Strategies and Advancements for Personalized Medicine in Gastric Cancer

by

Daniel Todd Skubleny

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Surgery
University of Alberta

Abstract

Five-year overall survival for gastric cancer in Canada is only 25%. Standard of care treatment using neoadjuvant chemotherapy and surgery achieves partial or complete treatment response in only 40% of patients. Additional strategies and methods to discover novel therapies and optimize existing treatment regimens are required. In this work we use a combination of immunohistochemistry-based biomarkers, 3-dimensional organoid models and omics based molecular classification to investigate and enhance personalized medicine strategies in gastric cancer.

Forty-three patients with gastric adenocarcinoma, of which 18 underwent neoadjuvant chemotherapy, were included in a prospective clinical cohort. Differences in expression of Galectin-3, E-cadherin, $CD4^+$ and $CD8^+$ molecules between tumours with and without treatment response to neoadjuvant chemotherapy were assessed with immunohistochemistry. To enhance procurement of fresh tissue for organoid culture we assessed the feasibility of shipping mouse stomach on ice for 24- or 48-hours using Hank's Balanced Salts Solution (HBSS), Histidine-tryptophan-ketoglutarate (HTK) or University of Wisconsin solutions as transport media. The effect of transport time and transport media on organoid viability, growth rate and stem cell gene expression of LGR5 and TROY were assessed using cell counting with Trypan Blue and quantitative real-time PCR, respectively. Multivariable generalized additive models (GAM) were used to assess these outcomes over 12 organoid passages. Using publicly available whole-transcriptome data we developed supervised machine learning models to assign molecular subtypes from The Cancer Genome Atlas (TCGA), Asian Cancer Research Group (ACRG) and Tumour Microenvironment score (TME) classification systems to 2,202 patients. Overall survival was assessed using a multivariable Cox proportional hazards model. Using genes informed by these models we developed a custom Nanostring codeset to assign molecular subtypes to our 43-patient cohort and 10 tumour and tumour-organoid pairs. The accuracy of molecular subtype

models and our Nanostring test were assessed relative to gold-standard Epstein-Barr encoded early RNAs *in-situ* hybridization and pentaplex polymerase chain reaction for Epstein-Barr Virus (EBV) and Microsatellite instability (MSI) tumours, respectively.

The ratio between CD4 and CD8 lymphocytes was significantly greater in treatment responsive tumours (Wilcoxon, $p=0.03$). In univariate models CD4/CD8 ratio was the only biomarker that provided significant predictive value (Accuracy 86%, $p<0.001$). Mock-shipment of mouse stomach tissue for 24- or 48-hours significantly decreased growth rate relative to freshly prepared organoids but did not affect viability (GAM, $p<0.001$). Transport media did not affect growth rate or viability. Gene expression of LGR5 and TROY was not affected by transport time or media but significantly decreased upon tissue dissociation and subsequently increased in successive passages to regain endogenous expression levels by passage 6 (Kruskal-Wallis, $p<0.001$ with post-hoc Dunn's Test). Two human gastric cancer organoids were developed following 24- or 48-hour transport in HBSS solution. Classification models for TCGA (57 genes), TME (50 genes) and ACRG (39 genes) had a mean accuracy ± standard deviation of 89.5% ± 0.04, 89.4% ± 0.01 and 84.66% ± 0.04, respectively. Improved prognosis was observed for TME High tumours (Hazard Ratio 0.61, [95% Confidence Interval 0.46, 0.79]) and the TME score was the only statistically significant prognostic classification system in a multivariable Cox survival model (Global Wald Test, $p<0.001$). In a public cohort of 2,202 patients our models demonstrated a 98.7% and 99.3% accuracy for EBV and MSI subtypes with respect to gold-standard tests. In 10 patient and patient-derived organoid samples the Nanostring test was 100% accurate for EBV and MSI subtypes.

In this study, a multimodal approach is applied to investigate personalized medicine strategies for gastric cancer. These results demonstrate that $CD4^+/CD8^+$ Ratio is a promising IHC-based biomarker with therapeutic implications for response to neoadjuvant chemotherapy in locally advanced gastric cancer. From a translational perspective we found

that cold shipment of fresh gastric tissue for 24- or 48-hours from mouse and human is a feasible and reliable method to increase procurement of primary organoid tissue. To our knowledge, at the time of publication, this study is the largest integrated analysis of TCGA, ACRG and TME molecular classification systems in gastric cancer. Using a custom Nanostring codeset we successfully translate these machine learning models to our own population and organoid samples. Together, these findings form a foundation to enhance future investigation of personalized medicine in gastric cancer.

Preface

This thesis is an original work by Daniel Skubleny. All human clinical participants consented according to the approved ethics protocol granted by the Health Research Ethics Board of Alberta, under the project title "Expression of Biomarkers in Gastric Cancer and the Effect of Neoadjuvant Chemotherapy" and study identification HREBA.CC-17-0228.

Ethics approval of mouse experiments conducted in Chapter 4 was granted by the University of Alberta Research Ethics Office under Category A Exemptions for review by the Animal Care Use Committee on the principle of Reduction. Ethics was granted under the project title "Optimization of fresh tissue preservation for three-dimensional cell-culture organoids."

In Chapters 4 and 5, we used a custom gene panel manufactured by Nanostring®. All genes selected for the code set were informed from models developed by myself. Gene codeset chips were manufactured and supplied by Nanostring®.

Whole-transcriptome data in this thesis is publicly accessible on Gene Expression Omnibus and available for Unrestricted-Use as stated by the National Center for Biotechnology Information and the National Institutes for Health Genomic Data Sharing Policy. Whole-transcriptome data accessible from cBioPortal is available under an open source license via Github.

Publicly available cancer cell line data genetic data from the Cancer Cell Line Encyclopedia is publicly accessible according to the "Broad Institute's commitment to Open Science," in which the "data made available were generated for research purposes and are not intended for clinical use."

Cancer cell line dose-response data from the Genomics of Drug Sensitivity in Cancer project hosted by the Wellcome Sanger Institute are available for non-exclusive, non-transferable right to use for internal proprietary research and educational purposes including target, biomarker and drug discovery.

Table of Contents

1.1 List of Tables

1.2 List of Figures

List of Abbreviations

Abbreviation	Term
ACRG	Asian Cancer Research Group
AICc	Corrected Akaike's information criterion
AUC	Area under the curve
C-Index	Concordance Index
cDNA	Complementary DNA
CI	Confidence interval
CIN	Chromosomal Instability
conf-ECE	Confidence estimated calibration error
conf-MCE	Confidence maximum calibration error
Cox-KP	Cox proportional hazards with Kalbfleisch-Prentice extensions
CoxEN-KP	Cox proportional hazards ElasticNet with Kalbfleisch-Prentice extensions
CV	Cross validation
cw-ECE	Class wise estimated calibration error
cw-MCE	Class wise maximum calibration error
DMSO	Dimethyl sulfoxide
DSS	Drug sensitivity score
EBER	Epstein-Barr encoded early RNAs
EBV	Epstein-Barr Virus
EMT	Epithelial to Mesenchymal Transition
FFPE	Formalin-Fixed Paraffin-Embedded
FLOT	5-Fluorouracil, Leucovorin, Oxaliplatin, Docetaxel
FOLFOX	Folinic-Acid (Leucovorin), 5-Fluorouracil, Oxaliplatin
FSQN	Feature specific quantile normalization
GAM	Generalized Additive Model
GBM	Gradient Boosting Machine
GS	Genomically stable
HBSS	Hank's Balanced Salts Solution
HR	Hazard Ratio
HTK	Histidine-tryptophan-ketoglutarate
IC50	Half-maximal inhibitory concentration
IFC	Immunofluorescence

IHC	Immunohistochemistry
Int-varHet	Integrated variance heterogeneity
IQR	Interquartile range
ISD	Individual survival distributions
ISH	In-situ hybridization
ITH	Intratumour heterogeneity
KNN	k-Nearest Neighbours
LASSO	Least absolute shrinkage and selection operator
mRNA	Messenger Ribonucleic Acid
MSI	Microsattelite Instability
MSS	Microsatellite stable
MSS TP53-	Microsatellite stable TP53 negative
MSS TP53+	Microsatellite stable TP53 postive
MTLR	Multi-task logistic regression
NB	Naïve Bayes
NIR	No information rate
NNET	Neural network
PBS	Phosphate buffered saline
pCR	Pathologic complete response
RF	Random Forest
ROC	Receiver operating characteristic
RSEM	RNA-Seq by Expectation-Maximization
sc-varHet	Single classifier variance heterogeneity
SKCE	Squared kernel calibration error
SVM	Support Vector Machine
TCGA	The Cancer Genome Atlas
TILs	Tumour infiltrating lymphocytes
TMA	Tissue microarray
TME	Tumour Microenvironment
TNM	Tumour, Node, Metastasis
UW	University of Wisconsin solution
WISP	Weight In-Silico Patholoy

Glossary of Terms

Term	Definition
Adjuvant	Administration of anti-cancer surgery only after surgical removal of the cancer
Aneuploidy	An abnormal number of chromosomes. Pertains to gains or losses in chromosomes or chromosome segments.
Feature	A variable or covariate in a model or data frame
Feature selection	A method by which the most relevant features or variables to predict an outcome are selected using various mathematical or statistical properties.
Microarray	A gene expression microarray that hybridizes complementary DNA to probes on a chip. Gene expression is measured by fluorescence of the chip.
Neoadjuvant	Administration of anti-cancer therapy before and after surgical removal of the cancer
Normalization	Removal of experimental and technological effects on a sample in order to realize true biological effects
Organoid	A collection of cells grown in-vitro which resembles a mini-organ. In contrast to 2-dimensional cell culture, organoids grow in 3-dimensional space within a gel matrix.
Patient-derived	Pertaining to the growth of a patient's actual tissue biopsy in the lab
RNA-seq	Next generation sequencing of all mRNA transcripts in an organism using complementary DNA.
Supervised (machine learning)	Pertaining to learning data to match a specified outcome. The learning is supervised because the model is trying to find patterns to match to a pre-specified outcome. In contrast, unsupervised learning identifies patterns in the data unrelated to any specified outcome
Transcriptome	The entire messenger RNA expressed within an organism. Includes coding and noncoding regions

Chapter 1: Introduction

1.1 Overview

Personalized medicine strives to achieve "the right treatment, for the right patient at the right time". The experiments conducted in this thesis centre around the concept of personalized medicine, which is born from a vision of providing superior medical care based off the unique characteristics of a patient and their tumour. Although the concept of personalized medicine is frequently stated in clear terms, significant questions remain on the utility and feasibility of a personalized approach. Within each null hypothesis tested in this work, is a signal pertaining to the advance or detriment of the vision of personalized medicine.

Over the previous three decades, gastric cancer diagnosis, treatment, and surgical management has significantly improved (see: Epidemiology, Presentation and Diagnosis,

Management). Coinciding with these advances has been a decline in incidence related to environmental factors. Yet gastric cancer remains a deadly disease. This is especially true in Canada, where most gastric cancer patients present with advanced disease.

Outcomes in Asia have improved due to the high incidence of gastric cancer, which provides increased societal awareness and allows affordable implementation of population-based screening programs (see: Screening). In North America, the debate regarding the cost-utility benefit of screening programs has spanned decades. No screening program currently exists in Canada, even among first-generation Asian immigrants.

Neoadjuvant chemotherapy exists as a viable treatment strategy to combat advanced disease in Canadian populations (see: Chemotherapy and chemoradiotherapy). The MAGIC trial, published in 2006, established epirubicin, cisplatin and 5-fluorouracil as an effective therapy that downstages bulky tumours to facilitate adequate surgical resection and improves survival. In 2019 a new treatment called FLOT4, which added docetaxel to oxaliplatin, 5-fluorouracil and leucovorin, was found to improve overall survival compared to MAGIC. In light of this success, the clinical trial also found that FLOT4 is relatively toxic and complete pathologic response only occurs in approximately 15% of tumours. There is no method or biomarker capable of determining which patient receives a favourable treatment response. FLOT4 is the current standard of care and our institution's neoadjuvant therapy of

choice. Thus, this therapy is a centrepiece in clinical and *in-vitro* investigations in this thesis.

In addition to improving cytotoxic chemotherapy regimens, immune checkpoint inhibitor therapy in microsatellite instability cancers emerged in 2017 (see: Targeted molecular therapy and immunotherapy). For the first time, the FDA approved a cancer treatment based on a tumour biomarker as opposed to approval on a cancer-by-cancer basis. The Keynote-059 and Keynote-062 trials demonstrated immune checkpoint inhibitor safety and efficacy in Stage IV gastric adenocarcinoma. In this population, first-line Pembrolizumab (anti-PD-L1) monotherapy or combination with chemotherapy provides non-inferior outcomes and a more favourable side-effect profile compared to chemotherapy alone. Investigation into the utility of Pembrolizumab (anti-PD-L1) in the neoadjuvant setting is ongoing with the Keynote-585 and NCT-3257163 trials.[1,2]

Despite these advancements, there remains a significant opportunity to enhance outcomes in gastric cancer patients. Some themes outside of the direct scope of his thesis include the need for improved methods of early detection and treatment strategies for dealing with peritoneal disease or cancer with a high risk of peritoneal disease.[3–6] This thesis does, however, contribute insight into (1) biological and therapeutic considerations of microsatellite instability (MSI) and other molecular subtypes in gastric cancer, (2) investigation of potential tumour biomarkers to predict neoadjuvant treatment response and (3) strategies to investigate and communicate personalized medicine in both the wet lab and the clinic (see: Methods to Advance Personalized Medicine in Gastric Cancer).

Recently, a debate has centred on tumour MSI status and survival outcomes following neoadjuvant chemotherapy. Post-hoc analysis of clinical trial data has strongly suggested that neoadjuvant chemotherapy may harm patients with MSI-High tumours.[7] According to these data, patients with MSI-High tumours should proceed straight to surgical treatment without chemotherapy. However, statistical evidence supporting this conclusion is limited by small sample sizes and conflicting evidence within other gastric cancer studies and similar gastrointestinal cancers such as colon cancer.[4,8–10] This thesis provides additional insight into MSI-related tumour biology by highlighting the relationship between MSI-High tumours and their propensity to interact effectively with the patient's immune system.

