30E-02
ATRAID	17.999	-0.376	0.137	2.740	6.15E-03	4.31E-02
NTMT1	54.751	0.378	0.138	-2.738	6.18E-03	4.33E-02
PUS10	6.230	0.523	0.191	-2.738	6.19E-03	4.33E-02
CERS4	6.575	0.702	0.257	-2.737	6.19E-03	4.33E-02
SGCD	8.169	-0.602	0.220	2.737	6 20E-03	4.33E-02
HABP4	3.423	0.658	0.241	-2.737	6 20E-03	4.33E-02
DMBX1	1.013	1.147	0.419	-2.737	6 20E-03	4.33E-02
CYP4A22-AS1	0.868	1.054	0.385	-2.736	6 21E-03	4.34E-02
RFPL1S	0.995	0.994	0.363	-2.735	6 24E-03	4.36E-02
GLRX3	20.613	-0.352	0.129	2.733	6 28E-03	4.38E-02
TEF	9.407	-0.464	0.170	2.733	6 29E-03	4.38E-02
IGHG3	336.323	0.533	0.195	-2.731	6 30E-03	4.39E-02
ARHGEF38-IT1	0.795	1.131	0.414	-2.732	6 30E-03	4.39E-02
GTF3A	48.001	-0.342	0.125	2.731	6 32E-03	4.40E-02
KRT79	0.418	1.455	0.533	-2.731	6 32E-03	4.40E-02
CSNK1G2-AS1	4.268	0.803	0.294	-2.730	6 33E-03	4.40E-02
NIN	54.575	0.368	0.135	-2.730	6 34E-03	4.41E-02
FAM193A	48.718	0.377	0.138	-2.729	6 34E-03	4.41E-02
ZNF12	18.337	-0.340	0.125	2.728	6 37E-03	4.42E-02
PRKAG1	116.459	-0.167	0.061	2.728	6 36E-03	4.42E-02
TMEM41B	10.804	-0.482	0.177	2.728	6 37E-03	4.42E-02
M6PR	7.703	0.560	0.205	-2.728	6 38E-03	4.42E-02
SLC25A45	18.868	0.649	0.238	-2.728	6 38E-03	4.42E-02
UGGT1	56.590	0.241	0.088	-2.727	6 38E-03	4.42E-02
ETS1	113.264	0.371	0.136	-2.727	6 38E-03	4.42E-02
ETFB	17.755	0.438	0.161	-2.727	6 38E-03	4.42E-02
ABHD6	3.310	0.691	0.254	-2.727	6 39E-03	4.43E-02
CLCC1	4.920	0.662	0.243	-2.727	6.40E-03	4.43E-02
CUBN	5.219	0.797	0.292	-2.726	6.42E-03	4.44E-02
ZNF558	3.491	-0.663	0.243	2.724	6.45E-03	4.46E-02
DNAJC3	30.935	-0.287	0.105	2.724	6.46E-03	4.46E-02
ZNF746	4.245	-0.647	0.238	2.723	6.48E-03	4.47E-02
KLF10	5.400	-0.538	0.197	2.723	6.47E-03	4.47E-02
CLDN2	34.626	-0.526	0.193	2.723	6.48E-03	4.47E-02
HIP1	26.511	0.363	0.133	-2.723	6.48E-03	4.47E-02
MGEA5	192.038	0.233	0.086	-2.722	6.48E-03	4.47E-02
CYorf17	1.444	-0.931	0.342	2.722	6.49E-03	4.48E-02
DOCK8	62.074	0.658	0.242	-2.721	6 51E-03	4.49E-02
RASL10A	1.705	-0.885	0.325	2.720	6 52E-03	4.49E-02
GFER	17.008	0.478	0.176	-2.719	6 54E-03	4.50E-02
PRRC2C	510.468	0.166	0.061	-2.719	6 56E-03	4.51E-02
CCDC88C	109.911	0.278	0.102	-2.719	6 56E-03	4.51E-02
PEMT	20.496	0.443	0.163	-2.719	6 55E-03	4.51E-02
TERF1P5	1.039	0.952	0.350	-2.718	6 56E-03	4.51E-02
CLK2	31.659	0.322	0.119	-2.717	6 58E-03	4.53E-02
P2RX4	23.894	0.455	0.168	-2.714	6.64E-03	4.56E-02
TRAPPC10	40.134	0.259	0.095	-2.713	6.66E-03	4.57E-02
IL18	10.654	0.447	0.165	-2.713	6.66E-03	4.57E-02
SNRPB	123.355	-0.242	0.089	2.713	6.67E-03	4.57E-02
RHD	15.087	0.680	0.251	-2.713	6.67E-03	4.58E-02
PBX2	6.836	0.666	0.245	-2.712	6.68E-03	4.58E-02
GS1-124K5.4	0.647	-1.050	0.387	2.712	6.69E-03	4.59E-02
COLGALT1	86.975	-0.245	0.090	2.711	6.70E-03	4.59E-02
ZCRB1	73.288	0.229	0.084	-2.711	6.72E-03	4.60E-02
KDM8	5.112	0.670	0.247	-2.709	6.74E-03	4.61E-02
CDK7	5.860	0.532	0.197	-2.709	6.76E-03	4.62E-02
CYP2B7P	1.412	1.123	0.415	-2.709	6.76E-03	4.62E-02
LA16c-						
313D11.12	0.664	1.189	0.439	-2.708	6.77E-03	4.62E-02
LGMN	18.282	0.381	0.141	-2.707	6.78E-03	4.63E-02
RN7SL832P	8.436	0.735	0.272	-2.705	6 84E-03	4.67E-02
ZCCHC18	1.030	-0.975	0.361	2.704	6 85E-03	4.67E-02
TLR5	2.969	0.676	0.250	-2.704	6 85E-03	4.67E-02
CSNK2B	34.728	-0.352	0.130	2.702	6 89E-03	4.70E-02
TAPBP	675.715	0.339	0.126	-2.702	6 89E-03	4.70E-02
CCDC144B	13.567	0.617	0.228	-2.702	6 89E-03	4.70E-02
ANKRD28	26.290	0.270	0.100	-2.701	6 92E-03	4.71E-02
DENN2D3	34.311	0.379	0.140	-2.701	6 92E-03	4.71E-02
ACYP2	12.295	0.629	0.233	-2.698	6 97E-03	4.74E-02
PILRB	53.194	0.519	0.192	-2.698	6 98E-03	4.75E-02
OR10A4	3.172	0.783	0.290	-2.697	7 00E-03	4.76E-02

UBQLN1	56.191	-0.295	0.109	2.696	7.02E-03	4.77E-02
EIF1AX	113.441	-0.319	0.118	2.696	7.03E-03	4.78E-02
ANKRD18CP	0.654	1.082	0.402	-2.695	7.04E-03	4.78E-02
SCRIB	33.187	0.337	0.125	-2.694	7.05E-03	4.79E-02
TINCR	1.089	1.021	0.379	-2.694	7.05E-03	4.79E-02
PET100	13.215	-0.611	0.227	2.692	7.09E-03	4.81E-02
RFTN2	6.515	0.693	0.257	-2.692	7.10E-03	4.81E-02
RN7SL179P	4.880	0.707	0.263	-2.692	7.11E-03	4.82E-02
DDX39B	23.385	0.482	0.179	-2.691	7.12E-03	4.82E-02
C5orf28	7.291	-0.522	0.194	2.689	7.16E-03	4.84E-02
SQRDL	22.852	-0.453	0.168	2.689	7.16E-03	4.84E-02
RASL11B	1.458	0.974	0.362	-2.689	7.16E-03	4.84E-02
CHAD	0.728	1.288	0.479	-2.689	7.16E-03	4.84E-02
NDC80	10.207	0.573	0.213	-2.689	7.17E-03	4.85E-02
BANF1	49.305	-0.216	0.081	2.689	7.18E-03	4.85E-02
HS3ST1	8.747	0.518	0.193	-2.688	7.19E-03	4.86E-02
KB-1205A7.2	1.229	1.097	0.408	-2.687	7.20E-03	4.86E-02
TALDO1	106.687	-0.219	0.082	2.687	7.22E-03	4.87E-02
ERO1L	64.078	0.228	0.085	-2.687	7.22E-03	4.87E-02
CD99	45.830	-0.256	0.095	2.686	7.23E-03	4.88E-02
ZSCAN32	2.827	0.652	0.243	-2.686	7.23E-03	4.88E-02
TMSB4XP4	1.933	-0.760	0.283	2.684	7.26E-03	4.90E-02
ADAP2	36.048	0.516	0.192	-2.684	7.26E-03	4.90E-02
MEGF6	21.624	0.606	0.226	-2.684	7.27E-03	4.90E-02
TCP1	22.382	0.423	0.158	-2.684	7.27E-03	4.90E-02
PHYKPL	40.549	0.332	0.124	-2.683	7.29E-03	4.91E-02
FGD2	41.516	0.660	0.246	-2.683	7.30E-03	4.91E-02
POM121C	39.866	-0.437	0.163	2.682	7.31E-03	4.92E-02
KIAA1109	64.000	0.260	0.097	-2.682	7.32E-03	4.93E-02
NR1H2	10.192	0.492	0.183	-2.681	7.35E-03	4.94E-02
GSDMA	1.069	-1.024	0.382	2.681	7.35E-03	4.94E-02
CTD-2196E14.4	18.206	0.589	0.220	-2.680	7.36E-03	4.94E-02
RBM5	77.426	0.272	0.102	-2.677	7.44E-03	4.99E-02
LRTOMT	19.653	0.579	0.216	-2.677	7.44E-03	4.99E-02
PROSC	37.189	0.302	0.113	-2.676	7.44E-03	4.99E-02

Supplementary Table 10: Differential gene expression between stage 3 primary CRCs and lymph node metastases determined using DESeq2. Positive log fold change indicates higher expression in LN metastases.

Geneset	LNM	Primary	Difference	p value
Neg reg adaptive immunity	9.825	9.450	0.376	5.20E-07
Neg reg cytokine production	9.655	9.215	0.439	1.54E-04
Neg reg cytokine secretion	9.005	8.822	0.183	4.14E-01
Neg reg immune effector	9.740	9.443	0.296	2.91E-08
Neg reg immune response	9.764	9.476	0.288	2.33E-08
Neg reg innate immunity	9.932	9.694	0.237	7.92E-07
Neg reg molecular mediator production	9.647	9.278	0.369	4.33E-06
Pos reg adaptive immunity	9.631	9.261	0.370	2.64E-09
Pos reg cytokine production	9.564	9.278	0.285	1.75E-04
Pos reg cytokine secretion	9.871	9.827	0.044	9.78E-01
Pos reg immune effector	9.595	9.270	0.325	3.06E-09
Pos reg immune response	9.685	9.418	0.267	3.07E-11
Pos reg molecular mediator production	9.623	9.301	0.322	5.85E-07

Supplementary Table 11: Negative and positive regulation of immune response expression changes between lymph node metastases and stage 3 CRC primary tumours. Wilcoxon-rank sum tests were performed.

Geneset	Aggressive	Non-aggressive	Difference	Adjusted p value
TIGER 1	9.854	9.664	0.190	5.08E-03
TIGER 2	9.931	9.743	0.188	0.014
TIGER 3	9.658	9.342	0.316	3.12E-04
TIGER 4	9.352	8.982	0.370	NS
TIGER 5	9.758	9.434	0.324	5.90E-04
TIGER 6	9.493	8.974	0.519	1.28E-03
TIGER 7	9.851	9.630	0.221	3.64E-03
Apoptosis	9.854	9.664	0.190	2.18E-02
MHC class I antigen presentation	9.887	9.678	0.209	1.61E-02
MHC class II antigen presentation	9.976	9.808	0.167	NS
Activated CD4	9.302	8.863	0.438	6.61E-03
Activated CD8	9.853	9.545	0.308	NS
T cells	9.819	9.619	0.200	NS
Endothelial cell activation	9.352	8.982	0.370	NS
Central memory CD4	10.009	9.730	0.279	NS
Effector memory CD4	9.926	9.711	0.214	NS
Th1	10.035	9.790	0.245	NS
Th17	9.314	9.093	0.220	NS
Treg	9.574	9.357	0.217	NS
Central memory CD8	9.588	9.344	0.244	NS
Cytotoxic cells	10.101	9.754	0.347	NS
Effector memory CD8	9.826	9.293	0.533	4.52E-06
Gamma Delta T Cells (TGD)	9.590	9.096	0.494	4.15E-03
Natural Killer T Cells (NKT)	9.622	9.175	0.447	NS
Immune checkpoint	9.493	8.974	0.519	5.48E-03
Effector CD8 v exhausted CD8	9.904	9.729	0.175	NS
NK cytotoxicity	9.798	9.531	0.267	1.20E-02

Supplementary Table 12: Cancer Immunity Cycle expression changes in lymph node metastases based on lymph node aggressiveness (proliferative capacity). Wilcoxon-rank sum tests were performed, and *p* values corrected by applying a Bonferroni coefficient.

xCell Gene Signature	PC1	PC2		xCell Gene Signature	PC1	PC2
cDC	0.558	0.234		CD8+ Tem	0.029	0.007
iDC	0.383	0.385		NK cells	0.027	0.014
aDC	0.367	0.046		Mast cells	0.027	-0.016
HSC	0.285	0.104		mv Endothelial cells	0.022	0.018
MSC	0.174	-0.414		Memory B cells	0.021	-0.058
CD8+ Tcm	0.130	0.038		Myocytes	0.020	-0.020
CD8+ T cells	0.121	-0.030		Astrocytes	0.020	0.013
DC	0.103	0.086		Mesangial cells	0.019	0.025
Fibroblasts	0.102	0.051		CD4+ naive T cells	0.018	-0.035
Th2 cells	0.101	0.169		Melanocytes	0.016	-0.012
Basophils	0.100	-0.072		Preadipocytes	0.015	0.020
Macrophages	0.087	0.081		Plasma cells	0.015	-0.006
B cells	0.083	-0.241		CD4+ T cells	0.015	-0.025
pDC	0.082	0.009		Skeletal muscle	0.012	-0.019
Tregs	0.076	-0.086		naive B cells	0.011	-0.035
ly Endothelial cells	0.073	0.003		pro B cells	0.007	-0.093
Macrophages M2	0.068	-0.002		Neutrophils	0.007	0.013
Endothelial cells	0.058	0.011		CD4+ Tcm	0.007	-0.060
CMP	0.055	-0.006		Hepatocytes	0.006	-0.009
GMP	0.054	0.024		Adipocytes	0.005	-0.007
Chondrocytes	0.054	0.012		Erythrocytes	0.000	0.000
Platelets	0.051	-0.042		Tgd cells	0.000	-0.001
CD4+ memory T cells	0.048	0.021		Sebocytes	-0.003	0.007
Class switched memory B cells	0.047	-0.081		MPP	-0.003	-0.003
Monocytes	0.044	-0.004		Keratinocytes	-0.007	0.010
Macrophages M1	0.044	0.042		CD8+ naive T cells	-0.013	0.071
CD4+ Tem	0.042	0.000		MEP	-0.027	0.048
Megakaryocytes	0.037	0.015		Epithelial cells	-0.065	0.043
Eosinophils	0.037	-0.006		Osteoblast	-0.095	0.102
Pericytes	0.034	-0.018		Th1 cells	-0.131	0.150
NKT	0.032	-0.019		Smooth muscle	-0.233	0.501
Neurons	0.032	-0.023		CLP	-0.251	0.409

Supplementary Table 13: Determinants of xCell principal component analysis (PCA) for PCA axes 1 and

2. Dendritic cell signatures were found to be major drivers of variation in PC1.

Colon primary tumour v Lung metastasis

Geneset	Lung metastasis	Colon primary	Difference	p value
MHC CLASS II BIOSYNTHETIC PROCESS	8.773	8.732	0.041	0.6498
MHC CLASS II PROTEIN BINDING	9.016	8.125	0.892	0.0088
MHC CLASS II PROTEIN COMPLEX	10.268	9.618	0.651	0.0002
MHC CLASS II PROTEIN COMPLEX BINDING	10.572	10.223	0.348	0.0004
MHC CLASS II RECEPTOR ACTIVITY	10.130	9.247	0.883	0.0020
ANTIGEN PROCESSING AND PRESENTATION OF PEPTIDE ANTIGEN VIA MHC CLASS II	10.359	10.177	0.182	0.0008

Liver metastasis v Lung metastasis

Geneset	Lung metastasis	Liver metastasis	Difference	p value
MHC CLASS II BIOSYNTHETIC PROCESS	8.773	8.741	0.032	0.5614
MHC CLASS II PROTEIN BINDING	9.016	8.326	0.691	0.0115
MHC CLASS II PROTEIN COMPLEX	10.268	9.903	0.365	0.0061
MHC CLASS II PROTEIN COMPLEX BINDING	10.572	10.392	0.179	0.0103
MHC CLASS II RECEPTOR ACTIVITY	10.130	9.641	0.489	0.0161
ANTIGEN PROCESSING AND PRESENTATION OF PEPTIDE ANTIGEN VIA MHC CLASS II	10.359	10.361	-0.002	0.0090

Supplementary Table 14: Differences in class II MHC gene signature expression in lung metastases compared to CRC primary tumours and liver metastases. Wilcoxon-rank sum tests were performed.

	Haematological malignancies (n=227)	Solid tumour malignancies (n=817)	Univariable		Multivariable adjusted	
			Odds Ratio (95% CI)	p value	Odds Ratio (95% CI)	p value
Demographics						
Male	148 (65.2%)	447 (54.7%)	1.53 (1.13-2.09)	0.006	1.58 (1.16-2.16)	0.004
Female	79 (34.8%)	366 (44.8%)				
Other	0 (0.0%)	4 (0.5%)				
Median age/years (std)	69.0 (14.16)	70.0 (13.09)		0.034		
Co-morbidities						
Cardiovascular disease	21 (9.3%)	124 (15.2%)	0.56 (0.34-0.91)	0.023	0.62 (0.36-1.01)	0.065
COPD	7 (3.1%)	73 (8.9%)	0.32 (0.13-0.66)	0.005	0.35 (0.14-0.72)	0.009
Diabetes	33 (14.5%)	145 (17.7%)	0.79 (0.51-1.18)	0.262	0.78 (0.50-1.17)	0.243
Hypertension	60 (26.4%)	283 (34.6%)	0.66 (0.47-0.92)	0.017	0.68 (0.47-0.97)	0.033
None	52 (22.9%)	153 (18.7%)	1.32 (0.91-1.90)	0.138	1.31 (0.87-1.96)	0.189
No data	39 (17.2%)	136 (16.6%)				
Smoking status						
Current smoker	7 (3.1%)	38 (4.7%)	0.99 (0.39-2.17)	0.984	0.77 (0.30-1.74)	0.558
Ex-smoker	32 (14.1%)	234 (28.6%)	0.59 (0.37-0.94)	0.028	0.63 (0.38-1.03)	0.067
Never smoker	52 (22.9%)	218 (26.7%)	1.66 (1.06-2.63)	0.027	1.67 (1.04-2.70)	0.035
No data	136 (59.9%)	327 (40.0%)				
Patient outcome						
Death (all cause)	82 (36.1%)	237 (29.0%)	1.61 (1.15-2.24)	0.005	1.74 (1.21-2.48)	0.002
Death (COVID-19)	80 (35.2%)	215 (26.3%)	1.77 (1.27-2.48)	0.001	1.93 (1.35-2.77)	<0.001
Death (Cancer)	1 (0.4%)	18 (2.2%)	0.21 (0.01-1.02)	0.129	0.22 (0.01-1.06)	0.138
Death (other)	1 (0.4%)	4 (0.5%)	0.96 (0.05-6.54)	0.972	1.12 (0.06-7.79)	0.923
Hospitalisation	5 (2.2%)	36 (4.4%)	0.52 (0.18-1.23)	0.178	0.53 (0.18-1.26)	0.192
Cancer treatment within 4 weeks						
Chemotherapy	108 (47.6%)	241 (29.5%)	2.17 (1.60-2.93)	<0.0001	2.15 (1.57-2.95)	<0.0001
Immunotherapy	0 (0.0%)	39 (4.8%)	0.00 (0.00-2.90E+07)	0.9815	0.00 (0.00-3.17E+07)	0.9813
Radio herapy	2 (0.9%)	84 (10.3%)	0.08 (0.01-0.25)	0.0004	0.07 (0.01-0.24)	0.0003
Surgery	0 (0.0%)	36 (4.4%)	0.00 (0.00-8.99E+07)	0.9816	0.00 (0.00-1.39E+08)	0.9816
Targeted therapy	26 (11.5%)	65 (8.0%)	1.49 (0.91-2.39)	0.1018	1.45 (0.87-2.33)	0.1397

	Haematological malignancies (n=227)	Solid tumour malignancies (n=817)	Univariable		Multivariable adjusted	
			Odds Ratio (95% CI)	p value	Odds Ratio (95% CI)	p value
COVID-19 Symptoms						
Chills	9 (4.0%)	23 (2.8%)	1.49 (0.64-3.16)	0.324	1.45 (0.62-3.11)	0.357
Coryzal symptoms	13 (5.7%)	47 (5.8%)	1.04 (0.53-1.90)	0.911	1.03 (0.52-1.90)	0.931
Cough	93 (41.0%)	381 (46.6%)	0.83 (0.60-1.14)	0.255	0.82 (0.59-1.13)	0.217
Diarrhoea	26 (11.5%)	63 (7.7%)	1.63 (0.99-2.62)	0.050	1.67 (1.01-2.70)	0.041
Fatigue	46 (20.3%)	150 (18.4%)	1.19 (0.81-1.73)	0.359	1.22 (0.83-1.77)	0.307
Fever	133 (58.6%)	450 (55.1%)	1.34 (0.96-1.90)	0.091	1.26 (0.89-1.79)	0.191
Headache	15 (6.6%)	28 (3.4%)	2.09 (1.07-3.94)	0.026	2.11 (1.07-4.00)	0.026
Myalgia	18 (7.9%)	60 (7.3%)	1.13 (0.64-1.93)	0.654	1.15 (0.64-1.97)	0.623
Nausea and/or vomiting	8 (3.5%)	43 (5.3%)	0.68 (0.29-1.40)	0.332	0.70 (0.30-1.45)	0.372
Shortness of breath	84 (37.0%)	324 (39.7%)	0.95 (0.69-1.30)	0.734	0.92 (0.66-1.27)	0.610
Sore throat	9 (4.0%)	32 (3.9%)	1.05 (0.47-2.16)	0.891	0.98 (0.43-2.04)	0.964
Asymptomatic	5 (2.2%)	39 (4.8%)	0.47 (0.16-1.09)	0.113	0.50 (0.17-1.18)	0.152
No data	35 (15.4%)	99 (12.1%)				
COVID-19 Severity						
Severe or critical	119 (52.4%)	339 (41.5%)	1.53 (1.13-2.06)	0.006	1.57 (1.15-2.15)	0.004
Mild	103 (45.4%)	448 (54.8%)				
No data	5 (2.2%)	30 (3.7%)				
COVID-19 treatment						
Antibiotics	145 (63.9%)	495 (60.6%)	1.35 (0.93-2.00)	0.120	1.35 (0.92-2.00)	0.129
Fluids	86 (37.9%)	247 (30.2%)	1.52 (1.10-2.11)	0.012	1.54 (1.10-2.14)	0.011
High Flow Oxygen	29 (12.8%)	61 (7.5%)	1.89 (1.16-3.01)	0.009	1.82 (1.11-2.94)	0.015
Intensive care admission, ventilation	19 (8.4%)	25 (3.1%)	3.00 (1.60-5.57)	0.001	2.73 (1.43-5.11)	0.002
Intensive care admission, no ven ila ion	7 (3.1%)	12 (1.5%)	2.19 (0.81-5.54)	0.104	2.16 (0.78-5.54)	0.118
Non-invasive Ventilation	19 (8.4%)	35 (4.3%)	2.11 (1.16-3.75)	0.012	2.10 (1.14-3.76)	0.014
Oxygen	99 (43.6%)	310 (37.9%)	1.38 (1.00-1.90)	0.054	1.41 (1.01-1.96)	0.044
None	26 (11.5%)	134 (16.4%)	0.67 (0.42-1.04)	0.083	0.65 (0.41-1.02)	0.071
No data	40 (17.6%)	128 (15.7%)				

Supplementary Table 15: Univariable and multivariable adjusted analyses of differences in patient demographics/symptoms/cancer treatments and clinical course of haematological compared to solid tumour malignancies. Analyses were performed with presence compared to absence (reference for each category) in haematological malignancies vs. non-haematological malignancies. Multivariable regression models were adjusted for patient age and sex.

SACT Regimen	SACT Type	SACT Regimen	SACT Class
ABEMACICLIB	Cytotoxic	ABEMACICLIB	A
ABVD	Cytotoxic	ABEMACICLIB + FULVESTRANT	A
AC	Cytotoxic	Abemaciclib + Letrozole	A
ACCEPT	Cytotoxic	ABIRATERONE	A
ACE (GERM CELL)	Cytotoxic	AC	A
ACELARATE	Cytotoxic	ACALABRUTINIB	A
ACTICCA-1	Cytotoxic	AFATINIB	A
ADE	Cytotoxic	ALECTINIB	A
ALEMTUZUMAB	Cytotoxic	ANAGRELIDE	A
ALL R T-CELL CONSOLIDATION	Cytotoxic	ARSENIC TRIOXIDE	A
ALL/BL T CELL RELAPSE NECTAR CYCLOPHOSPHAMIDE, ETOPOSIDE, NELARABINE AND IT	Cytotoxic	ARSENIC TRIOXIDE + TRETINOIN	A
AML RELAPSED FLA - IDARUBICIN	Cytotoxic	ASCIMINIB	A
AML19	Cytotoxic	AVAPRITINIB	A
AMSACRINE	Cytotoxic	AVELUMAB	A
APML	Cytotoxic	AXITINIB	A
ARSENICTRIOXIDE	Cytotoxic	BCG	A
ATEZOLIZUMAB	Cytotoxic	BEXAROTENE	A
BEACOPP	Cytotoxic	BINIMETINIB + ENCORAFENIB	A
BEAM	Cytotoxic	BLEOMYCIN	A
BENDAMUSTINE	Cytotoxic	BORTEZOMIB + PANOBINOSTAT	A
BEP	Cytotoxic	BORTEZOMIB + PANOBINOSTAT + THALIDOMIDE	A
BEVACIZUMAB	Cytotoxic	BORTEZOMIB + THALIDOMIDE	A
BEXAROTENE	Cytotoxic	BOSUTINIB	A
BLEOMYCIN	Cytotoxic	BRIGATINIB	A
BOP	Cytotoxic	CABOZANTINIB	A
BUSULFAN	Cytotoxic	CAPECITABINE	A
BUSULPHAN	Cytotoxic	CAPECITABINE + IRINOTECAN	A
CABAZITAXEL	Cytotoxic	CAPECITABINE + IRINOTECAN + PANITUMUMAB	A
CAP	Cytotoxic	CAPECITABINE + MITOMYCIN	A
CAPE	Cytotoxic	CAPECITABINE + VINORELBINE	A
CAPECITABINE	Cytotoxic	CARBO + FLUOROURACIL	A
CARBO	Cytotoxic	CARDAMON TRIAL	A
CARBOPLATIN	Cytotoxic	CARFILZOMIB	A

CARMUSTINE	Cytotoxic	CEMPIIMAB	A
CAV	Cytotoxic	CERITINIB	A
CCEP	Cytotoxic	CHLORAMBUCIL	A
CETUXIMAB	Cytotoxic	CHLORAMBUCIL + VINBLASTINE	A
CEV	Cytotoxic	COSMIC TRIAL	A
CHLORAMBUCIL	Cytotoxic	CRIZOTINIB	A
CHLVPP	Cytotoxic	CYCLO + DOXORUBICIN + PACLITAXEL	A
CHLVPP/EVA	Cytotoxic	CYCLO + DOXORUBICIN + VINCRIPTINE	A
CHOEP	Cytotoxic	DABRAFENIB	A
CHOP R	Cytotoxic	DABRAFENIB + TRAMETINIB	A
CHOP	Cytotoxic	DACOMITINIB	A
CHOP-R	Cytotoxic	DACTINOMYCIN	A
CISPLATIN	Cytotoxic	DASATINIB	A
CLADRIBINE	Cytotoxic	DOXORUBICIN EMBOLISATION	A
CLODRONIC ACID	Cytotoxic	DURVALUMAB	A
CMF	Cytotoxic	EMA/CO	A
CMV	Cytotoxic	ENRICH TRIAL	A
CNS SIOP 2009	Cytotoxic	ENZALUTAMIDE	A
CODOX M	Cytotoxic	EPIRUBICIN INTRAVESICULAR	A
COMPARE	Cytotoxic	ERLOTINIB	A
CTD	Cytotoxic	EVEROLIMUS	A
CVAD	Cytotoxic	EVEROLIMUS + LENVATINIB	A
CVD	Cytotoxic	FBC	A
CVP R	Cytotoxic	FLUOROURACIL	A
CVP	Cytotoxic	FLUOROURACIL + MITOMYCIN	A
CYCLO	Cytotoxic	FOCUS4 TRIAL	A
CYCLOPHOSHAMIDE	Cytotoxic	Fulvestrant + Palbociclib	A
CYCLOPHOSPHAMIDE	Cytotoxic	GEFITINIB	A
CYTARABINE (VYXEOS)	Cytotoxic	GILTERITINIB	A
CYTARABINE HIGH DOSE	Cytotoxic	HYDROXYCARBAMIDE	A
CYTARABINE	Cytotoxic	ICON9 TRIAL	A
CYTARABINEIT	Cytotoxic	IMMUNOGLOBULIN	A
DA	Cytotoxic	IRINOTECAN	A
DACTINOMYCIN	Cytotoxic	IRINOTECAN + MDG	A

DA-EPOCH	Cytotoxic	IRINOTECAN + MDG + PANITUMUMAB	A
DARATUMUMAB	Cytotoxic	IRINOTECAN + PANITUMUMAB + MDG	A
DAUNORUBICIN	Cytotoxic	IXAZOMIB	A
DE ANGELIS	Cytotoxic	IXAZOMIB + LENALIDOMIDE	A
DHAP	Cytotoxic	LANREOTIDE	A
DICE	Cytotoxic	LENALIDOMIDE	A
DOCETAXEL	Cytotoxic	LENVATINIB	A
DOX	Cytotoxic	LETROZOLE + RIBOCICLIB	A
DTPACE	Cytotoxic	LORLATINIB	A
EC	Cytotoxic	LUTETIUM-177	A
ECARBOX	Cytotoxic	MEDROXYPROGESTERONE	A
ECF	Cytotoxic	METHOTREXATE INTRATHECAL	A
ECX	Cytotoxic	MIDOSTAURIN	A
EMA/CO	Cytotoxic	MIFAMURTIDE	A
EOF	Cytotoxic	NELARABINE	A
EOX	Cytotoxic	NEMO TRIAL	A
EP	Cytotoxic	NERATINIB	A
EP/EMA	Cytotoxic	NILOTINIB	A
EPIRUBICIN	Cytotoxic	NINTEDANIB	A
ERIBULIN	Cytotoxic	NIRAPARIB	A
ERWINIA ASPARAGINASE	Cytotoxic	NIVOLUMAB	A
ESCALATED BEACOP DAC	Cytotoxic	OCTREOTIDE	A
ESHAP	Cytotoxic	OLAPARIB	A
ETOPOSIDE	Cytotoxic	OSIMERTINIB	A
EURAMOS	Cytotoxic	PACLITAXEL	A
EURO-EWING	Cytotoxic	PACLITAXEL + PERTUZUMAB + TRASTUZUMAB	A
EW REL REECUR IFOSFAMIDE	Cytotoxic	PACLITAXEL + TRASTUZUMAB	A
FAB LMB	Cytotoxic	PACRITINIB	A
FCARBOST	Cytotoxic	PALBOCICLIB	A
FCIST	Cytotoxic	PANITUMUMAB	A
FCR	Cytotoxic	PARTNER TRIAL	A
FEC	Cytotoxic	PATHOS TRIAL	A
FLAG	Cytotoxic	PAZOPANIB	A
FLUDARABINE	Cytotoxic	PEG INTERFERON	A

FLUOROURACIL	Cytotoxic	PEMBROLIZUMAB	A
FOXFIRE	Cytotoxic	PEMBROLIZUMAB + PEMETREXED	A
FU	Cytotoxic	PEMETREXED	A
G-CVP	Cytotoxic	PERTUZUMAB	A
GDCVP	Cytotoxic	PERTUZUMAB + TRASTUZUMAB	A
GDP	Cytotoxic	POMALIDOMIDE	A
GEMCARBO	Cytotoxic	PONATINIB	A
GEMCITABINE	Cytotoxic	PRISM TRIAL	A
HCX	Cytotoxic	PROCARBAZINE	A
HD EURONET-PHL-C2 OEPA	Cytotoxic	QUIZARTINIB	A
HLH INDUCTION W3-8 ETOPOSIDE/ DEXAMETHASONE/CICLOSPORIN	Cytotoxic	RADIUM 223	A
HLHCONSOLETOP/DEXAMTHASONE/ CICLOSPORIN	Cytotoxic	RALTITREXED	A
HYDROXYCARBAMIDE	Cytotoxic	REGORAFENIB	A
ICE	Cytotoxic	RIBOCICLIB	A
ICE-R	Cytotoxic	ROMICAR TRIAL	A
ICON8BTRIAL	Cytotoxic	RUXOLITINIB	A
IDARUBICIN	Cytotoxic	SORAFENIB	A
IFOSFAMIDE	Cytotoxic	SUNITINIB	A
IGEV	Cytotoxic	THALIDOMIDE	A
INCA	Cytotoxic	TOPOTECAN	A
INTERLACE	Cytotoxic	TRABECTEDIN	A
IPM	Cytotoxic	TRAMETINIB	A
IPO	Cytotoxic	TRASTUZUMAB	A
IRINOTECAN	Cytotoxic	TRIPLE INTRATHECAL	A
IVA	Cytotoxic	VINBLASTINE	A
IVAC	Cytotoxic	VINORELBINE	A
IVADO	Cytotoxic	ABVD R	B
IVE-R	Cytotoxic	ACCEPT TRIAL	B
LACE	Cytotoxic	ACE (GERM CELL)	B
LEAM	Cytotoxic	ACTICCA-1 TRIAL	B
LENALIDOMIDE	Cytotoxic	ATEZOLIZUMAB	B
LI-1	Cytotoxic	ATEZOLIZUMAB + BEVACIZUMAB	B
LIPOSOMAL DAUNORUBICIN	Cytotoxic	ATEZOLIZUMAB + BEVACIZUMAB + CARBOPLATIN + PACLITAXEL	B
LIPOSOMAL DOXORUBICIN	Cytotoxic	ATEZOLIZUMAB + CARBOPLATIN + ETOPOSIDE	B

LOMUSTINE	Cytotoxic	ATEZOLIZUMAB + NAB PACLITAXEL	B
LOPP	Cytotoxic	AZACITIDINE	B
MA	Cytotoxic	AZACITIDINE + VENETOCLAX	B
MACE	Cytotoxic	BEACOPP	B
MAP	Cytotoxic	BENDAMUSTINE	B
MDG	Cytotoxic	BENDAMUSTINE + BRENTUXIMAB	B
MELPHALAN	Cytotoxic	BENDAMUSTINE + GEMCITABINE + VINORELBINE	B
MEOC	Cytotoxic	BENDAMUSTINE + OBINUTUZUMAB	B
MERCAPTOPURINE	Cytotoxic	BENDAMUSTINE + RITUXIMAB	B
METHOTREXATE	Cytotoxic	BENDAMUSTINE + THALIDOMIDE	B
MIC	Cytotoxic	BEVACIZUMAB	B
MIDAC	Cytotoxic	BEVACIZUMAB + CARBO + PACLITAXEL	B
MITOMYCIN	Cytotoxic	BEVACIZUMAB + CARBOPLATIN + NAB-PACLITAXEL	B
MITOMYCININ TRAVESICULAR	Cytotoxic	BEVACIZUMAB + CARBOPLATIN + PACLITAXEL	B
MITOTANE	Cytotoxic	BEVACIZUMAB + CISPLATIN + PEMETREXED	B
MMM	Cytotoxic	BEVACIZUMAB + IRINOTECAN + MDG	B
MPT	Cytotoxic	BLEOMYCIN + CARBOPLATIN + ETOPOSIDE	B
MPV	Cytotoxic	BLEOMYCIN + CISPLATIN + ETOPOSIDE	B
MTX	Cytotoxic	BLEOMYCIN + CISPLATIN + VINCISTINE	B
MUK7	Cytotoxic	BLEOMYCIN + CYCLOPHOSPHAMIDE + ETOPOSIDE + GEMCITABINE + RITUXIMAB + VINBLASTINE	B
MUK9	Cytotoxic	BLINATUMOMAB	B
MVAC	Cytotoxic	BORTEZOMIB + DARATUMUMAB	B
NAB-PACLITAXEL	Cytotoxic	BORTEZOMIB + MELPHALAN	B
NELARABINE	Cytotoxic	BRENTUXIMAB	B
NEO-AEGIS	Cytotoxic	CABAZITAXEL	B
NHL 2003	Cytotoxic	CAP	B
NINTEDANIB	Cytotoxic	CAP + RT	B
NORDIC	Cytotoxic	CAPE + OXALIPLATIN + TRASTUZUMAB	B
OBINUTUZUMAB	Cytotoxic	CAPECITABINE + CARBOPLATIN	B
OEPA	Cytotoxic	CAPECITABINE + CARBOPLATIN + RT	B
OXALIPATIN	Cytotoxic	CAPECITABINE + CARBOPLATIN + TRASTUZUMAB	B
OXALIPLATIN	Cytotoxic	CAPECITABINE + CETUXIMAB + IRINOTECAN	B
PACLITAXEL	Cytotoxic	CAPECITABINE + CETUXIMAB + OXALIPATIN	B
PAD	Cytotoxic	CAPECITABINE + CISPLATIN	B

PANITUMUMAB	Cytotoxic	CAPECITABINE + CISPLATIN + EPIRUBICIN	B
PATHOS	Cytotoxic	CAPECITABINE + CISPLATIN + RT	B
PCV	Cytotoxic	CAPECITABINE + EPIRUBICIN + OXALIPLATIN	B
PEGASPARAGINASE	Cytotoxic	CAPECITABINE + GEMCITABINE	B
PEGASPARGASE	Cytotoxic	CAPECITABINE + LOMUSTINE	B
PEI	Cytotoxic	CAPECITABINE + MITOMYCIN + RT	B
PEMBROLIZUMAB	Cytotoxic	CAPECITABINE + OXALIPLATIN	B
PENTOSTATIN	Cytotoxic	CAPECITABINE + OXALIPLATIN + PANITUMUMAB	B
PERTUZUMAB	Cytotoxic	CAPECITABINE + RT	B
PHITT	Cytotoxic	CAPECITABINE + TEMOZOLOMIDE	B
PIXANTRONE	Cytotoxic	CARBO + DOCETAXEL + TRASTUZUMAB	B
PMITCEBO-R	Cytotoxic	CARBOPLATIN	B
POLEM	Cytotoxic	CARBOPLATIN + CETUXIMAB + FLUOROURACIL	B
POMB/ACE	Cytotoxic	CARBOPLATIN + CETUXIMAB + FU	B
POUT	Cytotoxic	CARBOPLATIN + CYCLOPHOSPHAMIDE + DOXORUBICIN + PACLITAXEL	B
PRIMUS 002	Cytotoxic	CARBOPLATIN + CYCLOPHOSPHAMIDE + EPIRUBICIN + PACLITAXEL	B
PROCARBAZINE	Cytotoxic	CARBOPLATIN + CYCLOPHOSPHAMIDE + EPIRUBICIN + PACLITAXEL	B
R CODOX	Cytotoxic	CARBOPLATIN + DOCETAXEL	B
R IVAC	Cytotoxic	CARBOPLATIN + DOCETAXEL + FLUOROURACIL	B
RALTITREXED	Cytotoxic	CARBOPLATIN + DOCETAXEL + PERTUZUMAB + TRASTUZUMAB	B
RCD	Cytotoxic	CARBOPLATIN + DOCETAXEL + TRASTUZUMAB	B
RCEOP	Cytotoxic	CARBOPLATIN + EPIRUBICIN + FLUOROURACIL	B
R-CHOP	Cytotoxic	CARBOPLATIN + EPIRUBICIN + PACLITAXEL	B
R-DHAP	Cytotoxic	CARBOPLATIN + ETOPOSIDE	B
R-EPOCH	Cytotoxic	CARBOPLATIN + ETOPOSIDE + IFOSFAMIDE	B
R-GEM-P	Cytotoxic	CARBOPLATIN + ETOPOSIDE + IFOSFAMIDE + RITUXIMAB	B
RITUXIMAB	Cytotoxic	CARBOPLATIN + FLUOROURACIL + TRASTUZUMAB	B
RMS 2005	Cytotoxic	CARBOPLATIN + GEMCITABINE + PACLITAXEL	B
ROSCO	Cytotoxic	CARBOPLATIN + LIPOSOMAL DOX	B
SCALOP-2	Cytotoxic	CARBOPLATIN + NAB-PACLITAXEL	B
SCOPE2	Cytotoxic	CARBOPLATIN + PACLITAXEL	B
SHAMASH LBCMVD-56	Cytotoxic	CARBOPLATIN + PACLITAXEL + RT	B
SMILE	Cytotoxic	CARBOPLATIN + PEMBROLIZUMAB + PEMETREXED	B