Survival outcomes in gastric cancer would likely be improved by discovering efficacious targeted therapy regimens. Clinical trials for targeted therapy in gastric cancer have found

limited success compared to those in other malignancies, such as melanoma, colon or breast (see: Targeted molecular therapy and immunotherapy). Several molecular classification systems for gastric cancer have been established to inform foundational biology and guide therapeutic hypotheses (see: Molecular classification of gastric cancer). However, there is limited clinical translation of these classification systems. Improvement in technological applications and cost reduction related to molecular classification would aid in the advancement of personalized medicine in gastric cancer. In this work, I actively explore strategies to translate multi-omics molecular classification systems to the clinic. This includes preliminary validation of a molecular test that effectively measures molecular subtypes in gastric cancer. I establish methods that provide insight into similarities and differences between molecular subtypes from distinct classification systems. Finally, I investigate the prognostic and therapeutic effects of multiple established gastric cancer classification frameworks.

The following introduction will provide appropriate context for the problems we will investigate. This thesis contains a heavy mixture of biology and mathematics; however, it is my intent to communicate this work towards a biological science audience. Citations within the methods sections will be available to examine the evidence supporting these approaches.

1.2 Epidemiology

Gastric cancer is the fifth most common cancer and the third most common cause of cancer death worldwide.[11–13] In the mid-1990s gastric cancer was the most common cause of cancer death.[14] Since the first global cancer epidemiology estimates from the 1970s[15] gastric cancer incidence and mortality has decreased, particularly within western and developed countries.[12,13] Despite these decreasing trends, gastric cancer still has poor survival with a mortality-to-incidence ratio ranging from 0.65 to 0.83 in developed and undeveloped countries respectively. Recent analyses reinforce the global trend of decreasing gastric cancer mortality, however in western women aged 30-49 years mortality has remained stable.[14] This suggests that traditional environmental factors contributing to improved gastric cancer mortality such as diet, smoking cessation, oncologic and *Helicobacter pylori* (*H. pylori*) treatment[13] have maximized their effect within this population and additional advances are now required.

The term gastric cancer generally refers to the most common gastric neoplasm, gastric adenocarcinoma, which is the focus of this study. Other less common forms of gastric

neoplasm includes gastric lymphoma, mucosa associated lymphoid tissue lymphoma, gastrointestinal stromal tumor and neuroendocrine tumors.[16] In contrast to gastric adenocarcinoma, these rarer gastric neoplasms are not declining.[14]

Nearly one million new cases of gastric cancer are diagnosed each year. Men are twice as likely to be diagnosed with gastric cancer compared to women.[11–13] The incidence of gastric cancer is highest in East Asia (35.4 per 100,000) and Eastern and Central Europe (20.3 per 100,000) with relatively a low incidence in North America (5.5 per 100,000) and Western Europe (8.8 per 100,000).[11] In absolute terms, gastric cancer is estimated to account for 26,240 new cases and 10,800 deaths in the United States in 2018.[17] According to the Canadian Cancer Statistics 2017, the age-standardized incidence risk for stomach cancer was estimated at 8.6 cases per 100,000 or 3,500 cases, making it the 14th most common cancer.[18]

Distinct epidemiological patterns of gastric cancer exist depending on the anatomical location of the tumor. Cancer occurring in the gastric fundus, corpus and pylorus represent the majority of stomach cancer worldwide. The incidence of fundus and pylorus cancers is characterized by significant geographic variation with the majority of these cancers occurring in Asia and Eastern Europe.[14] Gastric cardia cancer represents an anomaly in the recent history of stomach cancer. Cardia adenocarcinoma has been shown to be stable or increasing worldwide and has less geographic variation compared to other gastric cancer locations.[13,14] The increase in cardia cancer is specifically noted in Europe and North America and is thought to be due to gastroesophageal reflux disease related to rising rates of obesity.[19–22]

1.3 Pathogenesis of gastric cancer

The development of gastric cancer is complex and incompletely understood. In addition to the multifactorial nature of neoplastic disease, gastric cancer is exceptionally heterogenous.[23] Multiple classification schemes have been proposed to understand the histology, morphology and molecular intricacies of gastric cancer.[24]

Traditionally gastric adenocarcinoma histology has been characterized by the Lauren Classification.[25] Formulated in 1965 by Scandinavian pathologist Dr. Pekka Laurén, the classification system has also served as a framework to help explain the molecular and genetic basis of gastric cancer.[23,26–28] The Lauren Classification has clinically relevant

clinicopathologic correlations and has been demonstrated to be an independent prognostic factor.[29–31]

The Lauren Classification defines gastric cancer as intestinal-type or diffuse type. Intestinal-type adenocarcinoma is well differentiated with a preserved architecture of gastric glands. Intestinal-type cancers are known to have a more favorable prognosis, increasing incidence with age and undergoes hematogenous spread. Diffuse-type adenocarcinoma is undifferentiated by appearing disorganized with no glandular architecture on histology. These cancers are known to lack CDH-1 (ie. Epithelial Cadherin or E-Cadherin), occur at younger age, have higher rate of recurrence and are more aggressive leading to an increased mortality.[16,30]

Intestinal-type gastric adenocarcinoma has been theorized to develop through a series of precancerous lesions. The 'Correa Cascade' was developed through histologic observation of the progression of gastritis to intestinal metaplasia to gastric cancer (Figure 1-1).[32,33] The cascade proposes that chronic gastritis leads to the development of chronic atrophic gastritis due to the bacteria *H. pylori* or other environmental factors such as diet. Less commonly, chronic atrophic gastritis is caused by the T cell mediated disease Autoimmune Metaplastic Atrophic Gastritis (AMAG).[34] Atrophic gastritis is characterized by a loss of normal gastric glands. Stomach pH rises due to the decreasing population of acid producing parietal cells. A decrease in ascorbic acid (Vitamin C) levels is also observed. In response to these changes, chief cells increase the secretion of gastrin resulting in mucosal proliferation.[32,33] Hypochloridria and atrophic gastritis is also associated with a change in the gastric microbiome.[35,36] Anaerobic bacteria populations proliferate, which results in the production of carcinogenic nitrosamines.[13] Interestingly, acid lowering medications such as proton pump inhibitors (PPIs) have not been shown to increase gastric cancer risk even in *H. pylori* affected patients.[37,38] Chronic atrophic gastritis eventually leads to the development of intestinal metaplasia.

Intestinal metaplasia is a preneoplastic lesion identified by the presence of intestinal cells within the gastric mucosa such as mucus-secreting goblet cells.[39] Immunohistochemistry (IHC) also identifies small intestine digestive enzymes such as sucrase.[32] Intestinal metaplasia exists within a spectrum of disease ranging from complete type (Type I) characterized by small intestine epithelial cells to the higher risk incomplete type (Type II and Type III) characterized by colonic epithelial cells.[34] Metaplasia is associated with approximately 80% of intestinal-type gastric cancer and even up to 60% of diffuse-type

cancer, though the association with diffuse-type cancer is variable and incompletely understood.[40,41]

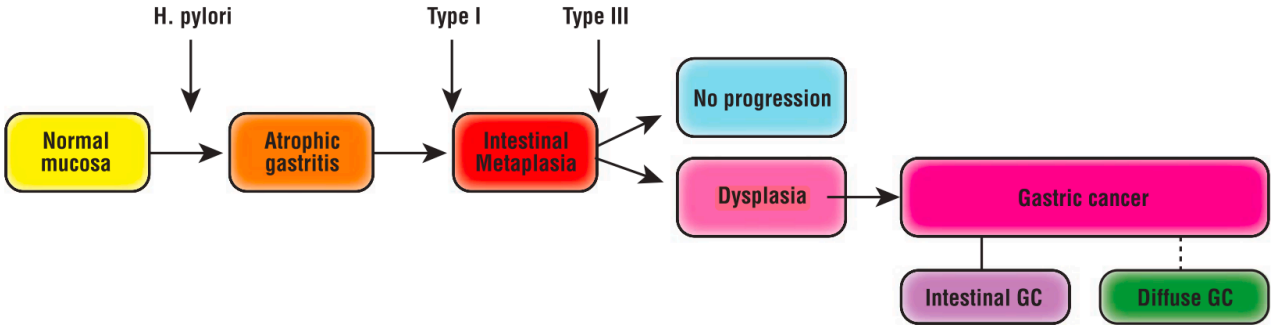


Figure 1-1 The 'Correa Cascade' demonstrating the pathogenesis of gastric cancer. (From Busuttil RA, Boussioutas A. Intestinal metaplasia: A premalignant lesion involved in gastric carcinogenesis. J Gastroenterol Hepatol. 2009;24(2):193-201.)

Intestinal metaplasia may represent a "point of no return" in the cascade as demonstrated by the discovery that *H. pylori* eradication does not prevent the progression to cancer in some patients with metaplasia.[42] Unfortunately gastric intestinal metaplasia is asymptomatic and can only be diagnosed via endoscopic biopsy at this time. A number of biomarkers including p53 and markers of microsatellite instability have been investigated and thus far they have failed to reliably identify intestinal metaplasia and which lesions may progress to cancer.[39]

Dysplasia is the final non-malignant lesion identified in the Correa Cascade. It is characterized by neoplastic type cell structure and architecture that is still confined by the basement membrane.[32] The degree of dysplasia may be specified as G1 (mild), G2 (moderate) or G3 (severe). In a multicenter prospective study containing one hundred twelve patients, the progression of mild, moderate and severe dysplasia to cancer was 21%, 33% and 57% respectively.[43]

1.4 Etiology

1.4.1 Genetic causes

The majority of gastric cancer is sporadic with 1-3% of gastric cancers arising from hereditary conditions. Genetics still play an appreciable role as 10% of gastric cancer can be attributed to familial aggregation. Gastric cancer risk is three-fold in individuals with a first-degree relative affected by gastric cancer. Genomic study has implicated single nucleotide polymorphisms (SNPs) and other genetic mutations to the familial and population based risk of gastric cancer.[13] Genetic variation of the Prostate Stem Cell Antigen (PSCA)[44] and Mucin 1 (MUC1)[45] genes increases susceptibility to gastric cancer. Among the high risk Japanese population these genotypes occur simultaneously in 56% of the population and as a single mutation in approximately 95% of the population.[46]

The majority of familial gastric cancer results from three syndromes: hereditary diffuse gastric cancer (HDGC), gastric adenocarcinoma and proximal polyposis of the stomach (GAPPS) and familial intestinal gastric cancer (FIGC). The specific genetic causes for GAPPS and FIGC are not yet explained. Gastric adenocarcinoma and proximal polyposis of the stomach is inherited in an autosomal dominant pattern and exhibits incomplete penetrance.[47] Autosomal dominant inheritance also likely explains FIGC transmission.[48]

Other notable inherited conditions that are known to cause gastric cancer, albeit less frequently, include Li-Fraumeni syndrome, Lynch syndrome/hereditary nonpolyposis colorectal cancer (HNPCC), Peutz-Jeghers syndrome (PJS) Cowden syndrome and familial adenomatous polyposis (FAP).[13,49] Germline mutations accounting for these diseases occur in well-known tumor suppressor genes and microsatellite instability markers including tumor protein p53 (p53), MutL homolog 1 (MLH1), serine/threonine kinase 11 (STK11), phosphatase and tensin homolog (PTEN) and adenomatous polyposis coli (APC).

Hereditary diffuse gastric cancer is a well characterized cause of stomach cancer. Central to the formation of HDGC is impaired cellular adhesion due to the loss of E-cadherin (CDH1 gene). The disease is inherited in an autosomal dominant fashion and exhibits high penetrance. The risk of gastric cancer by age 80 in patients with HDGC is greater than 80% and the average age of diagnosis is age 38.[50,51] HDGC is associated with an increased risk of lobular breast cancer, which is also characterized by the loss of E-cadherin.

E-cadherin is a homodimeric transmembrane cellular adhesion protein which interacts with catenins to form cellular adhesion complexes.[52] Carriers of the CDH1 mutation responsible

for HDGC lack one allele for the CDH1 gene. The second allele may become inactivated via mutation or loss of heterozygosity but is most commonly inactivated by promoter methylation.[53,54] HDGC progresses from non-neoplastic tissue to an *in situ* signet ring cell carcinoma with 'pagetoid' spread prior to forming an invasive signet-ring cell carcinoma.[55] Although the cytotoxin associated gene (CagA) positive *H. pylori* may affect the function of E-cadherin, *H. pylori* does not lead to HDGC.[56] Similarly, the majority of diffuse gastric cancer is not due to HDGC, but is associated with a sporadic loss in E-cadherin.[57]

A signet ring cell is identified on histology by the absence of cytoplasmic and nuclear contents (Figure 1-2). This is the result of the increased production of intracellular mucin, which pushes the intracellular contents to the periphery of the cell. Signet ring cell carcinoma has long been associated with a worse prognosis and a higher propensity for peritoneal carcinomatosis.[58] This progression makes the early diagnosis and screening of these individuals difficult and thus prophylactic gastrectomy is a reasonable treatment option.[16]

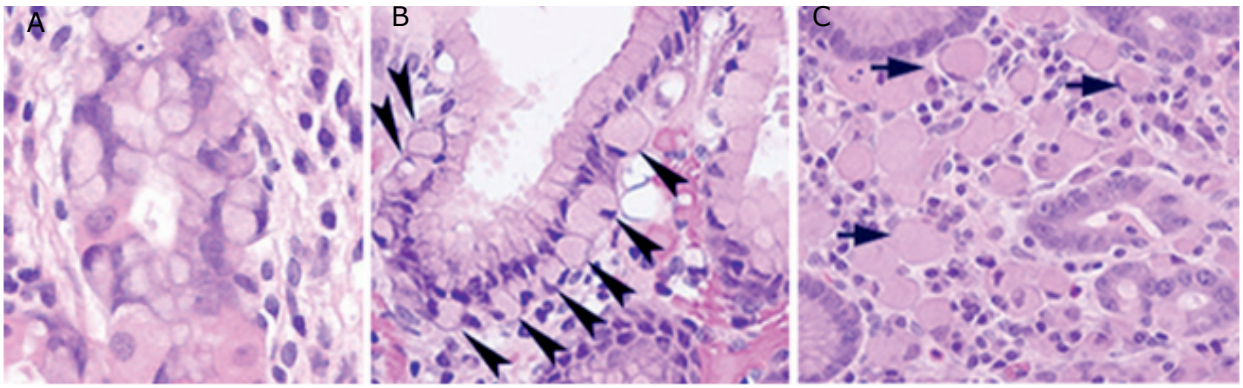


Figure 1-2 Representative images of signet ring cells. **A.** In situ signet ring cell carcinoma. **B.** Signet ring cells demonstrate 'pagetoid spread' along the basement membrane. Note the large vacuole of mucin displacing the nucleus to the periphery of the cell (arrows). **C.** Invasive signet ring cell carcinoma with signet ring cells present beyond the basement membrane in the lamina propria. (From Fitzgerald RC, Hardwick R, Huntsman D, et al. Hereditary diffuse gastric cancer: updated consensus guidelines for clinical management and directions for future research. J Med Genet. 2010;47(7):436-444.)