STAMPEDE	Cytotoxic	CARBOPLATIN + PEMETREXED	B
STS RHABDO VIT0910	Cytotoxic	CARBOPLATIN + RT	B
TEMOZOLOMIDE	Cytotoxic	CARBOPLATIN + VINORELBINE	B
THIOTEPA	Cytotoxic	CARBOPLATIN + VINORELBINE + RT	B
TIDE	Cytotoxic	CARMUSTINE + THIOTEPA	B
TIP	Cytotoxic	CAV	B
TIPIRACIL	Cytotoxic	CETUX MAB	B
TOPOTECAN	Cytotoxic	CETUX MAB + IRINOTECAN	B
TRANSPLANT ALEMTUZUMAB	Cytotoxic	CETUX MAB + IRINOTECAN + MDG	B
TRANSPLANT CYTARABINE	Cytotoxic	CETUX MAB + MDG	B
TRASTUZUMAB	Cytotoxic	CETUX MAB + OXALIPLATIN + MDG	B
TRETINOIN	Cytotoxic	CETUX MAB + RT	B
TRIFLURIDINE TIPIRACIL	Cytotoxic	CHLORAMBUCIL + ETOPOSIDE + LOMUSTINE	B
TRIFLURIDINE	Cytotoxic	CHLORAMBUCIL + LOMUSTINE	B
TRIPLE INTRATHECAL	Cytotoxic	CHLORAMBUCIL + OBINUTUZUMAB	B
UKALL2003	Cytotoxic	CHLORAMBUCIL + RITUXIMAB	B
UKALL2011	Cytotoxic	CHLVPP/EVA	B
UKALL2014	Cytotoxic	CHOEP	B
UKALL60	Cytotoxic	CHOP + OBINUTUZUMAB	B
UKALLR3	Cytotoxic	CHOP R + METHOTREXATE	B
UKALLXII	Cytotoxic	CISPLATIN	B
VAI	Cytotoxic	CISPLATIN + CYCLO + DOXORUBICIN	B
VDC	Cytotoxic	CISPLATIN + DACTINOMYCIN + ETOPOSIDE + FILGRASTIM + METHOTREXATE	B
VDT-PACE	Cytotoxic	CISPLATIN + DOCETAXEL	B
VEDEX	Cytotoxic	CISPLATIN + DOCETAXEL + FLUOROURACIL	B
VEIP	Cytotoxic	CISPLATIN + DOX + ETOPOSIDE	B
VEPEMB	Cytotoxic	CISPLATIN + DOXORUBICIN	B
VIDE	Cytotoxic	CISPLATIN + DOXORUBICIN + ETOPOSIDE	B
VINBLASTINE	Cytotoxic	CISPLATIN + EPIRUBICIN + FLUOROURACIL	B
VINCRISTINE	Cytotoxic	CISPLATIN + ETOPOSIDE	B
VINORELBINE	Cytotoxic	CISPLATIN + ETOPOSIDE + PACLITAXEL	B
VIP	Cytotoxic	CISPLATIN + ETOPOSIDE + RT	B
VNCRISTINE	Cytotoxic	CISPLATIN + FLUOROURACIL	B
WILMS REL ICE	Cytotoxic	CISPLATIN + FLUOROURACIL + RT	B

WT RELAPSE CARBO/CYCLO/ETOP	Cytotoxic	CISPLATIN + FU + TRASTUZUMAB	B
ABVD	B/T-Cell Antibody	CISPLATIN + GEMCITABINE	B
ACCEPT	B/T-Cell Antibody	CISPLATIN + GEMCITABINE + RITUXIMAB	B
ALL R T-CELL CONSOLIDATION	B/T-Cell Antibody	CISPLATIN + IFOSFAMIDE + PACLITAXEL	B
BEAM	B/T-Cell Antibody	CISPLATIN + LOMUSTINE + VINCRIStINE	B
BLINATUMOMAB	B/T-Cell Antibody	CISPLATIN + PEMBROLIZUMAB + PEMETREXED	B
BRIGATINIB	B/T-Cell Antibody	CISPLATIN + PEMETREXED	B
CVD	B/T-Cell Antibody	CISPLATIN + RT	B
DA-EPOCH	B/T-Cell Antibody	CISPLATIN + VINORELBINE	B
DHAP	B/T-Cell Antibody	CISPLATIN + VINORELBINE + RT	B
ELOQUENT2	B/T-Cell Antibody	CLADRIBINE	B
ESHAP	B/T-Cell Antibody	CMV	B
FCR	B/T-Cell Antibody	COMPARE TRIAL	B
FLAIR	B/T-Cell Antibody	CTD	B
GDP	B/T-Cell Antibody	CVP + OBINUTUZUMAB	B
ICE-R	B/T-Cell Antibody	CVP R	B
IGEV	B/T-Cell Antibody	CVP R + GEMCITABINE	B
IVE-R	B/T-Cell Antibody	CYCLO + DOCETAXEL + TRASTUZUMAB	B
MUK9	B/T-Cell Antibody	CYCLO + RITUXIMAB + VINCRIStINE	B
NORDIC	B/T-Cell Antibody	CYCLOPHOSPHAMIDE	B
OBINUTUZUMAB	B/T-Cell Antibody	CYCLOPHOSPHAMIDE + DOCETAXEL	B
PETREA	B/T-Cell Antibody	CYCLOPHOSPHAMIDE + DOCETAXEL + DOXORUBICIN	B
PMITCEBO-R	B/T-Cell Antibody	CYCLOPHOSPHAMIDE + DOCETAXEL + DOXORUBICIN + PERTUZUMAB + TRASTUZUMAB	B
R CODOX	B/T-Cell Antibody	CYCLOPHOSPHAMIDE + DOCETAXEL + DOXORUBICIN + TRASTUZUMAB	B
R IVAC	B/T-Cell Antibody	CYCLOPHOSPHAMIDE + DOCETAXEL + EPIRUBICIN	B
RCD	B/T-Cell Antibody	CYCLOPHOSPHAMIDE + DOCETAXEL + EPIRUBICIN + TRASTUZUMAB	B
RCEOP	B/T-Cell Antibody	CYCLOPHOSPHAMIDE + DOCETAXEL + PERTUZUMAB + TRASTUZUMAB	B
R-CHOP	B/T-Cell Antibody	CYCLOPHOSPHAMIDE + DOCETAXEL + TRASTUZUMAB	B
R-DHAP	B/T-Cell Antibody	CYCLOPHOSPHAMIDE + DOXORUBICIN	B
R-EPOCH	B/T-Cell Antibody	CYCLOPHOSPHAMIDE + DOXORUBICIN + RITUXIMAB	B
R-GEM-P	B/T-Cell Antibody	CYCLOPHOSPHAMIDE + EPIRUBICIN + FLUOROURACIL + PERTUZUMAB + TRASTUZUMAB	B
ACALABRUTINIB	Targeted	CYCLOPHOSPHAMIDE + ETOPOSIDE + LOMUSTINE	B
AFATINIB	Targeted	CYCLOPHOSPHAMIDE + ETOPOSIDE + PROCARBAZINE	B

ALECTINIB	Targeted	CYCLOPHOSPHAMIDE + GEMCITABINE + RITUXIMAB + VINCRISTINE	B
ASCIMINIB	Targeted	CYCLOPHOSPHAMIDE + GEMCITABINE + VINCRISTINE	B
AVAPRITINIB	Targeted	CYCLOPHOSPHAMIDE + LENALIDOMIDE	B
AXITINIB	Targeted	CYCLOPHOSPHAMIDE + MTX	B
BINIMETINIB	Targeted	CYCLOPHOSPHAMIDE + NAB-PACLITAXEL	B
BORTEZOMIB	Targeted	CYCLOPHOSPHAMIDE + RITUXIMAB	B
BOSUTINIB	Targeted	CYCLOPHOSPHAMIDE + RITUXIMAB + VINCRISTINE	B
BRENTUXIMAB	Targeted	CYCLOPHOSPHAMIDE + TOPOTECAN	B
CABOZANTINIB	Targeted	CYCLOPHOSPHAMIDE + VINCRISTINE	B
CARFILZOMIB	Targeted	CYCLOPHOSPHAMIDE + VINORELBINE	B
CERITINIB	Targeted	DACARBAZINE + DOXORUBICIN	B
COLUMBUS	Targeted	DACARBAZINE + DOXORUBICIN + VINBLASTINE	B
COMBI-AD	Targeted	DACARBAZINE + DOXORUBICIN + VINCRISTINE	B
COSMIC	Targeted	DARATUMUMAB	B
CRIZOTINIB	Targeted	DARATUMUMAB + LENALIDOMIDE	B
DACOMITINIB	Targeted	DAUNORUBICIN + PEGASPARAGINASE + VINCRISTINE	B
DASATINIB	Targeted	DAUNORUBICIN + VNCRISTINE	B
DICE	Targeted	DE ANGELIS	B
DINUTUXIMAB	Targeted	DHAP - R	B
DTPACE	Targeted	DOCETAXEL	B
ENCORAFENIB	Targeted	DOCETAXEL + FLUOROURACIL + OXALIPLATIN	B
ERLOTINIB	Targeted	DOCETAXEL + GEMCITABINE	B
EVEROLIMUS	Targeted	DOCETAXEL + NINTEDANIB	B
FLAIR	Targeted	DOCETAXEL + PERTUZUMAB + TRASTUZUMAB	B
FOCUS4	Targeted	DOCETAXEL + TRASTUZUMAB	B
FULVESTRANT	Targeted	DOXORUBICIN + IFOSFAMIDE	B
GEFITINIB	Targeted	DOXORUBICIN + METHOTREXATE	B
GEMTUZUMAB OZOGAMICIN	Targeted	DTPACE	B
GEMTUZUMAB	Targeted	EC	B
GILTERITINIB	Targeted	EC + DOCETAXEL	B
HCX	Targeted	EC + DOCETAXEL + PERTUZUMAB	B
IBRUTINIB	Targeted	EC + DOCETAXEL + TRASTUZUMAB	B
ICON9	Targeted	EC + PACLITAXEL	B
INOTUZUMAB	Targeted	EC + PACLITAXEL + PERTUZUMAB	B

IVOSIDENIB	Targeted	ECARBOX	B
IXAZOMIB	Targeted	ECF	B
LENVATINIB	Targeted	ECX	B
LI-1	Targeted	EOF	B
LORLATINIB	Targeted	EOX	B
MIDOSTAURIN	Targeted	EP	B
MUK9	Targeted	ERIBULIN	B
NEMO	Targeted	ESCALATED BEACOP DAC	B
NERATINIB	Targeted	ESHAP R	B
NILOTINIB	Targeted	ETOPOSIDE	B
NINTEDANIB	Targeted	ETOPOSIDE + IFOSFAMIDE	B
NIRAPARIB	Targeted	ETOPOSIDE + IFOSFAMIDE + RITUXIMAB + THIOTEPA	B
OLAPARIB	Targeted	ETOPOSIDE + MELPHALAN	B
OLYMPIA	Targeted	EURAMOS TRIAL	B
OSIMERTINIB	Targeted	EURO-EWING	B
PACRITINIB	Targeted	FCARBOST	B
PALBOCICLIB	Targeted	FCR	B
PANITUMUMAB	Targeted	FEC	B
PANOBINOSTAT	Targeted	FEC + DOCETAXEL	B
PARTNER	Targeted	FEC + DOCETAXEL + PERTUZUMAB + TRASTUZUMAB	B
PAZOPANIB	Targeted	FEC + DOCETAXEL + TRASTUZUMAB	B
PEMETREXED	Targeted	FEC + TRASTUZUMAB	B
PHITT	Targeted	FLAIR TRIAL	B
PONATINIB	Targeted	FLUDARABINE + MELPHALAN	B
PVD	Targeted	FLUOROURACIL + IRINOTECAN + OXALIPLATIN	B
QUIZARTINIB	Targeted	FLUOROURACIL + MITOMYCIN + RT	B
REGORAFENIB	Targeted	FOXFIRE TRIAL	B
RIBOCICLIB	Targeted	G-CVP	B
RUXOLITINIB	Targeted	GDCVP	B
SORAFENIB	Targeted	GDP R	B
SUNITINIB	Targeted	GEMCITABINE	B
TALAZOPARIB	Targeted	GEMCITABINE + NAB-PACLITAXEL	B
TIVOZANIB	Targeted	GEMCITABINE + OXALIPLATIN	B
TRAMETINIB	Targeted	GEMCITABINE + OXALIPLATIN + RITUXIMAB	B

TRASTUZUMAB EMTANSINE	Targeted	GEMCITABINE + PACLITAXEL	B
VANDETANIB	Targeted	GEMCITABINE + RT	B
VEMURAFENIB	Targeted	GEMTUZUMAB OZOGAMICIN	B
VENETOCLAX	Targeted	HGX	B
VISMODEGIB	Targeted	HLH Induction W3-8 Etoposide/Dexamethasone/Ciclosporin	B
AVELUMAB	Immunotherapy	IBRUTINIB	B
BCG	Immunotherapy	IBRUTINIB + RITUXIMAB	B
CEMPIIMAB	Immunotherapy	IBRUTINIB + VENETOCLAX	B
COMPARE	Immunotherapy	ICE	B
CONFIRM	Immunotherapy	ICE-R	B
COSMIC	Immunotherapy	Idarubicin + Tre inoin	B
CTD	Immunotherapy	IDELALISIB	B
DURVALUMAB	Immunotherapy	IFOSFAMIDE	B
INCA	Immunotherapy	IFOSFAMIDE + LIPOSOMAL DOXORUBICIN	B
KYMRIAH	Immunotherapy	IMATINIB	B
MPT	Immunotherapy	INOTUZUMAB OZOGAMICIN	B
MUK7	Immunotherapy	INTERLACE TRIAL	B
MUK9	Immunotherapy	IPILIMUMAB	B
NIVOLUMAB	Immunotherapy	IPILIMUMAB + NIVOLUMAB	B
POLEM	Immunotherapy	IRINOTECAN + TEMOZOLOMIDE	B
POMALIDOMIDE	Immunotherapy	IVA	B
PRISM	Immunotherapy	IVAC	B
TALIMOGENE LAHERPAREPVEC	Immunotherapy	IVE-R	B
YESCARTA	Immunotherapy	Kymriah + Fludarabine + Cyclophosphamide	B
ABIRATERONE	Hormonal	LACE	B
APALUTAMIDE	Hormonal	LENALIDOMIDE + RITUXIMAB	B
ENZALUTAMIDE	Hormonal	LIPOSOMAL DOXORUBICIN	B
LETROZOLE	Hormonal	LOMUSTINE	B
MEDROXYPROGESTERONE	Hormonal	LOMUSTINE + PROCARBAZINE	B
PARAGON	Hormonal	LOMUSTINE + PROCARBAZINE + VINCRIStINE	B
STAMPEDE	Hormonal	MAP	B
UNIRAD	Hormonal	MATRIX TRIAL	B
		MELPHALAN	B
		MERCAPTOPURINE	B

METHOTREXATE	B
METHOTREXATE + PEGASPARGASE	B
METHOTREXATE + RITUXIMAB	B
METHOTREXATE + RITUXIMAB + TEMOZOLOMIDE	B
METHOTREXATE + VINBLASTINE	B
METHOTREXATE HIGH DOSE + RITUXIMAB	B
MITOTANE	B
MPT	B
MPV	B
MUK 9 TRIAL	B
MVAC	B
MYELOMA XII TRIAL	B
NAB-PACLITAXEL	B
NAB-PACLITAXEL + TRASTUZUMAB	B
NEO-AEGIS TRIAL	B
NHL 2003	B
NORDIC	B
OBINUTUZUMAB	B
OEPA	B
OXALIPLATIN	B
OXALIPLATIN + MDG	B
OXALIPLATIN + MDG + PANITUMUMAB	B
OXALIPLATIN + RALTITREXED	B
PCV	B
PEI	B
PENTOSTATIN	B
PETREA TRIAL	B
POLEM TRIAL	B
PREDNISOLONE + VINBLASTINE + RITUXIMAB	B
PVD	B
R CODOX-M/ R IVAC	B
R IVAC	B
RCEOP	B
R-EPOCH	B

R-GEM-P	B
RITUXIMAB	B
RITUXIMAB + VENETOCLAX	B
RMS 2005 TRIAL	B
ROSCO TRIAL	B
SCOPE2 TRIAL	B
SHAMASH LBCMVD-56	B
STAMPEDE TRIAL	B
TEMOZOLOMIDE	B
TEMOZOLOMIDE + RT	B
TIVOZANIB	B
TRANSPLANT ALEMTUZUMAB + FLUDARABINE + MELPHALAN	B
TRASTUZUMAB EMTANSINE	B
VDC	B
VDT-PACE	B
VEDEX	B
VENETOCLAX	B
VEPEMB	B
VIP	B
ALEMTUZUMAB	C
ALEMTUZUMAB + BUSULFAN + CYCLOPHOSPHAMIDE	C
ALEMTUZUMAB + CYCLO + FLUDARABINE	C
ALEMTUZUMAB + CYCLOPHOSPHAMIDE	C
ALEMTUZUMAB + ETOPOSIDE	C
ALEMTUZUMAB + FLUDARABINE + MELPHALAN	C
AML Relapsed FLA - Idarubicin	C
AML19 TRIAL	C
AMSACRINE + BUSULFAN + CYTARABINE + FLUDARABINE	C
ATG + BUSULFAN + FLUDARABINE	C
ATG + BUSULFAN + FLUDARABINE + THIOTEPA	C
BENDAMUSTINE + CYTARABINE + RITUXIMAB	C
BEP	C
BUSULFAN + CYCLOPHOSPHAMIDE + FLUDARABINE + THIOTEPA	C
BUSULFAN + FLUDARABINE + METHOTREXATE	C

BUSULPHAN	C
CHOP-R + CYTARABINE HIGH DOSE	C
CODOX M	C
CYCLOPHOSPHAMIDE HIGH DOSE	C
CYTARABINE	C
CYTARABINE + DAUNORUBICIN + GEMTUZUMAB OZOGAMICIN	C
CYTARABINE + ETOPOSIDE + METHOTREXATE	C
CYTARABINE + Gemtuzumab Ozogamicin	C
CYTARABINE + METHOTREXATE + RITUXIMAB	C
CYTARABINE + METHOTREXATE + RITUXIMAB + THIOTEPA	C
CYTARABINE + MIDOSTAURIN	C
CYTARABINE + RITUXIMAB	C
CYTARABINE HIGH DOSE	C
CYTARABINE IT + METHOTREXATE	C
DA	C
DA-EPOCH	C
DAUNORUBICIN + CYTARABINE + MIDOSTAURIN	C
FAB LMB	C
FLAG	C
FLAG + IDARUBICIN	C
GEMCITABINE + TIP	C
LEAM	C
TIP	C
TRANSPLANT ALEMTUZUMAB + CYTARABINE + ETOPOSIDE + LOMUSTINE + MELPHALAN	C
TRANSPLANT CYTARABINE + ETOPOSIDE + LOMUSTINE + MELPHALAN	C
TRIFLURIDINE TIPIRACIL	C
Yescarta + Fludarabine + Cyclophosphamide	C

Supplementary Table 16: Systemic anti-cancer therapy (SACT) classifications for temporal evaluation of COVID-19 across the pandemic. SACT regimens were grouped into treatment type and class (A/B/C).

Cancer cohort					
Post-second dose (weeks)	Exposed (PCR-positive)		Not exposed (PCR-negative)		Vaccine Effectiveness (%)
	Post-2 nd dose (n)	Unvaccinated (N)	Post-2 nd dose (n)	Unvaccinated (N)	
-8-0 weeks (pre-dose 2)	2467	31649	269648	465982	86.5
0-8 weeks	1580	31649	245368	465982	90.5
8-16 weeks	5603	31649	248527	465982	66.8
16-24 weeks	8486	31649	217383	465982	42.5
24-32 weeks	2313	31649	52691	465982	35.4
32-40 weeks	268	31649	6563	465982	39.9
Overall	18292	31649	780054	465982	65.5
3-6 months	12513	31649	347414	465982	47.0

Non-cancer population (control)					
Post-second dose (weeks)	Exposed (PCR-positive)		Not exposed (PCR-negative)		Vaccine Effectiveness (%)
	Post-2 nd dose (n)	Unvaccinated (N)	Post-2 nd dose (n)	Unvaccinated (N)	
-8-0 weeks (pre-dose 2)	404256	3835981	9924741	26973145	71.4
0-8 weeks	270977	3835981	9888586	26973145	80.7
8-16 weeks	480592	3835981	9627760	26973145	64.9
16-24 weeks	410608	3835981	7436090	26973145	61.2
24-32 weeks	92645	3835981	2144116	26973145	69.6
32-40 weeks	10011	3835981	222746	26973145	68.4
Overall	1276049	3835981	29751784	26973145	69.8
3-6 months	678073	3835981	12359270	26973145	61.4

Supplementary Table 17: Number of PCR positive and negative test results and vaccine effectiveness following second dose vaccination in the cancer and control population cohort. Vaccine effectiveness was calculated in 8-week intervals.

Vaccine effectiveness post-second dose (overall)

Outcome measure	Exposed (PCR-positive)		Not exposed (PCR-negative)		Vaccine Effectiveness (%)
	Post-2 nd dose (n)	Unvaccinated (N)	Post-2 nd dose (n)	Unvaccinated (N)	
COVID-19 Breakthrough Infection	18292	31649	780054	465982	65.5% (65.1-65.9)
COVID-19 Hospital Admission	837	3227	780054	465982	84.5% (83.6-85.4)
COVID-19 Mortality	560	5139	780054	465982	93.5% (93.0-94.0)

Vaccine effectiveness 3-6 months post-second dose

Outcome measure	Exposed (PCR-positive)		Not exposed (PCR-negative)		Vaccine Effectiveness (%)
	Post-2 nd dose (n)	Unvaccinated (N)	Post-2 nd dose (n)	Unvaccinated (N)	
COVID-19 Breakthrough Infection	12513	31649	347414	465982	47.0% (46.3-47.6)
COVID-19 Hospital Admission	611	3227	347414	465982	74.6% (72.8-76.3)
COVID-19 Mortality	373	5139	347414	465982	90.3% (89.3-91.2)

Supplementary Table 18: Overall vaccine effectiveness against breakthrough infections, COVID-19 hospitalisation and death in the cancer cohort post 2nd vaccine dose and at 3-6 months post 2nd vaccine dose.

<i>Cancer Feature</i>	<i>Dose 2 overall</i>		<i>Dose 2 3-6 months</i>	
	OR (95% CI)	p value	OR (95% CI)	p value
Cancer Stage				
Stage 1	0.40 (0.39-0.42)	<0.0001	0.60 (0.57-0.64)	<0.0001
Stage 2	0.42 (0.40-0.45)	<0.0001	0.66 (0.62-0.70)	<0.0001
Stage 3	0.37 (0.35-0.40)	<0.0001	0.60 (0.56-0.64)	<0.0001
Stage 4	0.27 (0.25-0.30)	<0.0001	0.44 (0.40-0.48)	<0.0001
Cancer Diagnosis / Treatment				
Cancer diagnosis <12 months of data cut-off	0.45 (0.43-0.47)	<0.0001	0.73 (0.69-0.78)	<0.0001
Cancer diagnosis >12 months of data cut-off	0.37 (0.36-0.38)	<0.0001	0.56 (0.55-0.58)	<0.0001
Systemic Anti-Cancer Therapy (SACT)	0.49 (0.47-0.52)	<0.0001	0.80 (0.77-0.84)	<0.0001
No Systemic Anti-Cancer Therapy (SACT)	0.39 (0.38-0.39)	<0.0001	0.58 (0.57-0.60)	<0.0001
SACT <12 months of data cut-off	0.52 (0.50-0.55)	<0.0001	0.85 (0.81-0.90)	<0.0001
SACT >12 months of data cut-off	0.39 (0.38-0.39)	<0.0001	0.59 (0.57-0.60)	<0.0001
Radiotherapy	0.48 (0.46-0.51)	<0.0001	0.76 (0.72-0.81)	<0.0001
No Radiotherapy	0.40 (0.39-0.41)	<0.0001	0.61 (0.60-0.63)	<0.0001
Radiotherapy <12 months of data cut-off	0.54 (0.49-0.59)	<0.0001	0.90 (0.81-0.99)	0.031
Radiotherapy >12 months of data cut-off	0.40 (0.39-0.41)	<0.0001	0.61 (0.60-0.63)	<0.0001
Type of Malignancy				
Solid tumour malignancy	0.38 (0.37-0.39)	<0.0001	0.59 (0.58-0.61)	<0.0001
Haematological malignancy	0.62 (0.59-0.66)	<0.0001	0.95 (0.90-1.01)	0.115
Solid tumour malignancy subtype				
Lip, oral cavity & pharynx (C00-C14)	0.51 (0.44-0.58)	<0.0001	0.78 (0.67-0.91)	0.001
Non-colorectal gastrointestinal (C15-C17, C22-C26)	0.32 (0.29-0.35)	<0.0001	0.51 (0.46-0.57)	<0.0001
Colorectal gastrointestinal (C18-C21)	0.35 (0.33-0.37)	<0.0001	0.57 (0.53-0.61)	<0.0001
Lung (C34)	0.35 (0.33-0.38)	<0.0001	0.54 (0.50-0.59)	<0.0001
Respiratory & intrathoracic organs (C30-C33, C35-C39)	0.38 (0.31-0.48)	<0.0001	0.65 (0.51-0.83)	0.000
Bone and articular cartilage (C40-C41)	0.41 (0.27-0.63)	<0.0001	0.60 (0.36-1.01)	0.055
Mesothelial and soft tissue (C45-C49)	0.48 (0.40-0.58)	<0.0001	0.78 (0.63-0.96)	0.018
Breast (C50)	0.40 (0.38-0.42)	<0.0001	0.59 (0.56-0.63)	<0.0001
Female gynaecological (C51-C58)	0.37 (0.34-0.40)	<0.0001	0.55 (0.50-0.61)	<0.0001
Male urological (C60, C62, C63)	0.59 (0.49-0.71)	<0.0001	0.79 (0.63-0.98)	0.035
Prostate (C61)	0.35 (0.33-0.37)	<0.0001	0.54 (0.51-0.57)	<0.0001
Urinary tract (C64-C68)	0.40 (0.37-0.43)	<0.0001	0.62 (0.57-0.68)	<0.0001
Central nervous system (C69-C72)	0.35 (0.29-0.42)	<0.0001	0.51 (0.41-0.63)	<0.0001
Endocrine glands (C73-C75)	0.42 (0.35-0.50)	<0.0001	0.59 (0.48-0.74)	<0.0001

Cancer Feature	Dose 2 overall		Dose 2 3-6 months	
	OR (95% CI)	p value	OR (95% CI)	p value
Haematological malignancy subtype				
Lymphoma (C81-C85)	0.81 (0.75-0.87)	<0.0001	1.24 (1.14-1.35)	<0.0001
Myeloma (C90)	0.32 (0.28-0.36)	<0.0001	0.51 (0.45-0.59)	<0.0001
Leukaemia (C91-C95)	0.61 (0.55-0.68)	<0.0001	0.91 (0.80-1.02)	0.109

Supplementary Table 19. Multivariable logistic regression, adjusted for clinically important covariables (patient age, sex, ethnicity, and index of deprivation) and effect of cancer diagnosis and treatments. Regression models were done with presence compared with absence of each subgroup (reference). Multivariable vaccine effectiveness may be calculated by 1-Odds Ratio (OR).

Age Group	n	Mean (SD)	Cancer Stage	n	Mean (SD)
18-19	111	2.586 (4.911)	Stage 1	10160	2.039 (7.759)
20-29	889	2.021 (5.167)	Stage 2	6820	1.846 (5.459)
30-39	2364	2.475 (5.703)	Stage 3	6154	1.767 (6.812)
40-49	4948	2.602 (5.041)	Stage 4	3658	1.321 (4.363)
50-59	9841	2.309 (7.252)			
60-69	11785	1.842 (7.148)			
70-79	11913	1.445 (6.530)			
80-89	6700	0.910 (2.505)			
90+	1787	0.579 (1.242)			

Supplementary Table 20: Table demonstrating the relationship between number of contacts in the cancer cohort and effect of patient age (left) and cancer stage (right). Ten-year deciles were used for classifying age groups.

COVID-19 outcome	Post-dose 2 vaccine effectiveness (95% CI)		Post-dose 3 vaccine effectiveness (95% CI)	
	Cancer cohort	Non-cancer population	Cancer cohort	Non-cancer population
Breakthrough Infection	43.49% (43.28-43.71)	53.99% (53.94-54.04)	59.12% (58.75-59.49)	77.25% (77.16-77.34)
Symptomatic Infection	46.08% (45.86-46.30)	55.69% (55.64-55.75)	62.81% (62.43-63.19)	80.53% (80.44-80.62)
COVID-19 Hospitalisation	73.03% (71.20-74.77)	80.34% (80.10-80.57)	80.45% (77.29-83.19)	89.81% (89.42-90.18)
COVID-19 Death	88.88% (87.85-89.83)	83.34% (82.98-83.70)	94.54% (93.11-95.68)	91.64% (91.05-92.19)

Supplementary Table 21: Vaccine effectiveness in cancer cohort and non-cancer population following third dose booster (post dose 3) and following primary vaccination course (post dose 2).

COVID-19 outcome	Dose 1 & 2 BNT162b2, Dose 3 BNT162b2			Dose 1 & 2 ChAdOx1, Dose 3 BNT162b2		
	Post-3rd Dose	Not vaccinated	VE (95% CI)	Post-3rd Dose	Not vaccinated	VE (95% CI)
Breakthrough Infection	9290	174416	59.3% (58.8-59.8)	7260	174416	58.3% (57.7-58.9)
Symptomatic Infection	8368	167503	61.9% (61.4-62.4)	6150	167503	63.2% (62.6-63.8)
COVID-19 Hospitalisation	81	3272	81.1% (76.8-84.6)	68	3272	79.2% (73.9-83.4)
COVID-19 Death	36	5123	94.6% (92.6-96.1)	29	5123	94.3% (91.9-96.0)

Supplementary Table 22: Third dose booster vaccine effectiveness in cancer cohort and effect of different vaccination combinations. Most individuals the UKCCP dataset received either BNT162b2 or ChAdOx1 for primary vaccination (doses 1 and 2) and BNT162b2 for booster vaccination (dose 3).

	Breakthrough infection		Symptomatic infection	
	Post Dose 3	Post Dose 2	Post Dose 3	Post Dose 2
Cancer Diagnosis and Treatment				
Cancer diagnosis <12 months	52.49%	24.21%	58.27%	27.78%
Cancer diagnosis >12 months	65.69%	50.15%	68.40%	52.74%
Systemic Anti-Cancer Therapy	56.77%	35.42%	60.16%	39.24%
No Systemic Anti-Cancer Therapy	63.57%	45.24%	67.00%	47.56%
Systemic Anti-Cancer Therapy <12 months	51.14%	32.65%	54.33%	37.16%
Systemic Anti-Cancer Therapy >12 months	64.23%	33.47%	68.70%	35.66%
Radiotherapy	59.17%	37.82%	63.57%	42.63%
No Radiotherapy	64.46%	47.20%	67.47%	49.68%
Radiotherapy <12 months	43.30%	41.87%	50.23%	47.49%
Radiotherapy >12 months	68.70%	34.56%	71.47%	38.51%
Type of Malignancy				
Solid Tumour malignancy	65.98%	42.34%	69.62%	46.05%
Haematological malignancy	53.25%	40.81%	55.98%	41.78%
Solid tumour malignancy subtype				
Lip oral cavity and pharynx (C00-C14)	60.03%	-55.17%	76.14%	-49.88%
Non-colorectal gastrointestinal (C15-C17, C22-C26)	73.44%	42.99%	74.72%	45.99%
Colorectal gastrointestinal (C18-C21)	62.14%	40.23%	67.09%	43.93%
Respiratory and intrathoracic organs incl. lung (C30-C39)	51.60%	17.33%	55.73%	19.62%
Bone, mesothelial and soft tissue (C40-C41, C45-C49)	73.83%	58.36%	75.96%	62.05%
Breast (C50-C50)	68.94%	39.01%	72.77%	45.78%
Female gynaecological (C51-C58)	81.11%	52.23%	82.77%	54.06%
Male urological (C60, C62-C63)	77.21%	-33.97%	81.93%	-41.18%
Prostate (C61-C61)	79.37%	57.64%	83.45%	60.11%
Urinary tract (C64-C68)	63.77%	49.06%	67.34%	52.54%
Central nervous system (C69-C72)	85.42%	79.07%	88.20%	81.56%
Endocrine glands (C73-C75)	85.26%	60.41%	87.27%	63.48%
Haematological malignancy subtype				
Lymphoma (C81-C85)	10.52%	14.92%	13.60%	17.46%
Myeloma (C90-C90)	67.04%	56.05%	69.96%	56.27%
Leukaemia (C91-C95)	68.72%	37.83%	71.80%	38.82%

Supplementary Table 23: Vaccine effectiveness against breakthrough and symptomatic infections by predefined cancer subgroups (cancer diagnosis, treatment, type, and subtype). Subgroups include tumour stage, anticancer treatment, and primary tumour subtype.

	Combined Cohort	SARS-CoV-2 Negative	Breakthrough Infection	Symptomatic Infection	COVID-19 Hospitalisation	COVID-19 Death
Cancer status						
Cancer cohort	124998	108171	16827	14701	151	66
Non-Cancer population	3929844	3775399	154445	116088	2601	803
Age						
Median - Cancer cohort	71.0 years	71.0 years	69.0 years	70.0 years	72.0 years	76.0 years
Median - Non-cancer population	59.0 years	59.0 years	59.0 years	60.0 years	75.0 years	81.0 years
Sex						
Male	1345178	1260158	85020	68116	1696	577
Female	2709642	2623390	86252	62673	1056	292
Ethnicity						
White/White British	3534731	3387235	147496	112543	2456	775
Asian/Asian British	239847	230856	8991	6966	113	29
Black/Black British	87552	84227	3325	2865	32	11
Mixed/Other Ethnic Group	41416	40137	1279	984	12	6
Social Deprivation Index (IMD)						
IMD=Low (1-3)	1024540	983816	40724	32280	788	262
IMD=Medium (4-7)	1695888	1625441	70447	53581	1108	339
IMD=High (8-10)	1334219	1274120	60099	44927	856	268
Primary Vaccine (Dose 1 & 2)						
BNT162b2 (Pfizer)	2540940	2453467	87473	66770	1452	504
ChAdOx1 (AstraZeneca)	1478398	1396057	82341	62929	1278	352
Booster Vaccine (Dose 3)						
BNT162b2 (Pfizer)	3933435	3772068	161367	123679	2641	840
mRNA-1273 (Moderna)	85903	77456	8447	6020	89	16

Supplementary Table 24: Table showing baseline characteristics of the non-cancer population and cancer cohorts who had received their third dose booster.

	Dose 2	Dose 3	Dose 2 & 3
Cancer Stage			
<i>Stage 1</i>	535 (100.00)	235 (100.00)	770 (100.00)
- COVS Positive	535 (100.00)	231 (98.30)	766 (99.48)
- COVS Negative	0 (0.00)	4 (1.70)	4 (0.52)
- COVS Median	15255.0 (2995.5-25000.0)	24583.0 (10431.0-25000.0)	18250.0 (5078.8-25000.0)
<i>Stage 2</i>	341 (100.00)	181 (100.00)	522 (100.00)
- COVS Positive	339 (99.41)	180 (99.45)	519 (99.43)
- COVS Negative	2 (0.59)	1 (0.55)	3 (0.57)
- COVS Median	13908.0 (2769.0-25000.0)	18768.0 (8298.0-25000.0)	15891.0 (4157.0-25000.0)
<i>Stage 3</i>	246 (100.00)	130 (100.00)	376 (100.00)
- COVS Positive	245 (99.59)	128 (98.46)	373 (99.20)
- COVS Negative	1 (0.41)	2 (1.54)	3 (0.80)
- COVS Median	14934.0 (3123.3-25000.0)	20114.0 (6296.0-25000.0)	18028.0 (3650.8-25000.0)
<i>Stage 4</i>	61 (100.00)	59 (100.00)	120 (100.00)
- COVS Positive	61 (100.00)	59 (100.00)	120 (100.00)
- COVS Negative	0 (0.00)	0 (0.00)	0 (0.00)
- COVS Median	8069.0 (579.0-25000.0)	16821.0 (4429.5-25000.0)	11914.5 (2154.5-25000.0)
Cancer Treatment			
<i>Systemic Anti-Cancer Therapy (SACT)</i>	416 (100.00)	382 (100.00)	798 (100.00)
- COVS Positive	387 (93.03)	336 (87.96)	723 (90.60)
- COVS Negative	29 (6.97)	46 (12.04)	75 (9.40)
- COVS Median	6273.0 (369.9-25000.0)	10395.0 (1486.0-25000.0)	8131.0 (568.0-25000.0)
<i>No Systemic Anti-Cancer Therapy (SACT)</i>	1661 (100.00)	878 (100.00)	2539 (100.00)
- COVS Positive	1648 (99.22)	851 (96.92)	2499 (98.42)
- COVS Negative	13 (0.78)	27 (3.08)	40 (1.58)
- COVS Median	13213.0 (2032.0-25000.0)	19525.5 (7865.3-25000.0)	15443.0 (3323.5-25000.0)
<i>Radiotherapy</i>	314 (100.00)	173 (100.00)	487 (100.00)
- COVS Positive	309 (98.41)	169 (97.69)	478 (98.15)
- COVS Negative	5 (1.59)	4 (2.31)	9 (1.85)
- COVS Median	12943.5 (2233.8-25000.0)	17391.0 (5191.0-25000.0)	14520.0 (2953.0-25000.0)
<i>No Radiotherapy</i>	1763 (100.00)	1087 (100.00)	2850 (100.00)
- COVS Positive	1726 (97.90)	1018 (93.65)	2744 (96.28)
- COVS Negative	37 (2.10)	69 (6.35)	106 (3.72)
- COVS Median	11553.0 (1554.0-25000.0)	16937.0 (5110.0-25000.0)	13645.0 (2299.5-25000.0)

	Dose 2	Dose 3	Dose 2 & 3
Cancer Type			
<i>Solid Tumour Malignancy</i>	1847 (100.00)	944 (100.00)	2791 (100.00)
- COVS Positive	1843 (99.78)	938 (99.36)	2781 (99.64)
- COVS Negative	4 (0.22)	6 (0.64)	10 (0.36)
- COVS Median	13511.0 (2310.0-25000.0)	20649.5 (9413.0-25000.0)	16165.0 (3906.0-25000.0)
<i>Haematological Malignancy</i>	230 (100.00)	316 (100.00)	546 (100.00)
- COVS Positive	192 (83.48)	249 (78.80)	441 (80.77)
- COVS Negative	38 (16.52)	67 (21.20)	105 (19.23)
- COVS Median	1004.5 (19.0-12485.0)	4297.0 (11.4-18206.0)	1872.5 (12.8-16772.0)
Solid Tumour Subtypes			
<i>Lip, oral cavity, and pharynx (C00-C14)</i>	55 (100.00)	30 (100.00)	85 (100.00)
- COVS Positive	55 (100.00)	30 (100.00)	85 (100.00)
- COVS Negative	0 (0.00)	0 (0.00)	0 (0.00)
- COVS Median	14031.0 (2456.5-25000.0)	25000.0 (20288.8-25000.0)	21290.0 (6207.0-25000.0)
<i>Non-colorectal gastrointestinal (C15-C17, C22-C26)</i>	69 (100.00)	34 (100.00)	103 (100.00)
- COVS Positive	69 (100.00)	34 (100.00)	103 (100.00)
- COVS Negative	0 (0.00)	0 (0.00)	0 (0.00)
- COVS Median	14240.0 (2533.0-25000.0)	17698.5 (6586.0-25000.0)	17558.0 (2992.5-25000.0)
<i>Colorectal gastrointestinal (C18-C21)</i>	218 (100.00)	132 (100.00)	350 (100.00)
- COVS Positive	218 (100.00)	131 (99.24)	349 (99.71)
- COVS Negative	0 (0.00)	1 (0.76)	1 (0.29)
- COVS Median	15121.0 (2645.5-25000.0)	19071.5 (7605.8-25000.0)	17179.0 (3887.8-25000.0)
<i>Respiratory and intrathoracic (C30-C39)</i>	67 (100.00)	45 (100.00)	112 (100.00)
- COVS Positive	65 (97.01)	45 (100.00)	110 (98.21)
- COVS Negative	2 (2.99)	0 (0.00)	2 (1.79)
- COVS Median	8487.0 (1361.0-25000.0)	25000.0 (15921.0-25000.0)	18149.5 (5986.5-25000.0)
<i>Bone, mesothelial and soft tissue (C40-C41, C45-C49)</i>	20 (100.00)	14 (100.00)	34 (100.00)
- COVS Positive	20 (100.00)	14 (100.00)	34 (100.00)
- COVS Negative	0 (0.00)	0 (0.00)	0 (0.00)
- COVS Median	11912.0 (1102.8-25000.0)	18472.0 (8094.3-25000.0)	16371.5 (3287.0-25000.0)
<i>Breast (C50)</i>	678 (100.00)	346 (100.00)	1024 (100.00)
- COVS Positive	676 (99.71)	344 (99.42)	1020 (99.61)
- COVS Negative	2 (0.29)	2 (0.58)	4 (0.39)
- COVS Median	12723.5 (1993.3-25000.0)	19243.5 (7722.5-25000.0)	14557.5 (3324.8-25000.0)
<i>Female gynaecological (C51-C58)</i>	138 (100.00)	56 (100.00)	194 (100.00)
- COVS Positive	138 (100.00)	56 (100.00)	194 (100.00)
- COVS Negative	0 (0.00)	0 (0.00)	0 (0.00)
- COVS Median	10906.0 (2376.5-25000.0)	16222.5 (7065.0-25000.0)	12341.5 (2968.0-25000.0)