1.4.2 Environmental

Modifications of environmental factors for gastric cancer have contributed significantly to the worldwide decline in gastric cancer incidence. The World Health Organization (WHO)

identifies environmental factors to be essential in the primary prevention strategy for gastric cancer.

Foods preserved by salting or pickling and all meats increase the risk of gastric cancer.[13] Conversely, the consumption of fresh fruits and vegetables as well as the Mediterranean diet composed of legumes, seeds and olive oil is protective against cancer. Decreased incidence in gastric cancer is thus intimately related to the introduction of refrigeration in developed nations.[13,59] The carcinogenic effects of tobacco smoke also applies to gastric cancer. Specifically, the molecular carcinogenesis related to CagA *H. pylori* is potentiated by smoking. Obesity has been shown to be a risk for all gastric cancer types[60] and the International Agency for Cancer Research identified obesity as a cause of gastric cardia cancer only.[61] Gastric cancer is influenced by both low and high socioeconomic status. Increased gastric cancer has been attributed to lower socioeconomic status[16] and increased cardia cancer is identified specifically in those with higher socioeconomic class.[22]

1.4.3 Helicobacter pylori

Helicobacter pylori is a spiral shaped gram-negative bacterium that has been recognized on human gastric mucosa for over 100 years. It was first designated as *Campylobacter pylori* but was subsequently renamed *Helicobacter pylori* in 1989.[62] By the early 1990s significant evidence implicated *H. pylori* as a cause for gastric adenocarcinoma.[63]

The International Agency for Research on Cancer has designated *H. pylori* as a Group 1 carcinogen[64] as nearly all non-cardia cancer development may stem from *H. pylori* infection.[62] *H. pylori* has a global prevalence of approximately 50% making it the most common chronic infection worldwide.[65] The acquisition of infection often occurs in early childhood and persists until eradication.[13] Given the prevalence and implications for gastric cancer, *H. pylori* represents an intriguing target of research, screening programs and treatment for gastric cancer.

H. pylori has adapted to live in the low acidic environment present on gastric mucus-secreting cells and mobilize within the mucous layer using their distinct unipolar flagella. The bacteria are also able to enhance their micro-environment with the enzyme urease, which hydrolyzes urea with a concurrent release of ammonia that neutralizes stomach pH.[62,66] Clinically relevant pathogenesis of *H. pylori* and gastric cancer resides primarily within the genes CagA and vacuolating cytotoxin (VacA).[67] Specifically, the VacA s1/m1 strain and CagA genes are associated with the Correa cascade and the progression of

atrophic gastritis to gastric cancer.[25,60] VacA is contained in all *H. pylori*, however, it is not expressed unless the CagA pathogenicity island is co-expressed. VacA and CagA result in increased tissue inflammation and cytokine release in the gastric mucosa.[69] An important discrepancy among VacA strains then influences the underlying pathophysiology of gastric disease. In patients infected with CagA-positive VacA s1m1 *H. pylori* the resultant inflammation results in the development of preneoplastic lesions and progression to gastric cancer in keeping with the Correa cascade. Alternatively, CagA-negative VacA s2m2 expression results in gastric inflammation leading to non-atrophic corpus sparing gastritis and a higher risk of forming duodenal ulcers. This development of duodenal ulcers has actually been shown to have a decreased risk of gastric cancer.[70,71]

The effect of *H. pylori* eradication and the resulting decrease in gastric cancer has yielded mixed results. Meta-analyses have demonstrated that *H. pylori* treatment reduces gastric cancer incidence[72] while other data demonstrate that the incidence of metachronous gastric cancer is unchanged.[73,74] In a robust systematic review and meta-analysis published in 2016, Lee et al. demonstrated that *H. pylori* eradication reduced the incidence of gastric cancer across all levels of risk, however maximal benefit was in patients with asymptomatic infection and those treated after endoscopic resection of gastric cancer.[75] Evidently, the timing of eradication is important as *H. pylori* eradication does not prevent the progression to cancer in some patients who have already established metaplasia.[42]

The effect of *H. pylori* on causing gastric cancer at the population level is incompletely understood. Various epidemiological studies have demonstrated discrepancy between gastric cancer incidence and *H. pylori* infection.[65] Only a small number of individuals infected with H. pylori develop gastric cancer and despite regional variability in gastric cancer incidence, the incidence of *H. pylori* is relatively uniform[70,76,77] Modern classification systems have failed to designate a unifying molecular and genomic classification of *H. pylori* related cancer. On the genomic scale of gastric cancer other driver mutations and Epstein-Barr virus (EBV) infection have thus far been determined to represent significant sub-types of gastric cancer.[26,57] This effect is likely multifactorial in nature and is related to host, bacterial and environmental factors. Ongoing evaluation of *H. pylori* in epidemiological, clinical and basic science research is still required to fully understand the relationship of this bacteria to human disease.

1.4.4 Epstein-Barr Virus

Epstein-Barr virus is a relevant contributor to gastric cancer pathogenesis. Approximately 10% of gastric cancer specimens are infected with EBV.[78] Association of EBV infection with clinicopathologic indices and prognosis is well documented.[79] Despite the growing understanding of the molecular associations between EBV infection and gastric cancer a precise mechanism of pathogenesis is currently lacking.[80,81]

The molecular classification for gastric cancer has reinforced this notion. In The Cancer Genome Atlas (TCGA) project, EBV associated gastric cancer stratified into one of four notable subtypes of gastric cancer. They found that EBV related tumors occur most frequently in the fundus and body of the stomach and that the major mutational driver was related to DNA hypermethylation via CpG island methylator phenotypes.[82] When assessed at a genomic level, the TCGA classification of EBV-type gastric cancer was also found to have the best recurrence free survival and overall survival.[83] The Asian Cancer Research Group (ACRG) also demonstrated the significance of EBV-positive gastric cancer and associated mutations. However, ACRG did not discover an EBV-type gastric cancer group but instead found EBV related tumors to exist within the microsatellite stable Tumor Protein 53 positive (MSS TP53+) subtype.[26]

Improvement in survival is not fully understood but high prevalence of immune cells in proximity to the growth of EBV related cancer likely plays a role. The ACRG characterized EBV related tumors to have an increased cytokine signature.[26] The side-by-side growth of cancer with immune cells influenced by EBV may enhance the ability of native CD8 T lymphocytes to eradicate EBV related gastric cancer.[79] The molecular analysis of EBV related gastric cancer also suggests that EBV infection may relate to more favourable molecular mutations. For example, EBV positive gastric cancer has been shown to have higher expression of Phosphatidylinositonol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Alpha (PIK3CA), programmed cell death-ligand 1 (PD-L1) and programmed cell death-ligand 2 (PD-L2) expression.[84,85] The immunogenic profile of EBV related cancer makes it an intriguing target for immunotherapy with medications such pembrolizumab, which inhibits the interaction of PD-1 receptor to its ligand PD-L1/L2.

1.5 Presentation and Diagnosis

1.5.1 Clinical presentation

Gastric cancer is difficult to diagnose, especially in lower incidence regions. Early symptoms are often attributed to more common diseases such as peptic ulcer disease or GERD. Most common symptoms on presentation include abdominal pain, weight loss, nausea, anorexia and early satiety.[16,86] Due to the difficulty in diagnosis, gastric cancer is often diagnosed at an advanced stage.[87] Advanced disease may present with gastric outlet obstruction, dysphagia, bleeding or ascites. Anemia may develop from acute or chronic bleeding. Systemically advanced disease is also associated with cachexia, malaise and weakness.[88]

Focused physical exam and history is required. Approximately 25% of patients have previous history of gastric ulcer.[76] Family history may reveal hereditary gastric cancer syndromes or other genetic risks for gastric cancer. Metastatic disease may present on physical exam in one of five anatomic locations. Lymph node basins include the left supraclavicular node (Virchow's node) and left axillary node (Irish's node). Palpable metastases may present as a peri-umbilical bulge (Sister Mary Joseph nodule), via pelvic examination as ovarian deposits (Krukenberg tumour) or by digital rectal exam as anterior peritoneal metastasis (Blumer shelf).[16] Peri-gastric lymph nodes are not palpable.

1.5.2 Diagnostic investigation and staging

Investigation is warranted when appropriate suspicion is achieved to suspect the potential diagnosis of gastric cancer. Laboratory investigation includes basic complete blood count, electrolyte evaluation, coagulation tests and liver enzyme and function tests. The primary diagnostic modality of choice is flexible upper endoscopy.[16,87] This method is both diagnostic, and in cases of early gastric cancer and dysplastic lesions, therapeutic. Endoscopic appearance of masses, ulcerations, or suspicious lesions warrant biopsy. Endoscopic biopsies are then examined by a pathologist to determine the presence of gastric cancer. Other modalities of diagnosis include computed tomography (CT), positron emission tomography CT (PET CT), barium enhanced abdominal radiography or magnetic resonance imaging (MRI). Often these preliminary investigations result in the patient being referred for endoscopy to confirm and characterize the presence of disease.

Ideally, appropriate staging investigations are required prior to treatment. Systemic staging assesses for lung and abdominal metastasis. Chest and abdominal imaging is achieved with

a venous contrast enhanced CT.[16] Another approach is to use PET CT, however it is only indicated in individuals with PET avid tumours, which comprises only about 50% of gastric cancer tumours. Gastric cancer has a propensity for peritoneal metastasis. Peritoneal metastasis, also known as carcinomatosis, is difficult to identify on CT (sensitivity = 51%).[16] Approximately 30% of gastric cancer patients with negative cross-sectional imaging presents with peritoneal carcinomatosis.[89,90] Thus, in this population a staging laparoscopic evaluation of the abdominal cavity, biopsy of suspicious lesions and peritoneal washing cytology is warranted. Locoregional staging of depth of tumor invasion and perigastric lymph node involvement may be assessed via endoscopic ultrasound (EUS). Studies have demonstrated that EUS accuracy of diagnosis ranges from 46.2 – 80% for T stage and 66.7 – 85% for N stage.[16,87,91] T and N stages are discussed below (Pathologic staging of gastric cancer).

1.5.3 Screening

The effectiveness and affordability of gastric cancer screening programs is currently dependent on the regional incidence of disease. Since the introduction of Japanese gastric cancer screening programs in the 1970s, the mortality of gastric cancer has declined 50%.[16] In appropriate patient populations and high incidence regions both endoscopic and

Table 1-1 Gastric cancer TNM stage definitions per AJCC/UICC eight edition

Primary tumor (T)	
T category	T criteria
TX	Primary tumor cannot be assessed
T0	No evidence of primary tumor
Tis	Carcinoma in situ: Intraepithelial tumor without invasion of the lamina propria, high-grade dysplasia
T1	Tumor invades the lamina propria, muscularis mucosae, or submucosa
T1a	Tumor invades the lamina propria or muscularis mucosae
T1b	Tumor invades the submucosa
T2	Tumor invades the muscularis propria
T3	Tumor penetrates the subserosal connective tissue without invasion of the visceral peritoneum or adjacent structures
T4	Tumor invades the serosa (visceral peritoneum) or adjacent structures
T4a	Tumor invades the serosa (visceral peritoneum)
T4b	Tumor invades adjacent structures/organs
Regional lymph nodes (N)	
N category	N criteria
NX	Regional lymph node(s) cannot be assessed
N0	No regional lymph node metastasis
N1	Metastasis in one or two regional lymph nodes

N2	Metastasis in three to six regional lymph nodes
N3	Metastasis in seven or more regional lymph nodes
N3a	Metastasis in 7 to 15 regional lymph nodes
N3b	Metastasis in 16 or more regional lymph nodes
Distant metastasis (M)	
M category	M criteria
M0	No distant metastasis
M1	Distant metastasis

biomarker screening with serum pepsinogen have been demonstrated to be cost effective.[92,93] Recently, recommendations advocating the utility of gastric cancer screening programs in high-risk American Korean patients were proposed.[94] In this high-risk population, patients of Korean descent with atrophic gastritis, intestinal metaplasia or a family history of gastric cancer may benefit from early screening programs, however, whether this would improve overall survival is currently unknown.

1.6 Pathologic staging of gastric cancer

Formal clinical and pathologic gastric cancer classification systems have become essential to the prognosis and treatment of cancer. The American Joint Committee on Cancer (AJCC) and Union for International Cancer Control (UICC) is the most relevant staging system for gastric cancer worldwide.[16,95,96] Other staging systems include the WHO Classification of Tumours of the Digestive System[97], which stratifies stomach adenocarcinoma into papillary, tubular, mucinous, mixed carcinoma, and signet ring cell/poorly cohesive.[24] The utility of the WHO classification system is clinically limited[16] and thus the AJCC/UICC system and complimentary systems will be outlined.