	Dose 2	Dose 3	Dose 2 & 3
Solid Tumour Subtypes			
<i>Male urological (C60, C62-C63)</i>	23 (100.00)	10 (100.00)	33 (100.00)
- COVS Positive	23 (100.00)	10 (100.00)	33 (100.00)
- COVS Negative	0 (0.00)	0 (0.00)	0 (0.00)
- COVS Median	13474.0 (3665.0-25000.0)	24891.0 (17112.8-25000.0)	20913.0 (4268.0-25000.0)
<i>Prostate (C61)</i>	396 (100.00)	177 (100.00)	573 (100.00)
- COVS Positive	396 (100.00)	175 (98.87)	571 (99.65)
- COVS Negative	0 (0.00)	2 (1.13)	2 (0.35)
- COVS Median	16067.0 (2829.8-25000.0)	24583.0 (11989.0-25000.0)	19680.0 (4539.0-25000.0)
<i>Urinary Tract (C64-C68)</i>	106 (100.00)	68 (100.00)	174 (100.00)
- COVS Positive	106 (100.00)	67 (98.53)	173 (99.43)
- COVS Negative	0 (0.00)	1 (1.47)	1 (0.57)
- COVS Median	17166.5 (2827.8-25000.0)	25000.0 (9927.0-25000.0)	19140.5 (5780.3-25000.0)
<i>Central nervous system (C69-C72)</i>	29 (100.00)	7 (100.00)	36 (100.00)
- COVS Positive	29 (100.00)	7 (100.00)	36 (100.00)
- COVS Negative	0 (0.00)	0 (0.00)	0 (0.00)
- COVS Median	11997.0 (1666.0-18353.0)	10913.0 (7181.0-23361.5)	11942.5 (2065.3-21139.5)
<i>Endocrine glands (C73-C75)</i>	48 (100.00)	25 (100.00)	73 (100.00)
- COVS Positive	48 (100.00)	25 (100.00)	73 (100.00)
- COVS Negative	0 (0.00)	0 (0.00)	0 (0.00)
- COVS Median	12501.5 (1809.5-25000.0)	17828.0 (9296.0-25000.0)	15066.0 (4327.0-25000.0)
Haematological Subtypes			
<i>Lymphoma (C81-C85)</i>	125 (100.00)	153 (100.00)	278 (100.00)
- COVS Positive	103 (82.40)	105 (68.63)	208 (74.82)
- COVS Negative	22 (17.60)	48 (31.37)	70 (25.18)
- COVS Median	736.0 (6.0-17740.0)	799.0 (0.4-9918.0)	777.5 (0.7-11638.0)
<i>Myeloma (C90)</i>	33 (100.00)	60 (100.00)	93 (100.00)
- COVS Positive	32 (96.97)	59 (98.33)	91 (97.85)
- COVS Negative	1 (3.03)	1 (1.67)	2 (2.15)
- COVS Median	5617.0 (849.0-23768.0)	16124.5 (4272.0-25000.0)	10451.0 (2328.0-25000.0)
<i>Leukaemia (C91-C95)</i>	66 (100.00)	91 (100.00)	157 (100.00)
- COVS Positive	52 (78.79)	76 (83.52)	128 (81.53)
- COVS Negative	14 (21.21)	15 (16.48)	29 (18.47)
- COVS Median	631.6 (9.5-3563.3)	5100.0 (137.3-18361.5)	1846.0 (40.0-11385.0)

Supplementary Table 25: Cancer cohort subgroup analyses of median SARS-CoV-2 anti-spike (COVS) antibody titres and responses post-vaccination. Measures of detectable antibody levels (COVS positive or negative) and median quantified titres are shown. Subgroups include cancer stage, anticancer treatment status and primary tumour subtype.

	COVID-19 Breakthrough Infection				COVID-19 Hospitalisation			
	Univariable		Multivariable		Univariable		Multivariable	
	OR (95% CI)	p value	OR (95% CI)	p value	OR (95% CI)	p value	OR (95% CI)	p value
Overall	2.56 (1.69-4.00)	<0.0001	3.05 (1.96-4.72)	<0.0001	5.88 (3.13-11.11)	<0.0001	6.48 (3.31-6.48)	<0.0001
Sensitivity analysis 1								
- Dose 2 only	2.17 (0.97-4.76)	0.0580	2.25 (0.99-5.12)	0.0532	6.67 (2.27-20.00)	0.0006	6.71 (2.21-20.35)	0.0007
- Dose 3 only	2.56 (1.54-4.35)	0.0003	3.16 (1.86-5.37)	<0.0001	5.26 (2.27-12.0)	0.0001	6.28 (2.63-14.99)	<0.0001
Sensitivity analysis 2								
- Solid only	5.88 (1.25-25.00)	0.0250	5.33 (1.05-27.06)	0.0436	50.00 (9.09-inf)	<0.0001	50.50 (9.13-279.19)	<0.0001
- Haem only	1.59 (0.85-3.03)	0.1470	2.04 (1.04-3.98)	0.0376	3.23 (1.27-8.33)	0.0137	3.64 (1.35-9.81)	0.0106

Supplementary Table 26: Univariable and multivariable adjusted odds ratio for the risk of coronavirus breakthrough infection and coronavirus hospitalisation amongst test from cancer patients with undetectable coronavirus (COVS) antibody responses relative to those with positive antibody responses. Multivariable corrected models were fitted for the patient age, sex, ethnicity, and deprivation. Sensitivity analyses were done according to number of vaccine doses and on cancer type.

Demographics	Patients with cancer	Non-cancer population
Age group		
18-19	389 (0.20%)	591296 (3.25%)
20-29	3143 (1.58%)	3223020 (17.72%)
30-39	8703 (4.38%)	3781383 (20.79%)
40-49	17635 (8.87%)	3598039 (19.78%)
50-59	36705 (18.46%)	3200888 (17.60%)
60-69	48417 (24.35%)	1928813 (10.60%)
70-79	51591 (25.95%)	1088001 (5.98%)
80-89	26373 (13.26%)	559060 (3.07%)
90-99	5788 (2.91%)	208561 (1.15%)
100+	75 (0.04%)	9512 (0.05%)
Sex		
Female	98452 (49.52%)	10356254 (56.94%)
Male	100367 (50.48%)	7830844 (43.05%)
Ethnicity		
Asian/Asian British	8053 (4.05%)	1396446 (7.68%)
Black/Black British	4707 (2.37%)	504064 (2.77%)
White/White British	181765 (91.42%)	15253698 (83.86%)
Mixed or Other Ethnicity	1725 (0.87%)	311899 (1.71%)
Indices of Deprivation (IMD)		
Low (IMD 1-3)	45788 (23.03%)	5007809 (27.53%)
Medium (IMD 4-7)	75897 (38.17%)	7248607 (39.85%)
High (IMD 8-10)	62861 (31.62%)	5514627 (30.32%)
SARS-CoV-2 Vaccination		
Vaccinated (1+ Dose)	156187 (78.56%)	14028232 (77.13%)
Unvaccinated (0 Doses)	42632 (21.44%)	4160341 (22.87%)

Supplementary Table 27: Baseline demographics of the UKCCP cancer cohort and non-cancer population control at the SARS-CoV-2 positive test level. Percentages of the total cancer or non-cancer cohorts are shown in brackets.

Hospital Assessment

Time Period	<i>Patients with cancer</i>		<i>Non-cancer population</i>	
	Outcome/Total (n/N)	Outcome % (95% CI)	Outcome/Total (n/N)	Outcome % (95% CI)
Nov-Dec 2020	5732/19209	29.84% (29.20-30.49)	103983/1094943	9.50% (9.44-9.55)
Jan-Feb 2021	7189/20228	35.54% (34.88-36.20)	145402/1173794	12.39% (12.33-12.45)
Mar-Apr 2021	884/2474	35.73% (33.87-37.64)	16146/166778	9.68% (9.54-9.82)
May-Jun 2021	458/1693	27.05% (24.99-29.22)	16460/304909	5.40% (5.32-5.48)
Jul-Aug 2021	2420/10485	23.08% (22.28-23.90)	72421/1481046	4.89% (4.86-4.92)
Sep-Oct 2021	2733/13638	20.04% (19.38-20.72)	68613/1314307	5.22% (5.18-5.26)
Nov-Dec 2021	3154/23592	13.37% (12.94-13.81)	86812/3543980	2.45% (2.43-2.47)
Jan-Feb 2022	5731/38076	15.05% (14.70-15.41)	114847/4273918	2.69% (2.67-2.70)
Mar-Apr 2022	5606/42270	13.26% (12.94-13.59)	109938/3238009	3.40% (3.38-3.42)
May-Jun 2022	2401/11988	20.03% (19.32-20.75)	43897/723862	6.06% (6.01-6.12)
Jul-Aug 2022	2725/15166	17.97% (17.36-18.59)	58562/873027	6.71% (6.66-6.76)

Inpatient Hospitalisation

Time Period	<i>Patients with cancer</i>		<i>Non-cancer population</i>	
	Outcome/Total (n/N)	Outcome % (95% CI)	Outcome/Total (n/N)	Outcome % (95% CI)
Nov-Dec 2020	5036/19209	26.22% (25.60-26.84)	95525/1094943	8.72% (8.67-8.78)
Jan-Feb 2021	6185/20228	30.58% (29.95-31.21)	133795/1173794	11.40% (11.34-11.46)
Mar-Apr 2021	659/2474	26.64% (24.93-28.41)	13845/166778	8.30% (8.17-8.43)
May-Jun 2021	337/1693	19.91% (18.07-21.87)	13878/304909	4.55% (4.48-4.63)
Jul-Aug 2021	2001/10485	19.08% (18.34-19.85)	61994/1481046	4.19% (4.15-4.22)
Sep-Oct 2021	2193/13638	16.08% (15.47-16.71)	59450/1314307	4.52% (4.49-4.56)
Nov-Dec 2021	2267/23592	9.61% (9.24-9.99)	68884/3543980	1.94% (1.93-1.96)
Jan-Feb 2022	3280/38076	8.61% (8.34-8.90)	77858/4273918	1.82% (1.81-1.83)
Mar-Apr 2022	3149/42270	7.45% (7.20-7.70)	74623/3238009	2.30% (2.29-2.32)
May-Jun 2022	1424/11988	11.88% (11.31-12.47)	31203/723862	4.31% (4.26-4.36)
Jul-Aug 2022	1530/15166	10.09% (9.62-10.58)	40089/873027	4.59% (4.55-4.64)

Intensive Care Admission				
Time Period	Patients with cancer		Non-cancer population	
	Outcome/Total (n/N)	Outcome % (95% CI)	Outcome/Total (n/N)	Outcome % (95% CI)
Nov-Dec 2020	436/19209	2.27% (2.07-2.49)	13028/1094943	1.19% (1.17-1.21)
Jan-Feb 2021	514/20228	2.54% (2.33-2.77)	17812/1173794	1.52% (1.50-1.54)
Mar-Apr 2021	49/2474	1.98% (1.50-2.61)	1553/166778	0.93% (0.89-0.98)
May-Jun 2021	32/1693	1.89% (1.34-2.66)	2396/304909	0.79% (0.76-0.82)
Jul-Aug 2021	235/10485	2.24% (1.97-2.54)	8943/1481046	0.60% (0.59-0.62)
Sep-Oct 2021	278/13638	2.04% (1.81-2.29)	7521/1314307	0.57% (0.56-0.59)
Nov-Dec 2021	212/23592	0.90% (0.79-1.03)	7563/3543980	0.21% (0.21-0.22)
Jan-Feb 2022	179/38076	0.47% (0.41-0.54)	3966/4273918	0.09% (0.09-0.10)
Mar-Apr 2022	131/42270	0.31% (0.26-0.37)	2926/3238009	0.09% (0.09-0.09)
May-Jun 2022	60/11988	0.50% (0.39-0.64)	1263/723862	0.17% (0.17-0.18)
Jul-Aug 2022	42/15166	0.28% (0.20-0.37)	1117/873027	0.13% (0.12-0.14)

COVID-19 Mortality				
Time Period	Patients with cancer		Non-cancer population	
	Outcome/Total (n/N)	Outcome % (95% CI)	Outcome/Total (n/N)	Outcome % (95% CI)
Nov-Dec 2020	3961/19209	20.62% (20.05-21.20)	36681/1094943	3.35% (3.32-3.38)
Jan-Feb 2021	4152/20228	20.53% (19.98-21.09)	48839/1173794	4.16% (4.12-4.20)
Mar-Apr 2021	372/2474	15.04% (13.68-16.50)	3384/166778	2.03% (1.96-2.10)
May-Jun 2021	162/1693	9.57% (8.26-11.06)	1776/304909	0.58% (0.56-0.61)
Jul-Aug 2021	732/10485	6.98% (6.51-7.49)	10585/1481046	0.71% (0.70-0.73)
Sep-Oct 2021	907/13638	6.65% (6.24-7.08)	12877/1314307	0.98% (0.96-1.00)
Nov-Dec 2021	891/23592	3.78% (3.54-4.03)	14175/3543980	0.40% (0.39-0.41)
Jan-Feb 2022	1568/38076	4.12% (3.92-4.32)	20470/4273918	0.48% (0.47-0.49)
Mar-Apr 2022	1374/42270	3.25% (3.09-3.42)	19701/3238009	0.61% (0.60-0.62)
May-Jun 2022	534/11988	4.45% (4.10-4.84)	6679/723862	0.92% (0.90-0.94)
Jul-Aug 2022	625/15166	4.12% (3.82-4.45)	9331/873027	1.07% (1.05-1.09)

Supplementary Table 28: Temporal analysis of COVID-19 case-outcome rates in the UKCCP cancer cohort and non-cancer population control. Case-outcome rates (outcome %) were calculated in 2-month intervals from 1st November 2020-31st August 2022. 95% confidence intervals for case-outcome rates are provided in brackets.

Time period	Patients with cancer		Non-cancer population	
	Vaccinated/Total (n/N)	Uptake % (95% CI)	Vaccinated/Total (n/N)	Uptake % (95% CI)
Nov-Dec 2020	115/19209	0.60% (0.50-0.72)	6590/1094943	0.60% (0.59-0.62)
Jan-Feb 2021	3603/20228	17.81% (17.29-18.35)	130256/1173794	11.10% (11.04-11.15)
Mar-Apr 2021	1607/2474	64.96% (63.05-66.81)	58914/166778	35.32% (35.10-35.55)
May-Jun 2021	1379/1693	81.45% (79.53-83.23)	160930/304909	52.78% (52.60-52.96)
Jul-Aug 2021	9493/10485	90.54% (89.96-91.08)	1073603/1481046	72.49% (72.42-72.56)
Sep-Oct 2021	12859/13638	94.29% (93.89-94.67)	1100684/1314307	83.75% (83.68-83.81)
Nov-Dec 2021	22376/23592	94.85% (94.56-95.12)	3084237/3543980	87.03% (86.99-87.06)
Jan-Feb 2022	36704/38076	96.40% (96.20-96.58)	3851883/4273918	90.13% (90.10-90.15)
Mar-Apr 2022	41431/42270	98.02% (97.88-98.14)	3054173/3238009	94.32% (94.30-94.35)
May-Jun 2022	11759/11988	98.09% (97.83-98.32)	680972/723862	94.07% (94.02-94.13)
Jul-Aug 2022	14861/15166	97.99% (97.75-98.20)	825990/873027	94.61% (94.56-94.66)

Supplementary Table 29: Temporal comparison of SARS-CoV-2 vaccination uptake within patients within cancer and the non-cancer population. Vaccination uptake was calculated as the proportion of SARS-CoV-2 tests taken by individuals who had received at least one SARS-CoV-2 vaccination dose at each temporal timepoint.

Clinical Outcome	Univariable		Multivariable adjusted	
	OR (95% CI)	p value	OR (95% CI)	p value
Hospital assessment	5.34 (5.23-5.45)	<0.00001	3.02 (2.95-3.08)	<0.00001
Inpatient hospitalisation	4.22 (4.11-4.33)	<0.00001	2.10 (2.04-2.16)	<0.00001
Intensive care admission	4.22 (3.76-4.74)	<0.00001	2.53 (2.24-2.86)	<0.00001
COVID-19 mortality	7.12 (6.85-7.40)	<0.00001	2.54 (2.44-2.65)	<0.00001

Supplementary Table 30: Relative risk of patients with cancer to severe clinical outcomes compared to the non-cancer population following a COVID-19 infection. Multivariable logistic regression models were adjusted for age, sex, ethnicity, and deprivation.

Hospital Assessment

Time Period	<i>Haematological malignancy</i>		<i>Solid tumour malignancy</i>	
	Outcome/Total (n/N)	Outcome % (95% CI)	Outcome/Total (n/N)	Outcome % (95% CI)
Nov-Dec 2020	914/2472	36.97% (35.09-38.90)	4818/16737	28.79% (28.11-29.48)
Jan-Feb 2021	1303/2893	45.04% (43.23-46.86)	5886/17335	33.95% (33.25-34.66)
Mar-Apr 2021	269/643	41.84% (38.08-45.69)	615/1831	33.59% (31.46-35.78)
May-Jun 2021	158/416	37.98% (33.45-42.73)	300/1277	23.49% (21.25-25.90)
Jul-Aug 2021	640/1573	40.69% (38.28-43.13)	1780/8912	19.97% (19.16-20.82)
Sep-Oct 2021	758/1950	38.87% (36.73-41.06)	1975/11688	16.90% (16.23-17.59)
Nov-Dec 2021	985/3186	30.92% (29.34-32.54)	2169/20406	10.63% (10.21-11.06)
Jan-Feb 2022	1621/5378	30.14% (28.93-31.38)	4110/32698	12.57% (12.21-12.93)
Mar-Apr 2022	1406/6096	23.06% (22.02-24.14)	4200/36174	11.61% (11.28-11.94)
May-Jun 2022	696/2372	29.34% (27.54-31.21)	1705/9616	17.73% (16.98-18.51)
Jul-Aug 2022	702/2803	25.04% (23.48-26.68)	2023/12363	16.36% (15.72-17.03)

Inpatient Hospitalisation

Time Period	<i>Haematological malignancy</i>		<i>Solid tumour malignancy</i>	
	Outcome/Total (n/N)	Outcome % (95% CI)	Outcome/Total (n/N)	Outcome % (95% CI)
Nov-Dec 2020	792/2472	32.04% (30.23-33.90)	4244/16737	25.36% (24.70-26.02)
Jan-Feb 2021	1065/2893	36.81% (35.07-38.59)	5120/17335	29.54% (28.86-30.22)
Mar-Apr 2021	183/643	28.46% (25.11-32.07)	476/1831	26.00% (24.04-28.05)
May-Jun 2021	111/416	26.68% (22.66-31.13)	226/1277	17.70% (15.70-19.89)
Jul-Aug 2021	540/1573	34.33% (32.02-36.71)	1461/8912	16.39% (15.64-17.18)
Sep-Oct 2021	630/1950	32.31% (30.27-34.42)	1563/11688	13.37% (12.77-14.00)
Nov-Dec 2021	760/3186	23.85% (22.41-25.37)	1507/20406	7.39% (7.03-7.75)
Jan-Feb 2022	881/5378	16.38% (15.42-17.39)	2399/32698	7.34% (7.06-7.62)
Mar-Apr 2022	705/6096	11.56% (10.79-12.39)	2444/36174	6.76% (6.50-7.02)
May-Jun 2022	370/2372	15.60% (14.19-17.11)	1054/9616	10.96% (10.35-11.60)
Jul-Aug 2022	326/2803	11.63% (10.50-12.87)	1204/12363	9.74% (9.23-10.27)

Intensive Care Admission

Time Period	<i>Haematological malignancy</i>		<i>Solid tumour malignancy</i>	
	Outcome/Total (n/N)	Outcome % (95% CI)	Outcome/Total (n/N)	Outcome % (95% CI)
Nov-Dec 2020	120/2472	4.85% (4.07-5.77)	316/16737	1.89% (1.69-2.11)
Jan-Feb 2021	142/2893	4.91% (4.18-5.76)	372/17335	2.15% (1.94-2.37)
Mar-Apr 2021	29/643	4.51% (3.16-6.40)	20/1831	1.09% (0.71-1.68)
May-Jun 2021	23/416	5.53% (3.71-8.16)	9/1277	0.70% (0.37-1.33)
Jul-Aug 2021	104/1573	6.61% (5.49-7.95)	131/8912	1.47% (1.24-1.74)
Sep-Oct 2021	124/1950	6.36% (5.36-7.53)	154/11688	1.32% (1.13-1.54)
Nov-Dec 2021	120/3186	3.77% (3.16-4.49)	92/20406	0.45% (0.37-0.55)
Jan-Feb 2022	98/5378	1.82% (1.50-2.22)	81/32698	0.25% (0.20-0.31)
Mar-Apr 2022	33/6096	0.54% (0.39-0.76)	98/36174	0.27% (0.22-0.33)
May-Jun 2022	25/2372	1.05% (0.71-1.55)	35/9616	0.36% (0.26-0.51)
Jul-Aug 2022	10/2803	0.36% (0.19-0.66)	32/12363	0.26% (0.18-0.37)

COVID-19 Mortality

Time Period	<i>Haematological malignancy</i>		<i>Solid tumour malignancy</i>	
	Outcome/Total (n/N)	Outcome % (95% CI)	Outcome/Total (n/N)	Outcome % (95% CI)
Nov-Dec 2020	642/2472	25.97% (24.28-27.74)	3319/16737	19.83% (19.23-20.44)
Jan-Feb 2021	704/2893	24.33% (22.81-25.93)	3448/17335	19.89% (19.30-20.49)
Mar-Apr 2021	114/643	17.73% (14.97-20.87)	258/1831	14.09% (12.57-15.76)
May-Jun 2021	87/416	20.91% (17.28-25.08)	75/1277	5.87% (4.71-7.30)
Jul-Aug 2021	233/1573	14.81% (13.14-16.65)	499/8912	5.60% (5.14-6.10)
Sep-Oct 2021	246/1950	12.62% (11.21-14.16)	661/11688	5.66% (5.25-6.09)
Nov-Dec 2021	274/3186	8.60% (7.68-9.62)	617/20406	3.02% (2.80-3.27)
Jan-Feb 2022	371/5378	6.90% (6.25-7.61)	1197/32698	3.66% (3.46-3.87)
Mar-Apr 2022	286/6096	4.69% (4.19-5.25)	1088/36174	3.01% (2.84-3.19)
May-Jun 2022	154/2372	6.49% (5.57-7.56)	380/9616	3.95% (3.58-4.36)
Jul-Aug 2022	134/2803	4.78% (4.05-5.63)	491/12363	3.97% (3.64-4.33)

Supplementary Table 31: Temporal analysis of COVID-19 case-outcome rates in individuals with a haematological malignancy or solid tumour malignancy. Case-outcome rates (outcome %) were calculated in 2-month intervals from 1st November 2020-31st August 2022. 95% confidence intervals for case-outcome rates are provided in brackets.

Hospital Assessment

Time Period	Systemic anti-cancer therapy (SACT)		No systemic anti-cancer therapy	
	Outcome/Total (n/N)	Outcome % (95% CI)	Outcome/Total (n/N)	Outcome % (95% CI)
Nov-Dec 2020	1818/4680	38.85% (37.46-40.25)	3914/14529	26.94% (26.22-27.67)
Jan-Feb 2021	2265/5205	43.52% (42.17-44.87)	4924/15023	32.78% (32.03-33.53)
Mar-Apr 2021	342/802	42.64% (39.26-46.09)	542/1672	32.42% (30.22-34.70)
May-Jun 2021	190/545	34.86% (30.98-38.96)	268/1148	23.34% (20.99-25.88)
Jul-Aug 2021	772/2452	31.48% (29.68-33.35)	1648/8033	20.52% (19.65-21.41)
Sep-Oct 2021	886/2933	30.21% (28.57-31.89)	1847/10705	17.25% (16.55-17.98)
Nov-Dec 2021	1161/4818	24.10% (22.91-25.33)	1993/18774	10.62% (10.18-11.06)
Jan-Feb 2022	2085/8068	25.84% (24.90-26.81)	3646/30008	12.15% (11.79-12.52)
Mar-Apr 2022	1897/8365	22.68% (21.79-23.59)	3709/33905	10.94% (10.61-11.28)
May-Jun 2022	849/2881	29.47% (27.83-31.16)	1552/9107	17.04% (16.28-17.83)
Jul-Aug 2022	872/3505	24.88% (23.48-26.34)	1853/11661	15.89% (15.24-16.57)

Inpatient Hospitalisation

Time Period	Systemic anti-cancer therapy (SACT)		No systemic anti-cancer therapy	
	Outcome/Total (n/N)	Outcome % (95% CI)	Outcome/Total (n/N)	Outcome % (95% CI)
Nov-Dec 2020	1428/4680	30.51% (29.21-31.85)	3608/14529	24.83% (24.14-25.54)
Jan-Feb 2021	1708/5205	32.81% (31.55-34.10)	4477/15023	29.80% (29.07-30.54)
Mar-Apr 2021	234/802	29.18% (26.14-32.42)	425/1672	25.42% (23.39-27.56)
May-Jun 2021	121/545	22.20% (18.91-25.88)	216/1148	18.82% (16.66-21.18)
Jul-Aug 2021	608/2452	24.80% (23.13-26.54)	1393/8033	17.34% (16.53-18.18)
Sep-Oct 2021	629/2933	21.45% (20.00-22.97)	1564/10705	14.61% (13.95-15.29)
Nov-Dec 2021	738/4818	15.32% (14.33-16.36)	1529/18774	8.14% (7.76-8.54)
Jan-Feb 2022	1038/8068	12.87% (12.15-13.61)	2242/30008	7.47% (7.18-7.77)
Mar-Apr 2022	841/8365	10.05% (9.43-10.72)	2308/33905	6.81% (6.54-7.08)
May-Jun 2022	429/2881	14.89% (13.64-16.24)	995/9107	10.93% (10.30-11.58)
Jul-Aug 2022	404/3505	11.53% (10.51-12.63)	1126/11661	9.66% (9.13-10.21)

Intensive Care Admission

Time Period	Systemic anti-cancer therapy (SACT)		No systemic anti-cancer therapy	
	Outcome/Total (n/N)	Outcome % (95% CI)	Outcome/Total (n/N)	Outcome % (95% CI)
Nov-Dec 2020	139/4680	2.97% (2.52-3.50)	297/14529	2.04% (1.83-2.29)
Jan-Feb 2021	137/5205	2.63% (2.23-3.10)	377/15023	2.51% (2.27-2.77)
Mar-Apr 2021	19/802	2.37% (1.52-3.67)	30/1672	1.79% (1.26-2.55)
May-Jun 2021	16/545	2.94% (1.82-4.72)	16/1148	1.39% (0.86-2.25)
Jul-Aug 2021	83/2452	3.38% (2.74-4.18)	152/8033	1.89% (1.62-2.21)
Sep-Oct 2021	99/2933	3.38% (2.78-4.09)	179/10705	1.67% (1.45-1.93)
Nov-Dec 2021	75/4818	1.56% (1.24-1.95)	137/18774	0.73% (0.62-0.86)
Jan-Feb 2022	78/8068	0.97% (0.78-1.20)	101/30008	0.34% (0.28-0.41)
Mar-Apr 2022	30/8365	0.36% (0.25-0.51)	101/33905	0.30% (0.25-0.36)
May-Jun 2022	18/2881	0.62% (0.40-0.99)	42/9107	0.46% (0.34-0.62)
Jul-Aug 2022	8/3505	0.23% (0.12-0.45)	34/11661	0.29% (0.21-0.41)

COVID-19 Mortality

Time Period	Systemic anti-cancer therapy (SACT)		No systemic anti-cancer therapy	
	Outcome/Total (n/N)	Outcome % (95% CI)	Outcome/Total (n/N)	Outcome % (95% CI)
Nov-Dec 2020	1049/4680	22.41% (21.24-23.63)	2912/14529	20.04% (19.40-20.70)
Jan-Feb 2021	1189/5205	22.84% (21.72-24.00)	2963/15023	19.72% (19.09-20.37)
Mar-Apr 2021	135/802	16.83% (14.40-19.58)	237/1672	14.17% (12.58-15.93)
May-Jun 2021	75/545	13.76% (11.12-16.91)	87/1148	7.58% (6.18-9.25)
Jul-Aug 2021	278/2452	11.34% (10.14-12.65)	454/8033	5.65% (5.17-6.18)
Sep-Oct 2021	279/2933	9.51% (8.50-10.63)	628/10705	5.87% (5.44-6.33)
Nov-Dec 2021	317/4818	6.58% (5.91-7.31)	574/18774	3.06% (2.82-3.31)
Jan-Feb 2022	527/8068	6.53% (6.01-7.09)	1041/30008	3.47% (3.27-3.68)
Mar-Apr 2022	407/8365	4.87% (4.42-5.35)	967/33905	2.85% (2.68-3.03)
May-Jun 2022	202/2881	7.01% (6.14-8.00)	332/9107	3.65% (3.28-4.05)
Jul-Aug 2022	194/3505	5.53% (4.83-6.34)	431/11661	3.70% (3.37-4.05)

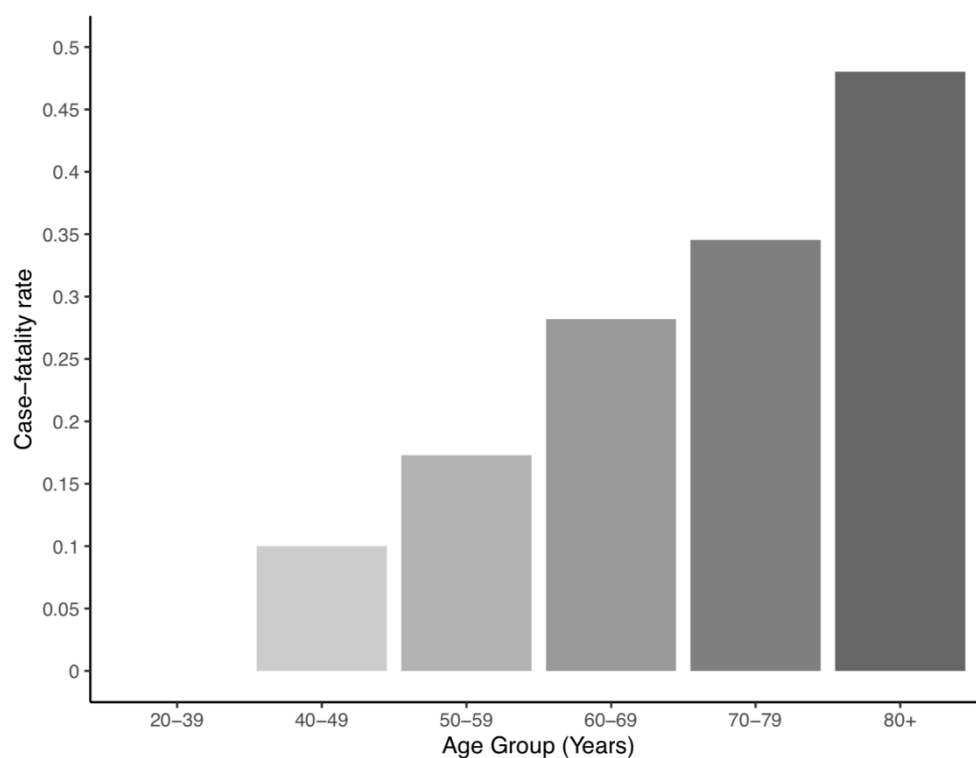
Supplementary Table 32: Temporal analysis of COVID-19 case-outcome rates of individuals receiving systemic anti-cancer therapy (SACT) and not receiving SACT. Case-outcome rates (outcome %) were calculated in 2-month intervals from 1st November 2020-31st August 2022. 95% confidence intervals for case-outcome rates are provided in brackets.

Systemic anti-cancer therapy (SACT)	Hospital assessment		Inpatient hospitalisation	
	Odds Ratio (95% CI)	p value	Odds Ratio (95% CI)	p value
Any SACT	2.29 (2.21-2.37)	<0.0001	1.58 (1.50-1.65)	<0.0001
SACT class				
SACT Class A	2.55 (2.40-2.70)	<0.0001	1.67 (1.54-1.80)	<0.0001
SACT Class B	2.46 (2.35-2.58)	<0.0001	1.68 (1.58-1.79)	<0.0001
SACT Class C	2.73 (2.20-3.38)	<0.0001	1.65 (1.22-2.21)	0.0013
SACT type				
Type = B/T Cell Antibody	2.41 (2.00-2.89)	<0.0001	1.77 (1.39-2.24)	<0.0001
Type = Cytotoxic	2.39 (2.29-2.49)	<0.0001	1.59 (1.51-1.68)	<0.0001
Type = Hormonal	3.00 (2.49-3.60)	<0.0001	2.00 (1.56-2.53)	<0.0001
Type = Immunotherapy	1.95 (1.45-2.60)	<0.0001	2.06 (1.45-2.88)	<0.0001
Type = Targeted	3.95 (3.63-4.31)	<0.0001	2.26 (2.01-2.52)	<0.0001

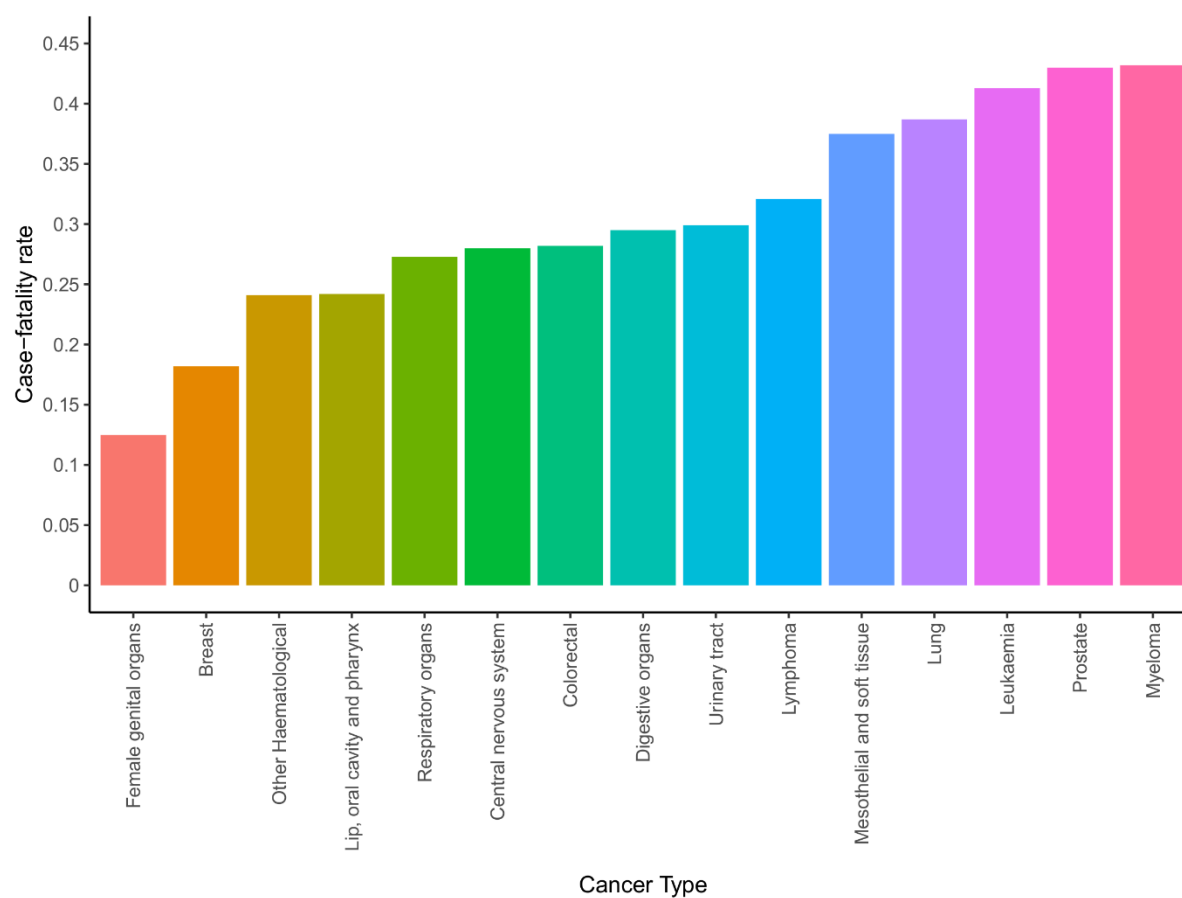
Systemic anti-cancer therapy (SACT)	Intensive care admission		Mortality	
	Odds Ratio (95% CI)	p value	Odds Ratio (95% CI)	p value
Any SACT	1.79 (1.45-2.21)	<0.0001	1.83 (1.71-1.96)	<0.0001
SACT class				
SACT Class A	1.27 (0.82-1.89)	0.2207	1.82 (1.62-2.03)	<0.0001
SACT Class B	1.81 (1.36-2.39)	<0.0001	1.89 (1.72-2.06)	<0.0001
SACT Class C	5.67 (2.41-11.44)	0.0001	1.87 (1.21-2.79)	0.0041
SACT type				
Type = B/T Cell Antibody	4.89 (2.31-9.19)	<0.0001	1.84 (1.28-2.57)	0.0010
Type = Cytotoxic	1.94 (1.53-2.44)	<0.0001	1.84 (1.70-1.98)	<0.0001
Type = Hormonal	0.53 (0.01-3.01)	1.0000	2.60 (1.88-3.53)	<0.0001
Type = Immunotherapy	0.00 (0.00-4.05)	1.0000	1.56 (0.84-2.66)	0.1258
Type = Targeted	2.28 (1.33-3.67)	0.0020	2.19 (1.85-2.58)	<0.0001

Supplementary Table 33: Relative risk of COVID-19 hospital assessment, inpatient hospitalisation, intensive care admission and mortality by systemic anti-cancer therapy (SACT) regimens in 2022 compared individuals not receiving SACT. Univariable Fisher's Exact testing was used to assess relative risk. Odds Ratios are provided for each SACT type or class with 95% confidence intervals provided in brackets.