The AJCC/UICC system stages cancer according to depth of tumor invasion (T stage), burden of lymph node disease (N stage) and the presence or absence of metastasis (M stage) (Table 1-1). Depending on the composition of TNM findings, cancer stage is then defined as Stage I-IV disease in which stage IV represents the most advanced stage of cancer (Table 1-2). The current eighth edition of the classification system was released in 2017. The formulation of newest edition was based off the International Gastric Cancer Association staging project in which the majority of cases were Japanese and Korean (84.8%) as opposed to western (8.8%). Additional studies validating the system in patients from the United States demonstrated 5 year overall survival rates ranging from 74.2% in Stage I disease to 6.7% in Stage IV disease.[95,96]

Table 1-2 Gastric cancer prognostic stage group allocation per AJCC/UICC eighth edition

When T is…	And N is…	And M is…	Then the stage group is…
Tis	N0	M0	0
T1	N0	M0	I
T2	N0	M0	I
T1	N1, N2, or N3	M0	IIA
T2	N1, N2, or N3	M0	IIA
T3	N0	M0	IIB
T4a	N0	M0	IIB
T3	N1, N2, or N3	M0	III
T4a	N1, N2, or N3	M0	III
T4b	Any N	M0	IVA
Any T	Any N	M1	IVB

1.7 Management

Approaches to the treatment of gastric cancer worldwide is heterogenous. Regardless, the mainstay of therapy typically consists of a combination of surgery, chemotherapy, radiotherapy and molecular therapy. Over the last two decades gastric cancer treatment has generally become more uniform internationally. Discrepancy between regional treatment can be summarized by the more frequent use of neoadjuvant chemotherapy in the west versus primary gastric resection followed by S-1 chemotherapy in the east.[16]

1.7.1 Surgery

Surgical treatment of gastric adenocarcinoma is largely dependent on cancer stage. For example, early gastric cancer may be treated with local endoscopic resection while advanced tumors may not be surgical candidates.

Early gastric cancer defined as T1a tumors which invade no deeper than the submucosa may be treated with endoscopic submucosal dissection (ESD) or endoscopic mucosal resection (EMR). Prior to endoscopic therapy, endoscopic ultrasound is performed to ensure appropriate patient selection.[98] Additional high-risk criteria have been developed for endoscopic therapy to ensure adequate treatment and low disease recurrence. Endoscopic therapy is contraindicated if tumor extends beyond mucosa, lymphovascular invasion is present, tumour is larger than 2 cm, ulceration is present or histopathology demonstrates undifferentiated cancer.[99] Using these criteria, retrospective data from Asia demonstrated no disease-specific mortality, 77% complete resection (R0) and 6% recurrence rate with 39 month-median follow up.[100] If R0 resection is achieved then generally no additional

chemotherapy or radiation is required. Recently, criteria for endoscopic therapy has been expanded[101] with similar recurrence rates between absolute and expanded criteria.[102]

Tumours invading the gastric submucosa (T1b) must be treated with abdominal surgery and either total or sub-total gastrectomy depending on tumor location. These tumours may proceed directly to surgery without prior neoadjuvant therapy.[103] Once tumour invasion reaches the muscularis propria (T2) neoadjuvant chemotherapy and subsequent surgery is recommended if possible. Locally advanced gastric cancer represents a therapeutic challenge, requiring multidisciplinary treatment consideration. In appropriately selected patients, surgery may be beneficial. In particular, neoadjuvant chemotherapy may be used to decrease tumour burden to allow a R0 resection.[16] The presence of metastasis including peritoneal carcinomatosis is almost universally not treated with curative intent surgery. Surgery can play a role however in palliation for symptom control.

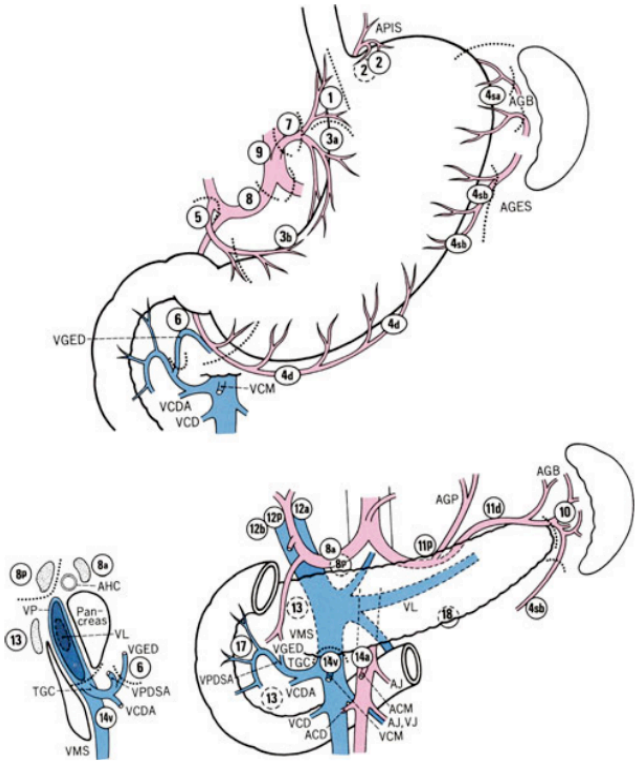


Figure 1-3 Lymph node stations of the stomach according to the Japanese Gastric Cancer Association. (Adapted from Sano T, Kodera Y. Japanese classification of gastric carcinoma: 3rd English edition. Gastric Cancer.)

Significant debate regarding the required extent of lymph node resection has resulted in multiple large randomized control studies.[104–106] Lymphadenectomy for gastric cancer can be classified broadly into D1 and D2 resections. A formal D1 resection includes the perigastric

nodes, which correspond to the Japanese lymph node stations 1 to 7 (Figure 1-3).[107] A D2 resection includes the perigastric nodes as well as nodes along the hepatic, left gastric, splenic and celiac arteries and the splenic hilum. The culmination of these studies resulted in preference given to the more extensive D2 resection over D1 resection.[16] In an experienced surgeon's hands, D2 resection with splenic preservation results in lower recurrence, increased disease-free survival but no difference in overall survival and does not significantly contribute to additional morbidity and mortality. Furthermore, accurate nodal staging has prognostic implications and thus 15 nodes are recommended in resected specimens so appropriate cancer staging can occur.[96]

1.7.2 Chemotherapy and chemoradiotherapy

The development of clinically beneficial chemotherapy regimens for gastric cancer has been challenging. However, data now demonstrate a variety of efficacious treatment regimens. Chemotherapy may be delivered after surgery (adjuvant chemotherapy) or before (neoadjuvant chemotherapy) and after.

The combination of chemotherapy and radiation therapy has demonstrated improved outcomes for gastric cancer. External beam radiotherapy and 5-fluorouracil was investigated in the Southwest Oncology Group (9008/INT-0116) randomized control trial (RCT).[108] This combination therapy was shown to improve three-year overall survival (41% versus 50%) and recurrence free survival (41% versus 64%). However, inadequate lymph node resection and inability to complete treatment due to toxicity confound this approach. Chemoradiotherapy treatment is primarily performed in oncology centers from the United States.

Various approaches to chemotherapy treatment have been investigated. Relevant combinations with clinical benefit include capecitabine plus oxaliplatin. In Asia, S-1 chemotherapy, consisting of tegafur (prodrug of 5-fluorouracil), 5-chloro-2, 4-dihydropyridine, and oxonic acid is a common treatment protocol.[109] The concept of neoadjuvant chemotherapy was developed in an attempt to improve patient tolerance and the need to down-stage locally advanced gastric tumors to achieve a R0 resection.[16,87] The MAGIC protocol consisting of epirubicin, cisplatin and 5-fluorouracil (ECF) delivered in three, three-week cycles preoperatively and postoperatively was a prominent neoadjuvant protocol. In a clinical trial, the MAGIC protocol resulted in improved preoperative treatment tolerance and decreased rate of recurrence, decreased metastasis and an improved 5-year overall survival compared to surgery alone.[110]

The most recent advancement in neoadjuvant chemotherapy is the FLOT protocol developed from the Arbeitsgemeinschaft Internistische Onkologie group. This protocol consists of 5-fluorouracil, leucovorin, oxaliplatin and docetaxel delivered in three, three week preoperative and postoperative cycles.[111] In a Phase II/III randomized control trial FLOT was compared to ECF (MAGIC) and epirubicin, cisplatin, and capecitabine (ECX). This trial demonstrated that FLOT results in higher rates of complete pathologic regression with a reasonable side-effect profile compared to ECF/ECX. In 2019, results of the Phase III trial established that FLOT significantly improved median overall survival (50 versus 35 months) and estimated five-year overall survival (45 versus 36%).[112,113] As a result of these findings, FLOT chemotherapy is now the primary treatment regimen used by our center, the Cross Cancer Institute at the University of Alberta.

1.7.3 Targeted molecular therapy and immunotherapy

Molecular therapy is a relatively new oncologic treatment modality. Advancements in molecular and genetic biology have resulted in the characterization and understanding of numerous driver mutations, genes, proteins and immune factors related to cancer pathogenesis.[114,115] New targeted therapy against specific molecules and immunotherapy designed to enhance or activate the host immune response to cancer has shown remarkable efficacy in a variety of cancers. As a result of these discoveries, the hope for personalized cancer treatment is expanding to gastric cancer therapy.

Numerous clinical trials investigating novel molecular therapy and immunotherapy approaches for gastric cancer are underway or completed. Thus far the investigation of these novel treatments has been mostly restricted to advanced gastric cancer.[116] The application of effective molecular therapy has proven difficult and even less evidence exists for first-line molecular therapy. A 2016 Cochrane review found that first-line molecular therapy for gastric cancer may have a small effect on survival, likely increases the risk of adverse events and that quality of evidence is low or very low.[117]

In the Avastin in Gastric Cancer (AVAGAST) study, bevacizumab, a vascular endothelial growth factor-A (VEGF-A) inhibitor, was used in conjunction with capecitabine and cisplatin. This phase III RCT failed to improve median overall survival. Interestingly, a sub-group analysis demonstrated that bevacizumab therapy revealed improved overall survival in non-Asian patients.[118] The Erbitux in Combination With Xeloda and cisplatin in Advanced Esophagogastric Cancer (EXPAND) trial failed to improve progression free survival in metastatic gastric cancer patients treated with cisplatin, capecitabine and the epidermal

growth factor receptor (EGFR) inhibitor cetuximab.[119] Inhibition of the mammalian target of rapamycin (mTOR) protein kinase via everolimus also failed to improve overall survival in the Gastric Anti-Tumor Trial with everolimus (GRANITE-1) phase III trial.[120] Gastric cancer requires additional biomarker research as many trials have demonstrated limited or no improvement in outcomes despite targeted treatment of these important molecular drivers of cancer.

Effective molecular therapy for gastric cancer does exist. Targeted therapy inhibiting human epidermal growth factor receptor 2 (HER2) and vascular endothelial growth factor receptor-2 (VEGFR-2) has yielded promising results in clinical trials. The evaluation of anti-HER2 treatment as a first-line therapy for gastric cancer is represented by two trials: Trastuzumab for Gastric Cancer (TOGA)[121] and Lapatinib Optimization Study in the HER2-positive Gastric Cancer (LOGiC).[122] Labatinib (small molecule inhibitor of epidermal growth factor type 1 and HER2) monotherapy and labatinib plus capecitabine and oxaliplatin failed to show improvement in overall survival and treatment response.[122] However, the TOGA trial demonstrated a statistically significant improvement in gastric cancer median overall survival when trastuzumab (anti-HER2 monoclonal antibody) was administered in combination with chemotherapy.[121] Treatment targeting VEGFR-2 was evaluated in the RAINBOW[123] trial and the REGARD trial.[124] The RAINBOW trial established that the VEGFR-2 antagonist ramucirumab in combination with paclitaxel versus paclitaxel plus placebo resulted in statistically significant increases in overall survival (median 9.6 months vs 7.4 months) for progressive advanced gastric or gastroesophageal junction (GEJ) cancer.[123] Ramucirumab was the first targeted therapy to be effective as a single agent as demonstrated in the REGARD trial.[124] Patients with progressive gastric or GEJ cancer following first-line chemotherapy were treated with ramucirumab or placebo. Treatment with ramucirumab yielded a median overall survival of 5.2 months versus 3.8 months in the placebo arm. Besides increased hypertension with anti-VEGFR-2 treatment, adverse events were similar between both groups.

Immunotherapy enhances the host immune response to eradicate cancer cells. The immune response may be augmented by activating tumor specific cytotoxic T cells, preventing immune cell exhaustion and inhibiting the ability of cancer cells to evade or diminish the immune response. In gastric cancer, the majority of immunotherapy evidence exists for checkpoint inhibition via blockade of programmed cell death-1 (PD-1) and cytotoxic T-lymphocyte-associated protein 4 (CTLA-4). Cancer cells that present programmed cell death ligand 1 (PD-L1) bind to PD-1 expressed on host immune T cells, which then causes a

downregulation in the host immune response. Similarly, CTLA-4 is expressed on our own immune cells and functions as a "brake" by downregulating T cell activation.[125]

In the phase Ib Keynote-012 and phase II Keynote-059 trial, pembrolizumab (anti-PD-1 monoclonal antibody) was used as a third-line medication in progressive advanced gastric cancer.[126,127] Keynote-059 found that the anti-PD-1 monotherapy response rate was 11.6% and median response duration was 8.4 months with a reasonable side-effect profile.[127] The combination of PD-1 and CTLA-4 inhibition was evaluated in the CheckMate-032 trial.[128] In this trial, third line esophagogastric therapy of nivolumab (anti-PD-1 monoclonal antibody) and nivolumab plus ipilimumab (anti-CTLA-4 antibody) was found to have an objective response rate of 12% for nivolumab only and 8-24% for combined therapy groups. The 12-month progression-free survival rates were 8% for nivolumab only and 10-17% for combined therapy. Furthermore, both trials demonstrated that therapy was effective in both PD-L1 positive and negative tumours.[127,128] In a Phase III trial, first-line pembrolizumab as monotherapy or in combination with chemotherapy was found to be non-inferior to chemotherapy alone in advanced gastric cancer patients.[129] Although immunotherapy did not provide significant improvement in overall survival or progression-free survival, patients treated with pembrolizumab experienced fewer adverse events.

The three clinically relevant targeted treatment modalities in 2018 are trastuzumab (anti-HER2), ramucirumab (anti-VEGFR2) and pembrolizumab (anti-PD-1). In 2017 the FDA approved the use of pembrolizumab for gastric cancer due to results of the Keynote-059 trial and the discovery that solid tumours with mutations resulting in mismatch repair deficiency (dMMR) were more susceptible to anti-PD-1 therapy.[127,130,131] The National Comprehensive Cancer Network (NCCN) guidelines for the systemic treatment of unresectable locally advanced, recurrent or metastatic disease list trastuzumab as a first-line treatment in combination with capecitabine/fluorouracil and cisplatin in tumours that overexpress HER2.[103] Ramucirumab may also be used as a second-line or subsequent therapy in combination with paclitaxel or as monotherapy. Progress in molecular therapy of gastric cancer is possible and has introduced modifications to how gastric cancer is now treated.