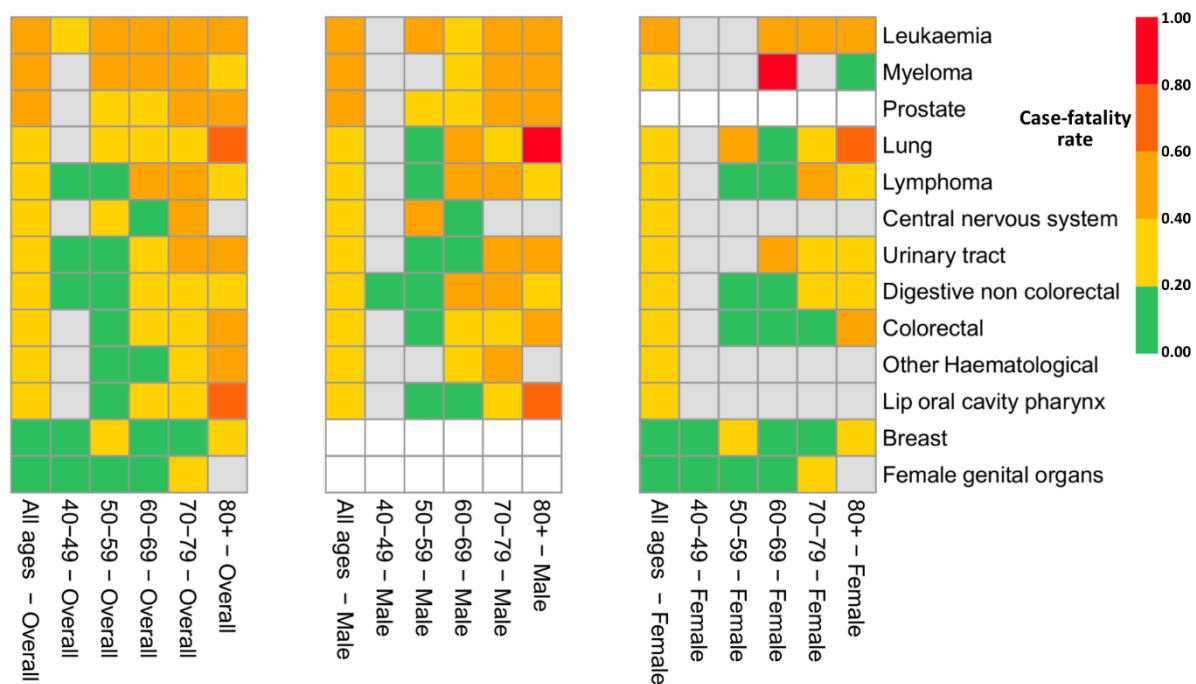
Supplementary figures



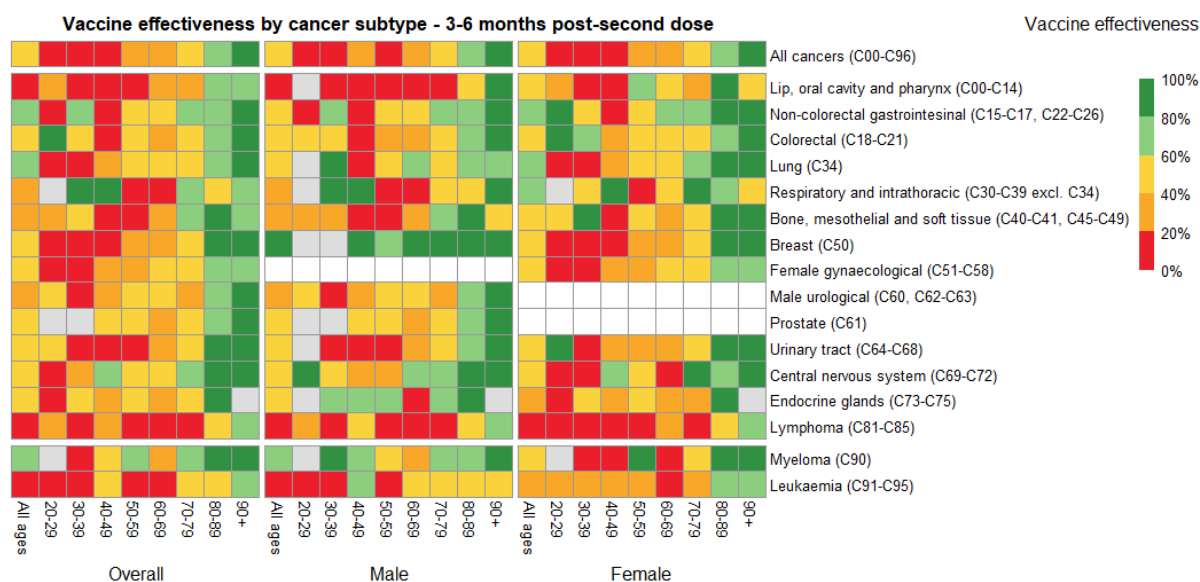
Supplementary Figure 1: Age and risk of all-cause case fatality rate of patients following a presenting with COVID-19 in the UKCCMP cohort. Case-fatality rate is calculated at the number of all-cause deaths over the total number of patients with cancer following COVID-19 within each age group.



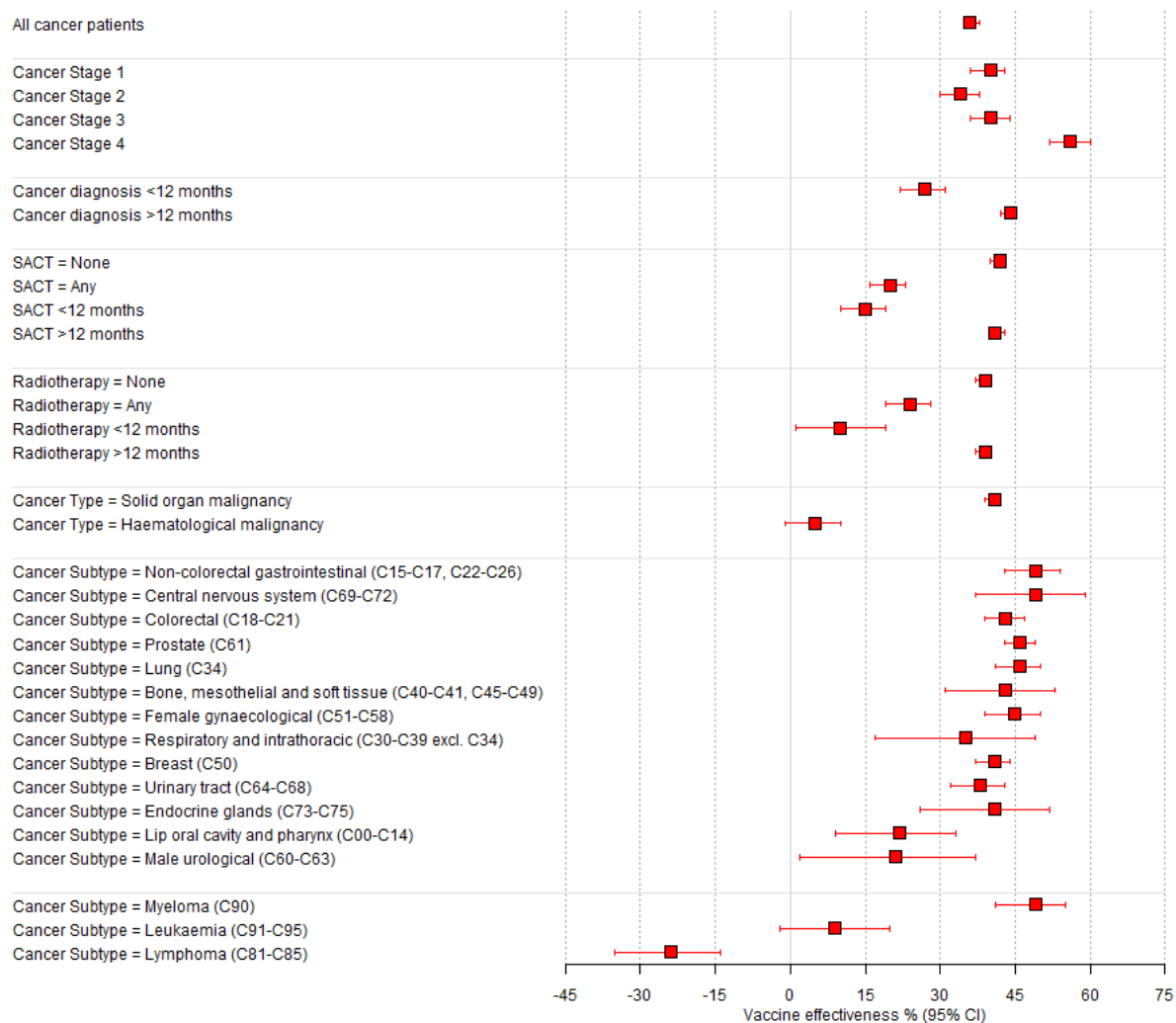
Supplementary Figure 2: Case fatality rate of patients following a presenting with COVID-19 in the UKCCMP cohort, assessed by tumour subtype. Case-fatality rate is calculated at the number of all-cause deaths over the total number of patients with cancer following COVID-19 within each primary tumour subtype.



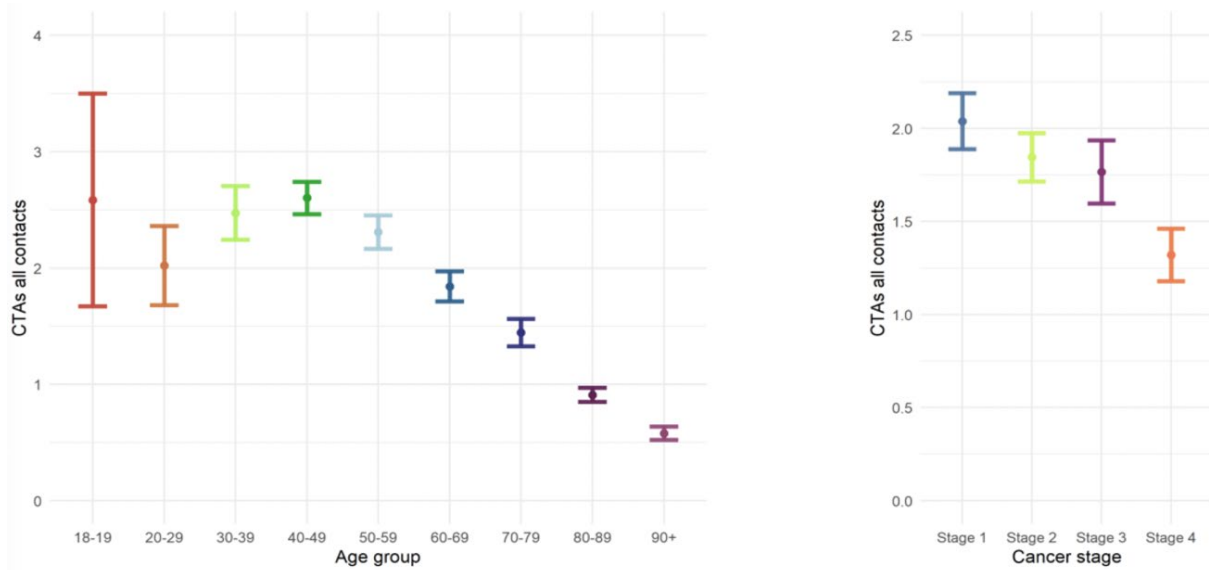
Supplementary Figure 3: Heatmap demonstrating case fatality rate following a COVID-19 presentation, broken down by tumour subtype, age and gender. Case-fatality rate is calculated at the number of all-cause deaths over the total number of patients with cancer following COVID-19 within each group. Grey bars represent where number of cases were less than 4.



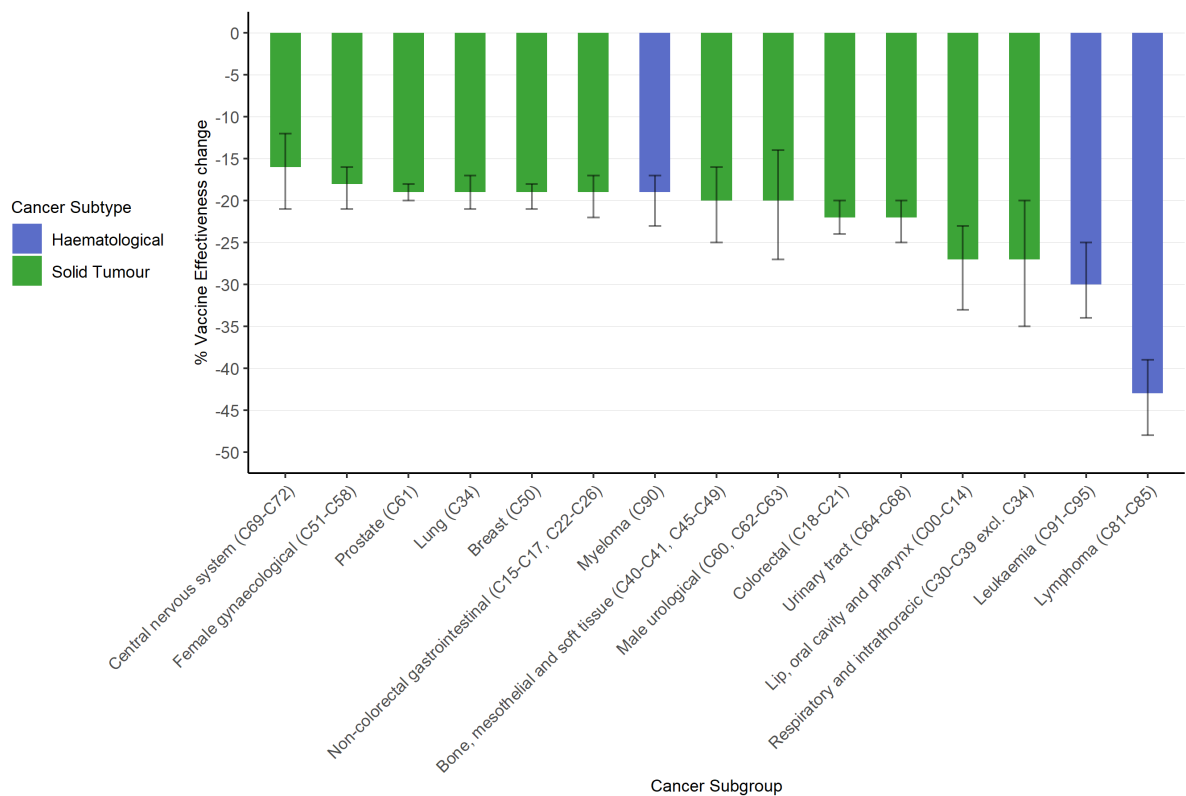
Supplementary Figure 4: Heatmap demonstrating vaccine effectiveness at 3-6m post 2nd dose vaccination and association to patient demographics (age and sex) and tumour subtype. Grey boxes denote insufficient data, white boxes denote inapplicable sections.



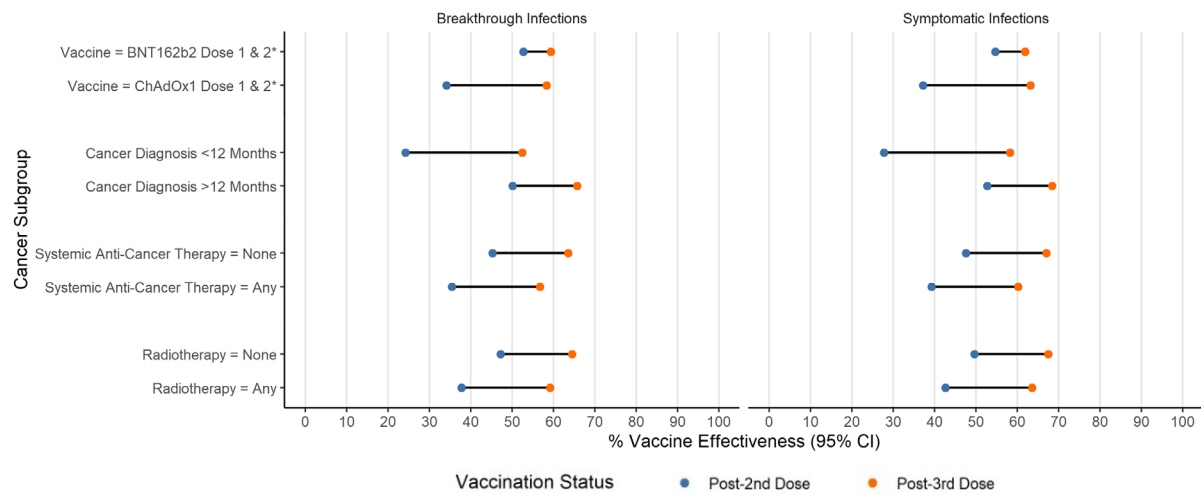
Supplementary Figure 5: Plot demonstrating multivariable corrected vaccine effectiveness at 3-6m amongst predefined cancer subgroups. Regression models were fitted for clinically relevant covariates of age, sex, deprivation index and ethnicity.



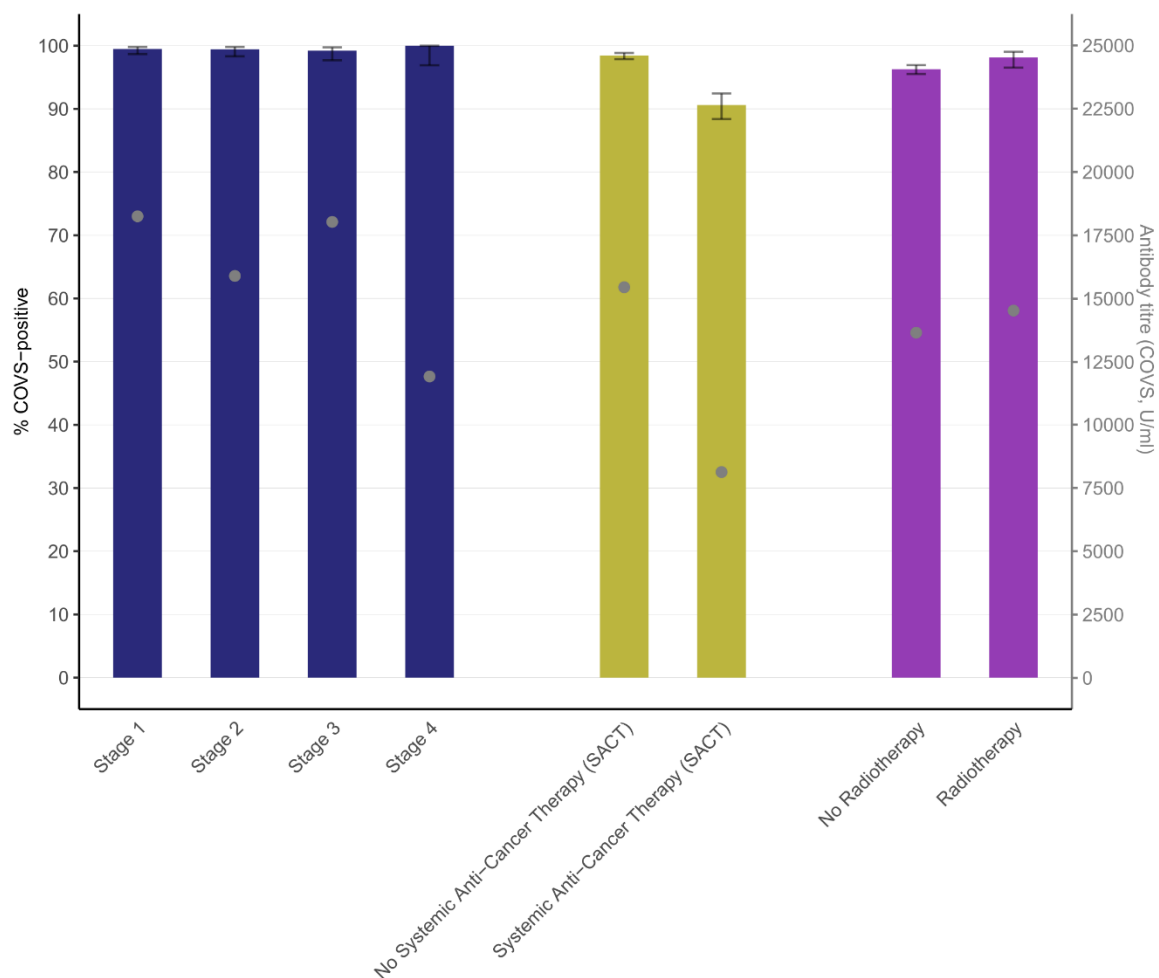
Supplementary Figure 6: Vertical dotplot showing the mean number of contacts in cancer patients with COVID-19 PCR tests by age group (left panels) and cancer stage (right panels). Bars denote 95% confidence interval bars.



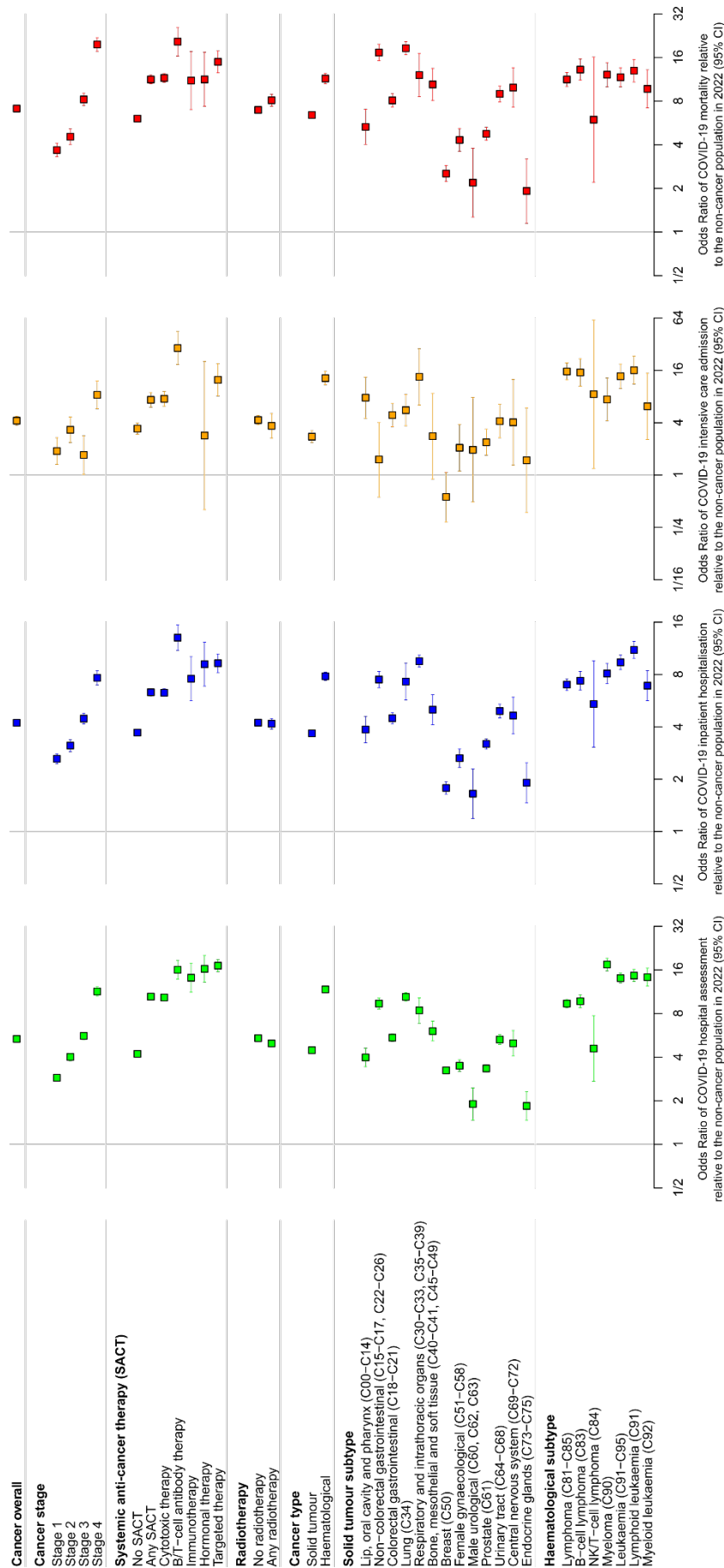
Supplementary Figure 7: Waterfall plot showing multivariable corrected vaccine effectiveness change in different cancer subgroups. Vaccine effectiveness was corrected for clinically relevant covariates (patient age, sex, ethnicity, and index of deprivation).