1.8 Methods to Advance Personalized Medicine in Gastric Cancer

1.8.1 Predictive, diagnostic, therapeutic and prognostic biomarkers in cancer

With the rapid proliferation of molecular and genetic analyses of cancer, increasing numbers of potential biomarkers are being discovered. An adherence to nomenclature is important in order to effectively characterize novel biomarkers. The Biomarkers Definition Working Group, composed of leaders in the National Institutes of Health, the Food and Drug Administration and subject experts, define a biomarker as a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention.[132,133]

Biomarkers may be further classified as predictive, prognostic, diagnostic and/or therapeutic.[116,134,135] A predictive biomarker is able to determine which patients may respond to a particular therapy. Prognostic markers may provide suggestions for further treatment as well as provide information regarding survival and particular therapy outcomes. A therapeutic biomarker is a potential molecular or immunologic target for treatment. Finally, a diagnostic biomarker may provide information regarding the presence of cancer or lesions with a risk of forming cancer.

1.8.2 Immunohistochemistry based biomarkers

Using immunohistochemistry (IHC) to identify the expression and localization of proteins in tumours is a proven method to discover biomarkers. In some cases, a handful of IHC markers can reasonably recapitulate complex molecular subtypes as demonstrated by the relationship between the Predication Analysis of Microarray 50 (PAM50) subtypes in breast cancer and the expression of estrogen, progesterone and HER2 proteins.[136] This technique also provides predictive and therapeutic guidance given that IHC spatially localizes protein expression within the tumour or cancer cell. For example, the allocation of HER2 therapy is contingent on the presence of adequate HER2 expression that must be present on the cell membrane.[137]

On April 18, 2022, the PubMed search "immunohistochemistry and biomarker and gastric cancer" provided over 5000 results. Despite the extensive investigation into IHC biomarkers in gastric cancer, clinically relevant use is limited mainly to HER2 and Mismatch Repair Protein expression.[137,138] Recent utility of IHC has been demonstrated in characterizing tumour infiltrating lymphocytes (TILs).[139–141] Lymphocytes are comprised of adaptive and innate immune cells including T-lymphocytes, B-lymphocytes, monocytes, eosinophils,

basophils and neutrophils, respectively. These cells significantly contribute to tumour immune response via direct killing or immune response coordination.[142,143]

Research of TILs has demonstrated that the story is more complex than the mere presence or absence of immune cells. Instead, the populations or mixture of different immune cell types *and* their location is proving to be more important.[143] This is highlighted by the characterization of "Hot" and "Cold" tumours. A "Hot" tumour contains TILs that are functionally active against the tumour. A "Cold" tumour lacks TILs. Cancer is also able to prevent access to the immune system as observed in "Excluded" TILs. In this case, immune cells are unable to penetrate the tumour despite appropriate recruitment of immune cells by the host immune system to the tumour location. Lastly, "Immunosuppressed" TILs may be present within a tumour but are unable to exert a meaningful effect. This phenomenon is related to the concept of immune exhaustion and immune checkpoint blockade.[144] Tumours with "Immunosuppressed" TILs are potential candidates for immune checkpoint blockade inhibitor therapy, which acts to remove the break of checkpoint inhibition and unleash the cytotoxic potential of the immune cells already present in the tumour. The Immunoscore in colon cancer uses IHC to identify $CD3^+$ and $CD8^+$ TILs.[141] Quantification of the presence of $CD3^+$ and $CD8^+$ T cells in combination with their spatial location provides a robust prognostic indicator. To date, the Immunoscore is limited to colon and rectal cancer and research on TILs within gastric cancer is comparatively limited.

The following sections describe the molecules we investigate using IHC and their relevance to gastric cancer pathology.

1.8.2.1 Epithelial Cadherin

Epithelial cadherin, also known as E-cadherin is encoded by the CDH1 gene located on chromosome 16q22.1. This transmembrane homodimeric protein is an integral calcium-dependent adhesion molecule.[52] It is arguably one of the most investigated molecules and is related to the development of a variety of cancers including gastric and breast cancer.

In hereditary diffuse gastric cancer, autosomal dominant inheritance of a single CDH1 allele allows E-cadherin to function as a tumor suppressor protein. Individuals are then exposed to a high risk of cancer after subsequent knockout of the single inherited gene allele.[54,55,145] The single CDH1 allele is most commonly inactivated via promoter hypermethylation and results in approximately 80% chance of gastric cancer by age 80.[50,51]

Sporadic mutation of E-cadherin is more common. At least 50 percent of diffuse and 26 percent of intestinal-type cancers contain a mutation in CDH1.[145–147] The loss of cellular adhesion regulation allows diffuse gastric cancer to spread under the gastric mucosa as a non-invasive neoplasm prior to becoming clinically recognizable as an invasive cancer. In some cases diffuse gastric cancer encompasses the entirety of gastric stroma with the absence of mucosal alteration resulting in the clinical presentation of linitis plastica.[148] Dysregulation of cellular adhesion in diffuse gastric cancer is thus associated with a worse disease prognosis and increased metastasis.[30]

Molecular classification has allocated loss of E-cadherin primarily with the Genomically Stable (GS) or microsatellite stable/epithelial-mesenchymal transition (MSS/EMT) variant of gastric cancer.[26,82] Both the TCGA and ACRG found that these subtypes of gastric cancer possess a poor prognosis. The use of E-cadherin as both a prognostic and predictive target is intriguing given the importance of this molecule in the pathogenesis of gastric cancer.

1.8.2.2 Galectin-3

Galectin-3 is a 30 kDa protein encoded by the LGALS3 gene located on chromosome 14q21-22.[149] Galectin-3 belongs to a large family of lectin proteins, which all contain a carbohydrate recognition domain that binds β-galactoside. Molecular studies show that Galectin-3 specifically binds with IgE and that IgE binding is inhibited by galactose or lactose.[150] The role of Galectin-3 in cancer is heterogenous and is dependent on cancer type and subcellular localization.[151] For example, the function of Galectin-3 is known to vary depending on whether it expressed in the nuclei, cytoplasm, cell surface or secreted as an extracellular protein.[152–154] This molecule has been implicated in a wide range of cellular processes including cell adhesion, apoptosis in immune cells, angiogenesis, tumorigenesis and inhibition of apoptosis in tumor cells.[155–157] As a targetable protein, anti-Galectin-3 drugs already exist in pre-clinical or clinical trials.[158,159]

Galectin-3 is an intriguing biomarker due to its surface expression on macrophages and its association with M2 macrophage function and migration.[160–162] Briefly, one mechanism of the innate immune response involves the migration of monocytes to a tumour. These monocytes undergo transformation to form tumour-associated macrophages (TAMs) that function as antigen presenting cells and provide regulation of the immune system via secretion of cytokines.[163] Two principle TAM subpopulations have been characterized. The M1 macrophage exerts anti-tumour effects via tumour phagocytosis, antigen presentation and secretion of pro-inflammatory cytokines such as Tumour Necrosis Factor and

interleukin-6. The M2 macrophage promotes an immunosuppressed tumour microenvironment, but also facilitates tumour angiogenesis and invasion.[164] In mouse models, increased Galectin-3 has been associated with the presence of M2 macrophages and subsequent tumour growth and angiogenesis.[161,162]

The significance of Galectin-3 in gastric cancer is currently under debate. Galectin-3 has been suggested to contribute to the carcinogenic effect of *H. pylori* infection by prolonging cell survival and affecting intracellular signalling. A meta-analysis published in 2018 consisting of eight studies involving 2093 patients with gastric cancer found that low expression of Galectin-3 was associated with a poor prognosis.[165] Furthermore, low levels of Galectin-3 were associated with lymphatic invasion, greater tumour depth, poorer tumour grade and TNM stage. Of the eight included studies only 2 occurred in populations outside of Asia (United States and Brazil). Due to the variety of mechanisms and unknown predictive value, Galectin-3 remains a molecule of interest that requires additional characterization in gastric cancer.

1.8.2.3 CD4 and CD8

The two main populations of T lymphocytes are identified by their respective T-cell co-receptor complexes CD4 or CD8. A $CD4^+$ T cell generally functions to modulate and coordinate the immune response. $CD4^+$ cells interact with peptide antigens presented on Major Histocompatibility Complex Class II receptors (MHC-II) by the antigen presenting cells (APCs). Depending on the peptide presented and the current cytokine milieu, $CD4^+$ cells modulate their function into distinct subpopulations such as T-regulatory (T_{reg}), Type 1 Helper (T_h1) and Type 2 Helper T cells (T_h2). The primary effect of $CD4^+$ cells is exerted by the secretion of numerous cytokines which may facilitate the recruitment of immune cells or regulate the activation or suppression of effector cells such as Natural Killer cells or $CD8^+$ cells.

Cytotoxic T cells are identified by the presence of the CD8 membrane receptor. $CD8^+$ cells facilitate direct cell death via the introduction of cytotoxins into cells that express foreign peptide on MHC Class I (MHC-I). In a meta-analysis, increasing $CD8^+$ TILs was found to be associated with improved overall survival and an improved anti-tumour response in the setting of immune checkpoint inhibitor therapy.[166]

Avoiding immune destruction is a Hallmark of Cancer.[167] Cancer evade cytotoxic T lymphocytes decreasing antigen presentation via downregulation of MHC-I.[168] This

adaptation prevents cytotoxic T cell engagement with tumour cells due to their obligate requirement of antigen presentation on MHC-I. Evasion of growth suppression via p53 mutations may also decrease $CD8^+$ T cell infiltration.[169,170] Cancer expression of immune co-inhibitory molecules such as PD-L1 and CTLA-4 promote T cell exhaustion and encourage immunosuppressive responses from T_{reg} cells.[168,171]

Given the complexity of tumour immune response, IHC may not be well suited to characterize the milieu of immune subpopulations. However, at a simple level, we can leverage an accurate quantification of modulatory $CD4^+$ cells and cytotoxic $CD8^+$ cells. The success of the Immunoscore in colon cancer suggests that the characterization of TILs in gastric cancer is a useful endeavour. To date, the utility of TILs in predicting tumour response to neoadjuvant chemotherapy in gastric cancer has not been investigated.

1.8.3 Three-dimensional cell culture (Organoid culture)

1.8.3.1 Organoid culture - a valuable translational tool

Traditional cell culture technique involves the growth of immortalized cancer cell lines on a 2-dimensional (2D) petri dish. This technique provides a streamlined and reproducible method for *in-vitro* interrogation of cancer biology. Criticism of 2D cell culture mainly centres around its limited fidelity in recapitulating *in-vivo* tumour characteristics and drug response.[172–174] Alternatives to 2D cell culture include xenograft models or patient-derived xenograft models, in which human cancer cells from cell lines or primary tumours are grown within a mammalian model such as a mouse.[175] The development of xenograft models is costly and similar issues can arise in accurately extrapolating drug responses in animal models to the human patient.[176]

Three-dimensional (3D) tissue culture of primary tumour cells into "mini-organs" termed, organoids, is a promising technology to augment current pre-clinical models.[177] A variety of methods for organoid development exists, but most commonly involves the extraction of tumour cells via surgical resection, biopsy or needle biopsy.[174,178] The tissue is subsequently dissociated in enzymes to smaller groups of cells that are then placed within a basement membrane matrix gel. The basement membrane matrix provides a scaffold in which tumour cells are able to migrate, propagate and eventually form organized units of differentiated cells reflective of the *in-vivo* tissue lineages.[179]

Molecular analysis of tumour organoids demonstrates that gene expression more closely resembles *in-vivo* tissue compared to 2D cell culture.[180] Furthermore, organoids provide a

reliable method of successfully generating and propagating primary tumour cells compared to 2D techniques.[181] Organoids can be cryopreserved in biobanks and successfully re-established to facilitate ongoing scientific inquiry to a given population of organoids.[182–184]

Organoid culture utilizes endogenous stem cell populations which reside in differentiated tissues. For example, in gastric tissue, intestinal organoids can be developed from a single $LGR5^+$ gastric stem cell retrieved from the crypt of gastric epithelium.[185] The essential components of organoid culture media include R-spondin, Wnt and noggin.[186] These molecules facilitate the stimulation and growth of stem cell populations, including TROY, LGR5, and AXIN2, among others.[179,187] All gastric cell lineages including chief, parietal, enteroendocrine, gland mucous and pit mucous cells can be established in organoid culture.[179,188]

Beyond providing greater epithelial cell diversity compared to 2D cell culture, tumour organoids also provide recapitulation and propagation of stromal tissue and the immune microenvironment.[189] In their landmark paper, The Hallmarks of Cancer: The Next Generation, Hanahan and Weinberg illustrate the importance of the tumour microenvironment as a micro-organ that collectively works to augment cancer growth.[167] Signaling between Cancer-Associated Fibroblasts (CAFs), cancer stem cells, endothelial cells and pericytes facilitates the development of malignant disease. In particular, CAFs have been demonstrated to contribute to chronic inflammation, immunosuppressive microenvironments, angiogenesis, and tumour invasion and metastasis.[164,190,191]

1.8.3.2 Future perspectives for organoid culture

Significant work involving gastric cancer organoids has preceded this thesis. These studies have provided established protocols for growth, propagation and research of gastric organoids.[178–180,182,184,192,193] Given the infancy and relative complexity of organoid culture, significant questions remain to be answered. Although clustering has identified similarity in gene-expression between parent tumours and organoids, the extent to which an organoid recapitulates the parent tumour's defined molecular subtype is unknown. We also do not fully understand the proportion of organoids that adequately recapitulate their parent tumour. In Yan et al. over half of the parent-organoid pairs were removed from subsequent analyses due to failure to cluster with parent tumours.[182] This limitation was attributed to organoids grown from samples with poor tumour purity. Confirmation of cancer growth is often challenging with current methods. Establishing whether an organoid culture contains cancer growth or has been overgrown with normal organoids relies on microscopic

morphology, histology and immunofluorescence techniques.[194] Finally, organoid media and growth protocols may also facilitate the growth of some molecular subtypes over others. Thus, the documented difficulty in growing some diffuse-type gastric cancer organoids could be attributed to molecular subtypes or tumour purity issues.[182]

Given the high fidelity of patient-derived tumour organoids, researchers have evaluated their role as a method to assign personalized cancer treatment. Several studies have documented a correlation between *in-vitro* patient-derived organoid dose-response assays and clinical tumour response.[182,183,195,196] In 2019, Ganesh et al. demonstrated a significant correlation between *in-vitro* dose-response and progression-free survival from 7 rectal cancer organoids treated with 5-fluorouracil or FOLFOX (5-fluorouracil, leucovorin, oxaliplatin). These early studies suggested that organoids are a valuable tool to inform personalized medicine. However, in 2021, the first prospective tumour organoid feasibility trial called the Selecting Cancer Patients for Treatment Using Tumour Organoids (SENSOR) failed to provide evidence of feasibility and utility in allocating cancer therapy.[174] Ongoing studies including the TUMOROID trial (NL49002.031.14) and the STRONG trial (NCT03307538) will provide additional prospective insight into the role of organoid culture in translational clinical models.[181,197]

1.8.4 Molecular classification of gastric cancer

Next-generation sequencing (NGS) and high throughput molecular analysis of various cancers have provided insight into foundational biological mechanisms of cancer. The premier example is the PanCancer Atlas project, which contains open source multi-omic and clinical data from over 11,000 tumours among 33 cancer types. From a pan-cancer perspective, these data have improved our understanding of molecular relationships between unique cancer types, cancer stemness, candidate driver genes, tumour immune response and the role of various oncogenic signalling pathways.[198–202] Alternative perspectives of cancer from the organ and system level, such as gastrointestinal carcinoma, has also revealed essential knowledge that guides analysis in this thesis.[203] In addition to facilitating advancement in sequencing technology and data science, these projects ultimately serve to guide medical science to more effective anti-cancer therapy or diagnosis.[198]

The immense therapeutic and prognostic potential of molecular classification frameworks derived from multi-omics data is juxtaposed by several significant limitations. Next generation sequencing or genome microarray technology is immensely expensive. Colloquial

discussion in medical research suggests NGS costs only $1000 per patient. These estimates do not reflect the reality of direct or indirect costs. A microcosting study from the United Kingdom evaluated the cost of implementing genome sequencing in 399 patients with cancer or rare disease per year.[201] For cancer-related NGS work, the cost of genome sequencing of germline and tumour samples was approximately £6,840 ($8,413.20 USD) or £3,420 ($4,206.60 USD) per sample. The direct cost of sequencing was only 76% of the total cost with an additional ~17% or $1,189.56 USD related to bioinformatics, reporting and data archiving alone.

Current pipelines for routine cancer omics data acquisition, analysis and communication do not exist outside of highly specialized centres such as Memorial Sloan Kettering. Implementation of molecular classification to mainstream clinical use is hindered by technological and computational complexity. Ideally, the process from the measurement of molecular data to usable clinical information should be streamlined through the use of user-friendly and accessible software.

In addition to limitations related to cost and logistics, the tangible clinical utility of molecular classification in cancer is unclear. For example, there has been limited comparative analysis between different classification schemes. It is entirely possible that slightly different approaches to molecular classification have yielded biologically similar subtypes between separate frameworks. Furthermore, attempts to simplify genome-wide classifications to a more manageable scale are ongoing.[204,205] Here I describe an overview of current molecular classifications systems in gastric cancer. Later in this work we will attempt to mitigate some of the previously stated issues affecting the clinical implementation of molecular classification.

1.8.4.1 The Cancer Genome Atlas

In 2014, The Cancer Genome Atlas, a multidisciplinary National Institutes of Health-funded molecular classification project, was the first to provide a robust and novel classification for gastric adenocarcinoma. In this landmark study, four main subtypes of gastric cancer were identified: EBV-type (EBV), chromosomal instability (CIN), microsatellite instability (MSI) and genomically stable (GS).[82] Subsequent work has revealed the prognostic utility of this novel classification system.[83] In particular, MSI and EBV cancer is associated with superior survival compared to CIN. Genomically stable cancer, which is enriched in diffuse type gastric cancer, carries the worst prognosis.

TCGA classification first isolated EBV and MSI tumours by their respective gold-standard molecular tests. EBV tumours were identified by Epstein-Barr encoding region *in-situ* hybridization. MSI tumours were defined by the Bethesda criteria following capillary electrophoresis measurement of established microsatellite regions of genomic DNA. Tumour aneuploidy or somatic chromosomal copy number aberration (SCNA) subsequently allocated tumours to CIN (High SCNA) or GS (Low SCNA) tumours. Chromosomal instability was the most prevalent (50%) subtype among the 295 tumours, followed by MSI (22%), GS (20%) and EBV (9%).

TCGA class tumours arguably carry the most evidence for clinical utility. In a Phase II trial 61 metastatic gastric cancer patients were treated with pembrolizumab immune checkpoint inhibitor therapy.[206] Tumours were sequenced and TCGA subtypes were assigned. EBV and MSI tumours were found to have an overall response rate of 100% and 85.7%, respectively.

1.8.4.2 The Asian Cancer Research Group Classification

The Asian Cancer Research Group characterized 300 Asian gastric cancer tumours into 4 distinct groups: microsatellite instable (MSI), microsatellite stable with TP53 overexpression (MSS TP53+), microsatellite stable with TP53 under expression (MSS TP53-) and microsatellite stable with epithelial to mesenchymal transition (MSS EMT).[26] The ACRG group used whole transcriptome microarray as the foundation for their classification system as opposed to the multi-omics approach used by TCGA. Using mRNA expression data, they associated the major trends of variance via Principle Component Analysis to known gene-signature profiles. This technique identified EMT and MSI-type tumours. The allocation of MSS TP53+ and MSS TP53- tumours was subsequently divided using a TP53 activation signature.

The ACRG classification carries prognostic value. Similar to TCGA, the MSI subtype was associated with the most favourable prognosis followed by MSS TP53+ and MSS TP53-. EMT tumours not only carried the worst prognosis but were also associated with greater recurrence and rates of peritoneal carcinomatosis.

1.8.4.3 The Tumour Microenvironment Score

The ACRG and TCGA classification systems attempted to classify gastric cancer from a bulk tumour perspective. That is, they evaluated the molecular and genomic data as is without the specific intent of isolating separate tumour components. The Tumour Microenvironment

Score (TMEscore) was a concrete effort to isolate tumour immune signatures related to the tumour microenvironment in gastric cancer.[207]

Using ACRG mRNA expression data, researchers isolated distinct immune signatures through a variety of statistical techniques. These immune signatures were subsequently applied to a number of publicly available gastric cancer transcriptome datasets including TCGA. Tumours with an activated, anti-tumour immune response enriched in interferon-gamma were defined as TME High subtype. TME Low tumours exhibited an immunosuppressive microenvironment and enriched signatures in epithelial to mesenchymal transition and transforming growth factor beta (TGF-beta). In multivariable Cox regression, TMEscore was an independent prognostic factor and TME High was associated with significantly greater survival in gastric cancer.

To test the classification system, TMEscore was applied to cohorts of urothelial cancer and melanoma patients who received immunotherapy. In both cases, patients with TME High tumours experienced improved immunotherapy responses. In a limited analysis of the ACRG cohort, improved survival associated with TME High tumours was found to exist regardless of adjuvant chemotherapy status.

1.8.4.4 Other molecular classification schemes

Other classification systems for gastric cancer exist. In 2018, Oh et al. proposed a binary classification system consisting of Mesenchymal and Epithelial subtypes. Once again, these subtypes demonstrated prognostic value across multiple public datasets. Furthermore, laboratory work provided evidence supporting the potential use of Insulin-like growth factor (IGF1)/IGF1 receptor (IGF1R) inhibitors in mesenchymal type cancers.[208] In 2021, a molecular pathway-based approach using TCGA, ACRG and other public data suggested three molecular classes for gastric cancer.[209] Finally, molecular classification based on single-cell sequencing has also been proposed.[210]

1.9 Summary of included chapters and hypotheses

This thesis is comprised of 7 chapters. Chapters 2-5 are written as self-sufficient scientific publications. A comprehensive conclusion to the problems presented in Chapter 1 and results in Chapters 2-5 are presented in Chapter 6. All supplemental material is compiled in the Appendix with appropriate subheadings for navigation to each respective chapter's

supplement. Of note, readers of electronically distributed versions of this thesis will be able to navigate to Figures and Tables via in-text references highlighted by blue font colour.

In Chapter 2, we performed a prospective pilot study to evaluate the immunohistochemistry-based expression of Galectin-3, E-cadherin, CD4 and CD8 in Canadian gastric cancer patients. To address the lack of predictive biomarkers informing neoadjuvant treatment response we measured biomarker expression in a subset of patients treated with FLOT4 chemotherapy. In Chapter 2 we tested the following hypotheses:

1. The selected biomarkers will be differentially expressed in cancer and normal adjacent gastric tissue.
2. The expression of selected biomarkers will significantly increase or decrease in response to neoadjuvant chemotherapy in paired tumour samples
3. The selected biomarkers will be differentially expressed in pre-treatment tumours that achieve complete or near-complete tumour response compared to tumours with partial response, poor response or progression to metastasis following neoadjuvant chemotherapy.

Chapter 3 focused on the investigation of pre-clinical translational gastric organoid models. Specifically, we worked to address issues associated with adequate and reliable tissue procurement. We evaluated the feasibility of shipping fresh mouse stomach tissue at 4 degrees Celsius over 24 or 48 hours for the purpose of establishing organoid culture. The experimental conditions tested the optimal transport media (Hanks Balanced Salt solution (HBSS), University of Wisconsin (UW) solution or Histidine-Tryptophan-Ketoglutarate (HTK) solution) and transport times (Fresh, 24 hours or 48 hours). The optimal shipping conditions were also tested using human gastric cancer tissue. In Chapter 3 we assessed the following hypotheses:

1. The success rate of establishing an organoid culture for 10 passages will differ between:
    a. Fresh, 24-hour and 48-hour tissues
    b. HBSS, UW and HTK solutions
2. The viability of organoid culture over 10 passages will differ between:
    a. Fresh, 24-hour and 48-hour tissues
    b. HBSS, UW and HTK solutions
3. The growth rate of organoid culture over 10 passages will differ between:

    a. Fresh, 24-hour and 48-hour tissues
    b. HBSS, UW and HTK solutions
4. The gene expression of LGR5 and TROY stem cell markers will differ between:
    a. Fresh, 24-hour and 48-hour tissues
    b. HBSS, UW and HTK solutions
5. The *in-vitro* organoid dose-response to the anti-cancer drug FOLFOX (folinic acid, 5-fluorouracil, oxaliplatin) will differ between:
    a. Fresh, 24-hour and 48-hour tissues
    b. HBSS, UW and HTK solutions
    c. Passage number in paired organoid samples

Chapters 4 and 5 focused on the translation of gastric cancer molecular classification systems to the clinic. Chapter 4 mainly concerned the development and retrospective analysis of classification systems for personalized medicine in gastric cancer. In Chapter 5 we expanded the models developed in Chapter 4 to a gene expression-based test. Next, the value of our test was assessed in our patient population and a series of patient-derived organoid samples.

In Chapter 4, we established machine-learning models to predict molecular subtypes from the TCGA, ACRG and TME classifications using whole-transcriptome data (see: Molecular classification of gastric cancer). We used feature selection techniques to decrease the number of genes required to classify gastric cancer from over 16000 to approximately 50 genes. Next, we examined model diagnostics including accuracy, Brier score and calibration. Using our models, we applied molecular subtypes to 2,202 publicly available gastric cancer patients and retrospectively evaluated the prognostic, therapeutic and predictive utility of these classification systems. In contrast to prior research, we tested whether using classification systems as discrete categories or continuous scores that approximate the probability of molecular subtypes is useful. Using this perspective, we attempted to develop a method to assess heterogeneity of molecular subtype classification. In Chapter 4 we evaluated the following hypotheses/questions:

1. Can complex whole-genome/transcriptome molecular classification systems in gastric cancer be reproduced in supervised machine-learning models?
2. Can machine-learning classes be properly assigned to additional patient samples with gene expression measurements obtained from a separate platform/technology using

Feature Specific Quantile Normalization (ie. Can molecular subtypes be learned across multiple gene expression platforms?)
    a. Is the assignment of molecular subtypes across platforms internally valid with respect to gold-standard benchmark tests?
3. How can we assess the effectiveness of binary and multi-class classification models? Are current model diagnostic methods sufficient?
4. How can we assess differences/similarities between multiple classification systems?
5. Does our novel continuous heterogeneity metric, called *Variance Heterogeneity* provide significant prognostic or diagnostic utility?
6. In a propensity score matched cohort, does the integration of multiple classification systems in gastric cancer identify prognostic, therapeutic or predictive "biomarkers" for adjuvant chemotherapy treatment?
7. Can the integration of multiple classification systems in gastric cancer facilitate the development of a meaningful personalized individual survival distribution model?

Using the genes selected in our models from Chapter 4, we created a custom Nanostring codeset to assign TCGA and TME subtypes to our patient population. In Chapter 5 we were primarily focused on the feasibility of this test in assigning subtypes in formalin-fixed paraffin embedded tissue and whether the test could be applied to translational patient-derived organoid models. Also, of paramount importance was the preliminary external validation of molecular subtypes using gold-standard reference tests. To guide future inquiry, we assessed the landscape of molecular subtypes produced by our model in 2-dimensional immortal gastric cancer cell lines from the Cancer Cell Line Encyclopedia. With these data we evaluated how well currently available cell lines approximate subtype probability scores in 2,202 clinical gastric cancer samples and performed a drug-sensitivity analysis according to molecular subtypes scores. In Chapter 5 we asked the following questions:

1. Can we develop patient-derived gastric cancer organoids from endoscopic biopsy or surgical specimens in our population?
    a. Do our organoids recapitulate gastric tissue?
2. Is our custom Nanostring codeset feasible for future study?
    a. Can we reliably isolate nucleic acid of sufficient quality from formalin-fixed paraffin embedded tissue samples for the purpose of a Nanostring gene expression test?

    b. Does our Nanostring test approximate gold-standard classification of matched parent tumour and tumour organoid samples?
    c. What is the optimal normalization procedure for allocating subtype scores to Nanostring-based gene expression data?
3. Are patient-derived gastric cancer organoids useful for personalized medicine?
    a. Using gold-standard tests and our Nanostring test, how well do patient-derived organoids recapitulate molecular subtypes of the parent tumour?
    b. Are patient-derived organoids feasible for translational dose-response assays?
    c. Does *in-vitro* organoid dose-response associate with molecular subtype scores?
4. What is the preliminary distribution of molecular subtypes in our patient population?
    a. Does the association of molecular subtype scores with basic tumour histology in our population approximate associations observed in an external cohort of 2,202 gastric cancer patients?
5. Is 2-dimensional *in-vitro* culture of immortalized cell lines useful for modern personalized medicine research?
    a. Do currently available 2-dimensional gastric cancer cell line molecular subtype scores approximate an external cohort of 2,202 gastric cancer patients?
    b. Do molecular subtype scores in cancer cell lines associate with dose-response in targeted anti-cancer therapy?