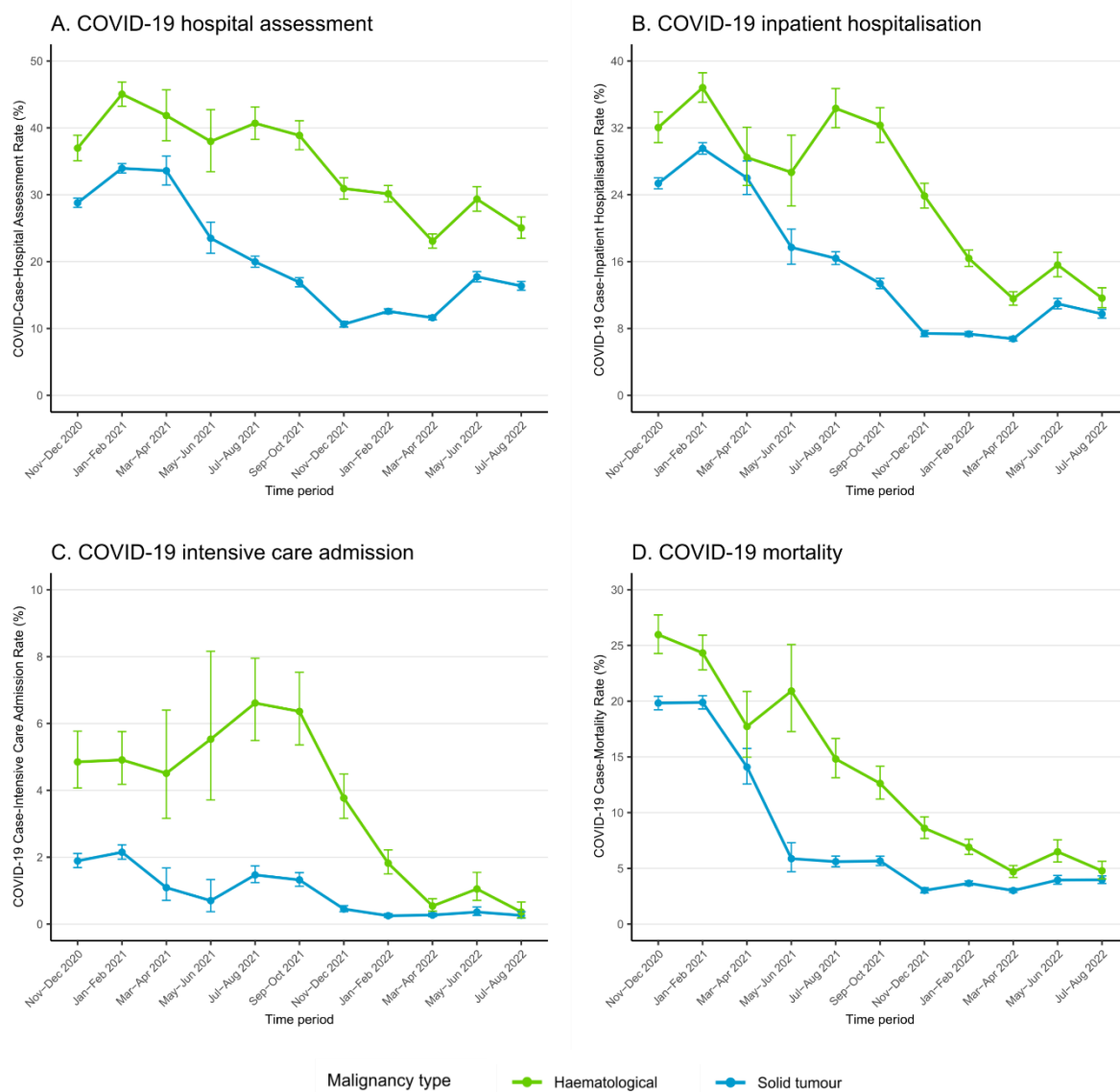
Supplementary Figure 8: Subgroup analyses of vaccine effectiveness in the cancer cohort post third dose booster and post second dose and effect of primary vaccination type, cancer diagnosis date and cancer treatments.



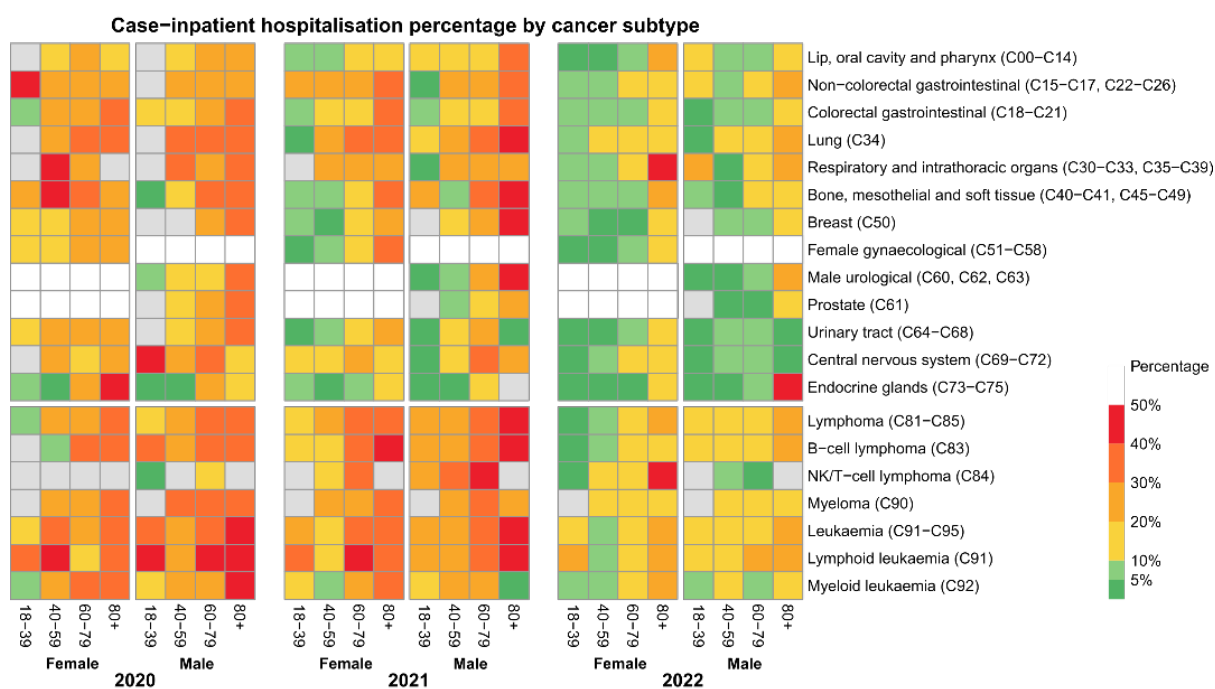
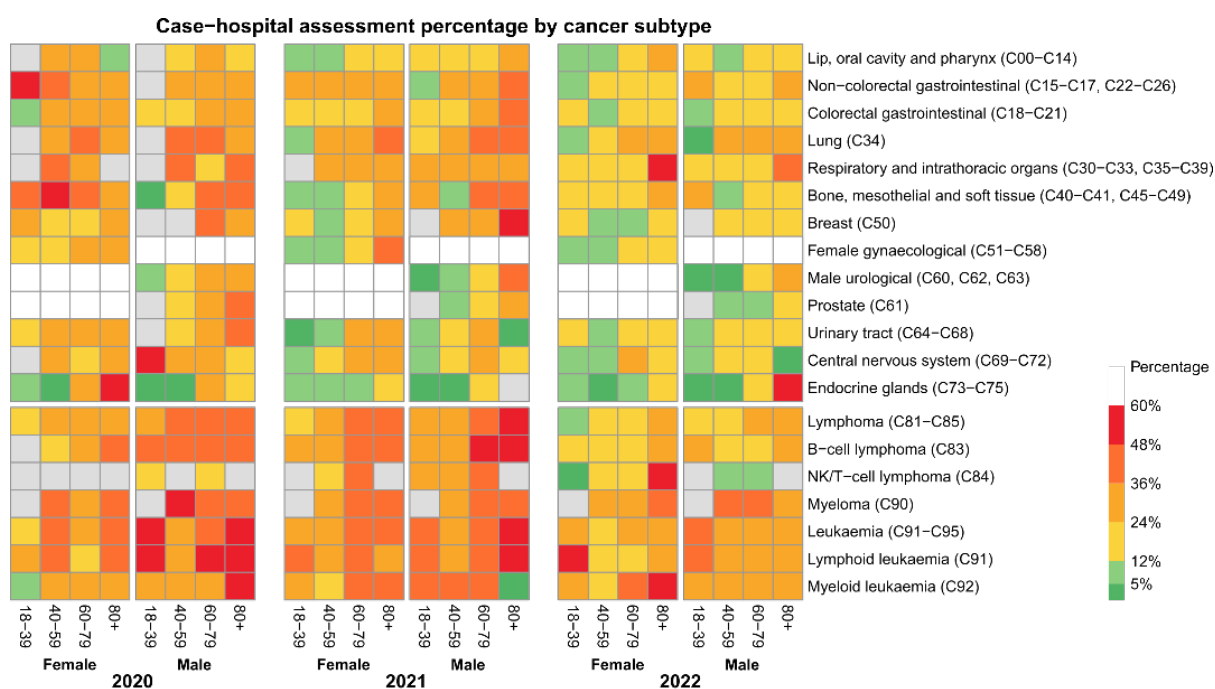
Supplementary Figure 9: Median coronavirus antibody titres and responses in cancer patients based on cancer and treatment. Bars represent % COVS positive results, dots represent antibody response titres (COVS/U/ml).

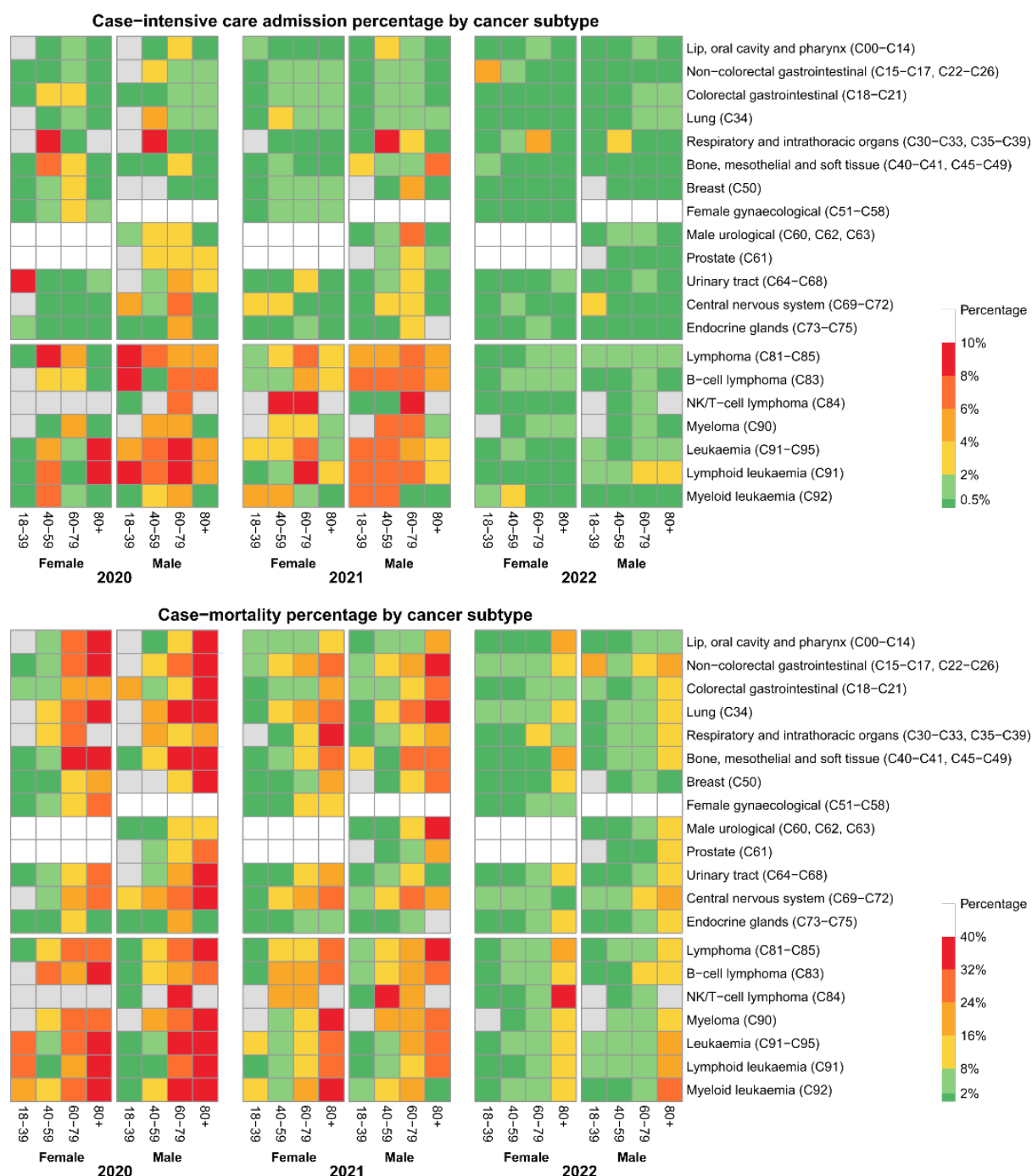


Supplementary Figure 10: Relative risk of subgroups of patients with cancer to severe clinical outcomes compared to the non-cancer population following a COVID-19 infection between January 2022 and August 2022. Univariable logistic regression models adjusting for age, sex, ethnicity and deprivation show the relative risk of COVID-19 hospital assessment (green), inpatient hospitalisation (blue), intensive care admission (orange) and mortality (red). Relative risk to the non-cancer population is shown as Odds Ratios (OR) with 95% confidence intervals. Corresponding ICD-10 codes for specific primary tumour subtypes are listed in brackets.

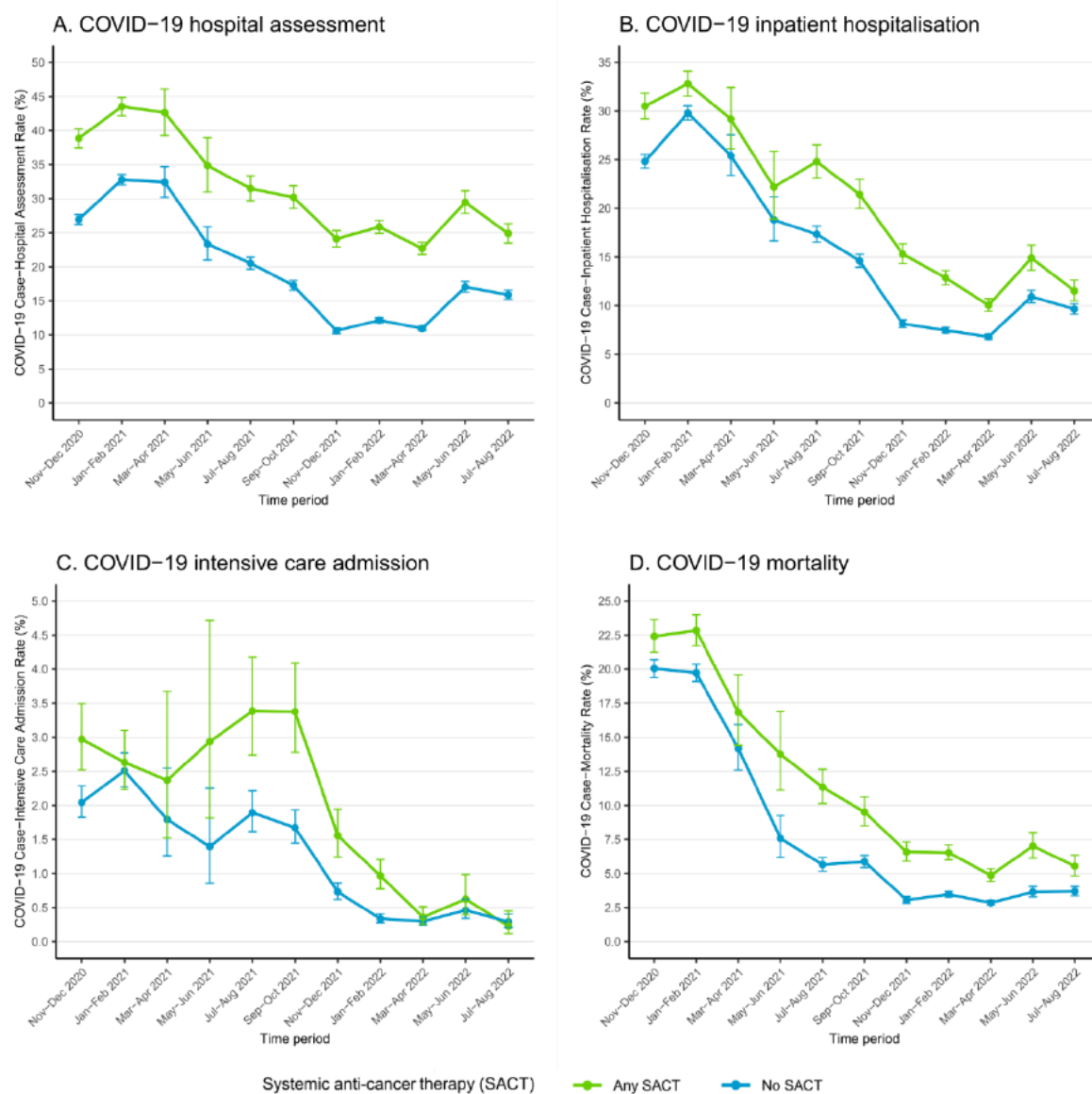


Supplementary Figure 11: Temporal trends of case-outcome rates between patients with haematological and solid tumour malignancies in England between November 2020 and August 2022. Case-outcome rates (%) with 95% confidence intervals are shown at 2-month intervals for COVID-19 hospital assessment (A), inpatient hospitalisation (B), intensive care admission (C) and mortality (D).





Supplementary Figure 12: Case-outcome rates of solid and haematological cancer subtypes by age group and sex (female or male) following a COVID-19 infection in 2020, 2021 and 2022. Case-outcome rates for COVID-19 hospital assessment, inpatient hospitalisation, intensive care admission and mortality are shown for each cancer subtype and are colour coded according to the scales below each heatmap. Grey boxes denote <10 SARS-CoV-2 tests within a subgroup. White boxes denote <10 SARS-CoV-2 tests within sex-specific cancer subtypes (female gynaecological, male urological, prostate cancers). Corresponding ICD-10 codes for specific primary tumour subtypes are listed in brackets.



Supplementary Figure 13: Temporal trends of case-outcome rates between patients with cancer in receipt of systemic anti-cancer therapy (SACT) and not receiving SACT in England between November 2020 and August 2022. Case-outcome rates (%) with 95% confidence intervals are shown at 2-month intervals for COVID-19 hospital assessment (A), inpatient hospitalisation (B), intensive care admission (C) and mortality (D).