Chapter 2: Can immunohistochemistry-based biomarkers provide predictive value in estimating tumour response to neoadjuvant FLOT in gastric cancer? A prospective pilot study

2.1 Introduction

Immunohistochemistry (IHC) is a proven molecular pathology technique with a record of providing prognostic and therapeutic biomarkers in oncology. In gastric cancer, prominent IHC-based biomarkers may be prognostic or therapeutic as in the case of E-cadherin and HER2, respectively.[121,211] However, there is a lack of predictive biomarkers to inform treatment response for neoadjuvant therapies.

In North America, the standard of care for locally advanced gastric cancer is neoadjuvant chemotherapy with FLOT4.[103] Despite advances in chemotherapy regimens, we currently possess limited information on which patient will receive optimal treatment benefit. Neoadjuvant chemotherapy is known to cause adverse events, including hospitalization in approximately 25% of patients. Prior evidence suggests that, in the absence of treatment response, neoadjuvant chemotherapy in gastric cancer contributes to adverse events without additional benefit compared to adjuvant treatment or surgery alone.[212,213]

Here we investigate a panel of biomarkers that we hypothesize may provide value in predicting tumour response. Galectin-3 is a lectin protein that facilitates cancer tumorigenesis and prognosis.[151,155–157] Pre-clinical models suggest that increased Galectin-3 expression is associated with chemotherapy resistance.[214,215] Recent work has implicated cell-surface expression of Galectin-3 with chemoresistance in gastrointestinal cancer stem cells.[216] E-cadherin is a cell-cell adhesion molecule that plays an important role in gastric cancer development, classification and prognosis.[26,30,82] In breast cancer, E-cadherin expression has been associated with chemotherapy response.[217,218] We also assess whether $CD4^+$ and $CD8^+$ tumour infiltrating lymphocytes (TILs) and the relative proportion of these cells influence neoadjuvant chemotherapy response. $CD4^+$ T cells function to regulate and potentiate cell-mediated immunity, while $CD8^+$ T cells perform tumour cell lysis. Increasing evidence recognises the association of greater TILs to favourable cancer prognosis and chemotherapy response in colon and gastric cancer.[139,140,219–222] To date, no studies have investigated the role $CD4^+$ or $CD8^+$ TILs in neoadjuvant chemotherapy response.

To guide future studies, we performed a prospective pilot study to evaluate if these selected biomarkers provide predictive value in evaluating treatment response following neoadjuvant FLOT chemotherapy.

2.2 Methods

2.2.1 Study design

We performed this single-center, prospective pilot study at the University of Alberta in Edmonton, Alberta, Canada from January 2018 to January 2022. All human clinical participants consented according to the approved ethics protocol granted by the Health Research Ethics Board of Alberta (Study ID: HREBA.CC-17-0228). Treatment naïve Stage I-IV sporadic gastric adenocarcinoma patients aged greater than 18 years were included. A subset of patients enrolled was allocated to a second cohort on the basis of receiving curative intent neoadjuvant FLOT chemotherapy. Patients with a known inherited oncogenic germline mutation or hereditary syndrome (ie. Familial Adenomatous Polyposis) were excluded.

Specimens were retrieved via endoscopic biopsy at the time of diagnosis, screening laparoscopy or at the time of surgical resection at the Walter C Mackenzie Health Sciences Centre or Royal Alexandra Hospital. The initial study protocol retrieved two tissue biopsies for permanent pathology, however, following interim review four biopsies were retrieved thereafter. The presence of cancer in specimens was confirmed by a gastrointestinal pathologist. In the absence of cancer, clinical formalin-fixed paraffin-embedded pathology blocks were retrieved when available. In clinical samples with treatment effect, residual cancer cells were detected using anti-pan cytokeratin (Abcam, clone C-11, ab7753) immunohistochemistry staining followed by the manual assembly of tissue microarray (TMA) blocks with 4mm cores of regions containing residual tumour.

Our primary outcome for all patients was the difference in expression of selected biomarkers between normal and cancer tissue. In the subgroup of patients receiving neoadjuvant chemotherapy, our primary outcome was the difference in expression between tumour treatment response and incomplete treatment response. We also evaluated the difference in expression of biomarkers in paired samples before and after chemotherapy treatment.

Treatment response was retrieved from clinical pathology reports. The Tumour Regression Score was graded according to the College of American Pathologists and National Comprehensive Cancer Network protocol on a 4-point scale (0 = Complete response, 1 =

near complete response, 2 = partial response, 3 = poor or no response).[103] In accordance with prior studies, treatment response was expressed as a binary variable consisting of response and incomplete response categories.[223] Responsive tumours included complete and near-complete responses, whereas incomplete responses included partial, and poor no response. Patients who progressed to metastasis while receiving neoadjuvant treatment were classified as an incomplete response.

2.2.2 Immunohistochemistry

Tissue specimens of normal and cancer tissue were fixed in zinc-formalin (Z-Fixx) for 24 hours, washed three times and stored in 70% ethanol prior to preservation in paraffin. Briefly, 4 µm tissue sections were deparaffinized in Histoclear (National Diagnostics) and rehydrated. Endogenous peroxidases were quenched using 3% hydrogen peroxide in methanol for 5 minutes. Microwave heat induced epitope retrieval was performed using Sodium Citrate (pH 6, heated to 94 degrees Celsius in 1-minute intervals followed by 9 minutes continuous heat) for E-cadherin and Tris-EDTA (pH 9, heated to 94 degrees Celsius in 1-minute intervals followed by 8 minutes 30 seconds continuous heat) for CD4 and CD8. Non-specific staining was blocked using 20% normal goat serum (Jackson Laboratories) for E-cadherin, CD4 and CD8 or 2% Fetal Bovine Serum (Gibco) in 1X phosphate buffered saline for Galectin-3 for 20 minutes followed by avidin and biotin blocking (Vector Laboratories, SP-2001) per manufacturer's protocol. Tissue sections were stained with primary antibodies anti-E-cadherin (1:25, 1.5 hours room temperature, ThermoFisher Scientific, clone 4A2C7, 33-4000), anti-Galectin-3 (1:200, 30 minutes room temperature, Cedarlane, clone M3/38, CL8942AP), anti-CD4 (1:200, overnight at 4 degrees Celsius, Abcam, clone EPR6855, ab133616) or anti-CD8 alpha (1:200, overnight at 4 degrees Celsius, Abcam, ab4055). All biotinylated IgG secondary antibodies were incubated at 1:200 for 30 minutes at room temperature, including rabbit anti-rat for Galectin-3 (Vector Laboratories, BA-4001), goat-anti-rabbit for CD4 (Vector Laboratories, BA-1000) and goat-anti-mouse for E-cadherin and CD8 (Jackson ImmunoResearch, 115-065-003). Antibody detection was performed using avidin-biotin complex/horseradish peroxidase (Vector Laboratories) and 3,3-diaminobenzidine tetrahydrochloride (DAB, Abcam, ab64238) per manufacturer's protocol. Stained tissue sections for E-Cadherin, CD4 and CD8 were counterstained with Harris' hematoxylin (Fisher Scientific) and Harris' hematoxylin and eosin (Fisher Scientific) for Galectin-3.

2.2.3 Histology imaging and quantification

Histology images were captured using a Leica Aperio CS2 digital slide scanner. Digital pathology quantification of antibody expression was performed using QuPath version 0.3.1 (Figure 2-1C).[224] Briefly, digital images were uploaded and the tumour and immediate tumour-host interface were annotated as a single region of interest. Stain vectors were estimated using default settings for each sample. For CD4, CD8 and Galectin-3, positive cells were detected using default nucleus DAB optical density settings. The CD4/CD8 ratio was calculated as the proportion of positively stained CD4 cells divided by the proportion of positively stained CD8 cells. For E-cadherin, both the proportion of positive cells and H-score was calculated. Annotated cell regions were assessed for accuracy and in the event of background or non-specific staining positive cell threshold values were adjusted to reflect true positive staining. The H-score provides a consensus scoring method for evaluating immunostaining across a gradient of intensity (Equation 2-1). As defined in McClelland et al., H, M and L denotes high, medium and low intensity staining. Cells without staining are denoted N for negative staining.[225]

Equation 2-1

$$H - score = 100 \times \frac{3H + 2M + L}{H + M + L + N}$$

2.2.4 Statistical analysis

Statistical analyses were completed using R version 4.1.2.[226] Differences between groups were assessed with a Wilcoxon two-sample test for independent samples and two-tailed paired Wilcoxon test for paired samples. Statistical significance was defined at alpha = 0.05. Multiple comparisons corrections were not made for our main outcomes given our prespecified analyses, but the possibility of false positive results is noted. Summary of continuous variables is expressed as median with interquartile range (IQR). Categorical variables are expressed as absolute number of cases and percent proportions.

The ability of biomarkers to predict treatment response was assessed using the caret package in R.[227] Briefly, out-of-sample resampling accuracy was estimated for each biomarker as well as the combination of all biomarkers using 1000 bootstraps with replacement. Continuous variables were centered and scaled. Logistic regression models were used for single biomarker estimates and a regularized ElasticNet model implemented

in glmnet was used for estimates containing all biomarkers.[228] Model significance was tested using a one-sided binomial test comparing the estimated model accuracy to the No Information Rate (NIR). The NIR is defined as the largest proportion of observed classes, or the maximum accuracy of a classifier if it predicted the majority class every time.

Sample size calculations were performed using the MKpower package in R. Two-sample Wilcoxon distributions were generated using the mean and standard deviation from our pilot study sample. The normality of the distribution for each biomarker's expression levels was confirmed with a Shapiro-Wilk test. Random sampling from a truncated normal distribution constrained between 0 and 100 was performed for a series of samples sizes ranging from 10 to 120, in intervals of 10. The empirical power (beta) for each sample size was calculated using Monte-Carlo simulations with 1000 iterations for a specified type-I error rate (alpha) = 0.05.

2.3 Results

2.3.1 Patient demographics

Fifty-three patients were consented for this pilot study. Ten patients were excluded: one patient was diagnosed with Familial Adenomatous Polyposis, one was found to have neuroendocrine tumour pathology, one gastroesophageal junction tumour received alternative neoadjuvant therapy and seven patients were excluded due to inadequate tissue biopsies. Of note, an interim analysis of our protocol after enrolling the first 20 patients determined a biopsy accuracy rate of 60% for treatment naïve specimens and 25% for biopsy following neoadjuvant treatment. This prompted a change in study protocol to retrieve 4-8 tissue biopsies per sample.

A total of 43 patients were available for analysis, of which 18 (42%) underwent neoadjuvant chemotherapy during our study period. Baseline demographics are included in Table 2-1. Median age was 65 (60, 75) and the majority of patients were male (70%). Tumour pathology was represented by all TNM stages but a preponderance of high grade (72%), proximal stomach (60%) and diffuse type (63%) tumours were present. *H. pylori* status was available for 32 patients, of which the majority were negative (69%) and one was previously treated. Total gastrectomy was performed in nearly half of all patients and comprised 59% of all surgical resections.

Table 2-1 Baseline Demographics

Characteristic	N = 43[1]	n/N (Missing %)
Age (Years)	65 (60, 75)	43 / 43 (0%)
Sex		43 / 43 (0%)
Female	13 (30%)	
Male	30 (70%)	
Stage		43 / 43 (0%)
I	11 (26%)	
II	10 (23%)	
III	8 (19%)	
IV	14 (33%)	
Grade		43 / 43 (0%)
G1	1 (2.3%)	
G2	10 (23%)	
G3	31 (72%)	
Gx	1 (2.3%)	
Tumour Location		43 / 43 (0%)
Distal	14 (33%)	
Proximal	26 (60%)	
Whole stomach	3 (7.0%)	
Lauren Classification		41 / 43 (4.7%)
Diffuse	26 (63%)	
Intestinal	13 (32%)	
Mixed	2 (4.9%)	
Signet Ring Cell (Present)	26 (63%)	41 / 43 (4.7%)
H. pylori history		32 / 43 (26%)
Negative	22 (69%)	
Positive	9 (28%)	
Treated	1 (3.1%)	
Smoker		40 / 43 (7.0%)
Yes	9 (22%)	
No	14 (35%)	
Ex	17 (42%)	
Smoker (Pack Years)	7 (0, 32)	39 / 43 (9.3%)
Surgery		43 / 43 (0%)
Total Gastrectomy	20 (47%)	
Distal Gastrectomy	14 (33%)	
No resection	9 (21%)	
Neoadjuvant Chemotherapy	18 (42%)	43 / 43 (0%)
CD4/CD8 Ratio (% Positive)	1.7 (1.2, 2.8)	42 / 43 (2.3%)
CD4 (% Positive)	14 (7, 24)	43 / 43 (0%)
CD8 (% Positive)	8 (5, 11)	42 / 43 (2.3%)
Galectin-3 (% Positive)	46 (30, 57)	43 / 43 (0%)
E-cadherin (% Positive)	18 (6, 28)	43 / 43 (0%)
E-cadherin H-score	22 (7, 40)	43 / 43 (0%)

[1]Median (IQR); n (%)

2.3.1 Expression of biomarkers in normal and cancer tissues

Representative images of each immunohistochemistry stain within the 75th and 25th percentile of expression is presented in Figure 2-1A. Staining for E-cadherin was only

identified on cell membranes of gastric epithelium. Galectin-3 exhibited heterogeneous staining and was identified in nuclei, cytoplasm, and surrounding tumour stroma. The presence of Galectin-3 was often sporadic with distinct regions representing intense positive stain followed by fairly abrupt transition to moderate positivity. CD4 and CD8 positive staining was identified on the cell membrane of lymphocytes.

Galectin-3 was the most abundant molecule with a median expression of 46 (30, 57), followed by E-cadherin, CD4 and CD8 (Table 2-1). The E-cadherin H-score (median 22 (7, 40)) closely approximated the proportion of E-cadherin positive cells (median 18 (6, 28)). Greater H-score values in the upper quartile reflects the presence of high staining intensity in positive cells.

Significantly increased expression of CD4, Galectin-3 and CD4/CD8 Ratio was identified in cancer tissue relative to normal adjacent tissue controls (Wilcoxon, $p=0.035$, $p=0.020$ and $p=0.018$ respectively) (Figure 2-1B). The distribution of IHC scores between normal and cancer tissue for CD4 and Galectin-3 was relatively uniform, whereas differences in CD4/CD8 Ratios were dominated by sample outliers with large cancer IHC scores. In agreement with historical study, E-cadherin positivity and H-score was significant decreased in cancer tissue relative to normal. (Wilcoxon, $p<0.0001$ and $p<0.001$, respectively).

There were no statistically significant associations between relevant clinicopathologic factors and the expression of any biomarker for stage, lymphovascular invasion, perineural invasion, carcinomatosis, tumour grade or location (Appendix .B.1). The proportion of E-cadherin positive cells was significantly different according to Lauren Class, with relatively fewer positive cells present in diffuse and mixed type cancers (Kruskal-Wallis, $p = 0.043$).

2.3.1 Association of biomarker expression with exposure to neoadjuvant chemotherapy

We compared the expression of biomarkers in 15 paired tumour samples from the same patient before and after neoadjuvant FLOT to evaluate the effect of treatment on biomarker expression. All pre-treatment specimens were obtained by endoscopic biopsy and thus were restricted mainly to the mucosa and lamina propria. The majority of post-treatment samples were analyzed as TMA cores from surgical resection specimens (TMA cores = 87% vs. biopsy = 13%) in which residual tumour was present in mucosa, submucosa and muscularis.

Table 2-2 Clinicopathologic factors according to treatment response

Characteristic	Incomplete response, N = 14[1]	Response, N = 4[1]	p-value[2]
Age (Years)	60 (57, 63)	60 (52, 67)	>0.9
Sex			0.3
F	5 (36%)	0 (0%)	
M	9 (64%)	4 (100%)	
Stage			0.6
I	2 (14%)	2 (50%)	
II	5 (36%)	1 (25%)	
III	6 (43%)	1 (25%)	
IV	1 (7.1%)	0 (0%)	
Grade			0.6
G1	1 (7.1%)	0 (0%)	
G2	2 (14%)	2 (50%)	
G3	10 (71%)	2 (50%)	
Gx	1 (7.1%)	0 (0%)	
Tumour Location			>0.9
Distal	3 (21%)	1 (25%)	
Proximal	10 (71%)	3 (75%)	
Whole stomach	1 (7.1%)	0 (0%)	
Lauren Classification			0.5
Diffuse	10 (71%)	1 (33%)	
Intestinal	4 (29%)	2 (67%)	
Signet Ring Cell (Present)	9 (64%)	1 (33%)	0.5
H. pylori history			>0.9
Negative	8 (57%)	2 (50%)	
Positive	3 (21%)	1 (25%)	
Unknown	3 (21%)	1 (25%)	
Smoker			>0.9
Yes	4 (33%)	2 (50%)	
No	4 (33%)	1 (25%)	
Ex	4 (33%)	1 (25%)	
Smoker (Pack Years)	13 (0, 40)	36 (25, 42)	0.5
Surgery			>0.9
Total Gastrectomy	10 (71%)	3 (75%)	
Distal Gastrectomy	3 (21%)	1 (25%)	
No resection	1 (7.1%)	0 (0%)	

[1]Median (IQR); n (%)
[2]Wilcoxon rank sum test; Fisher's exact test

We found significantly increased association of tumour cells with $CD4^+$ and $CD8^+$ TILs following neoadjuvant chemotherapy (Paired Wilcoxon, p=0.002 and p = 0.008, respectively) (Figure 2-2A). In contrast, E-cadherin positivity and H-score significantly decreased in post-treatment samples (Paired Wilcoxon, p=0.035 and p=0.04, respectively). This was likely in part due to differences in tumour cell depth of invasion between pre-treatment biopsy and post-treatment TMA cores. CD4/CD8 Ratio expression remained relatively stable within samples except for one patient (orange line, Figure 2-2A, top left).

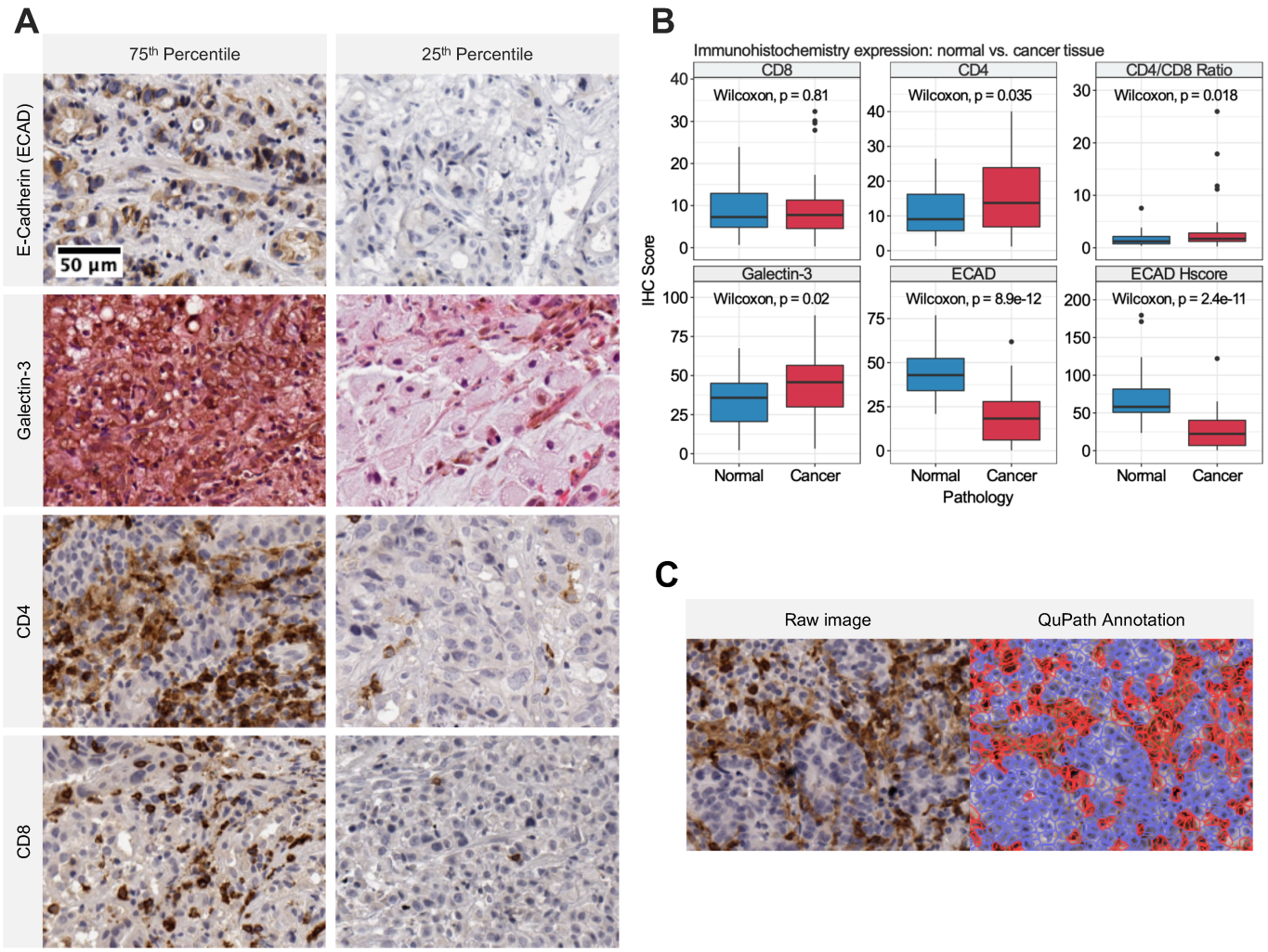
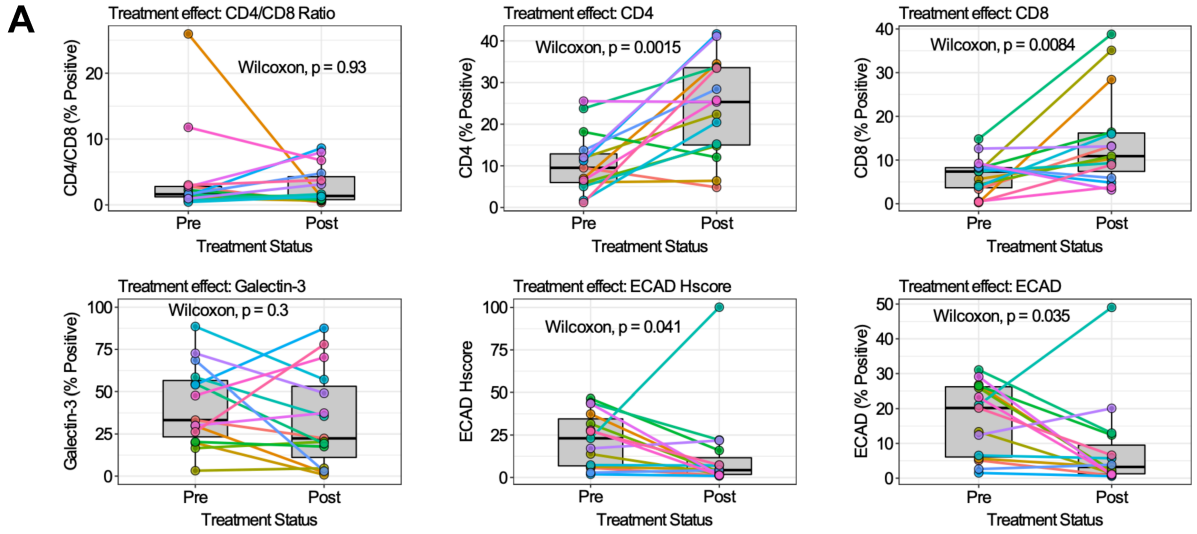
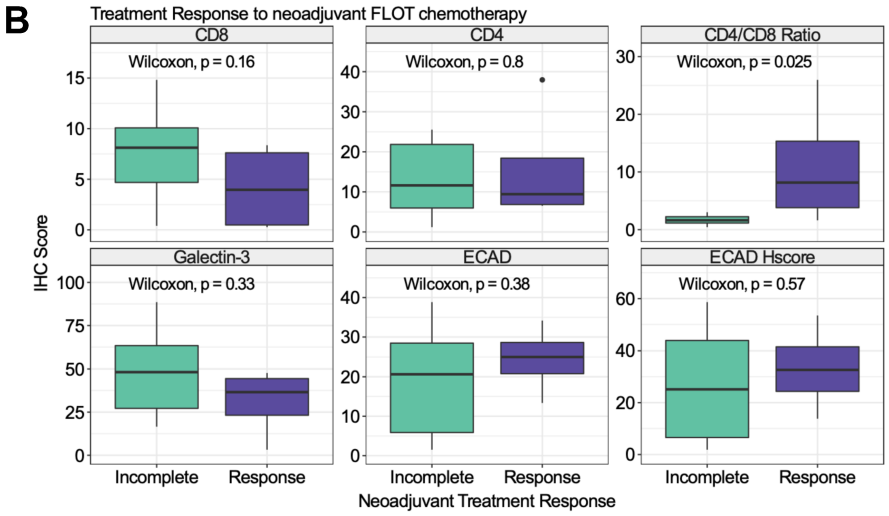


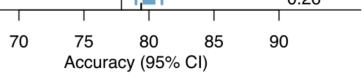
Figure 2-1 Immunohistochemistry stains and expression of biomarkers in treatment naïve normal and cancer tissue. **A.** Representative IHC images taken at 10X magnification for each respective biomarker identified on the y-axis. Images of expression values within the 75th and 25th percentile are presented in the left and right columns, respectively. **B.** Boxplot comparison of expression for each respective biomarker in treatment naive normal and cancer tissue. The IHC biomarker is labeled on the heading of each graph. The y-axis represents IHC score, which is the percent of positive stained cells for Galectin-3, CD4, CD8 and ECAD and the H-score for ECAD H-score plot. The x-axis labels the distribution corresponding to normal (blue) and cancer (red) tissue. The raw p-value for Wilcoxon tests is annotated in each panel. **C.** Representative images of QuPath digital pathology annotation using CD4 IHC at 10X magnification. Raw images (left) are processed and regions of interest are identified according to our methods. The annotated image (right) demonstrates the calculation of positive stained cells (red) and negative cells (blue).

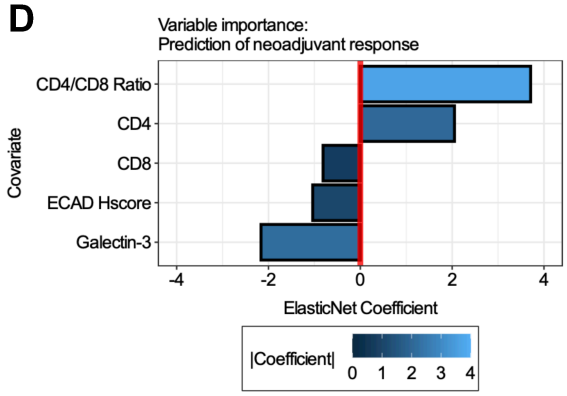




C

Model	Variables	Sensitivity (%)	Specificity (%)	Accuracy (95% CI)	P-Value
1	~ CD4/CD8 Ratio	98	44	86 (85, 87)	<0.001***
2	~ CD4	94	0	73 (72, 74)	1.00
3	~ CD8	90	16	74 (73, 75)	1.00
4	~ Galectin-3	91	6	72 (71, 73)	1.00
5	~ ECAD Hscore	89	1	70 (69, 71)	1.00
6	~ ECAD	88	3	70 (68, 71)	1.00
7	~ All IHC Markers	92	31	80 (79, 81)	0.26





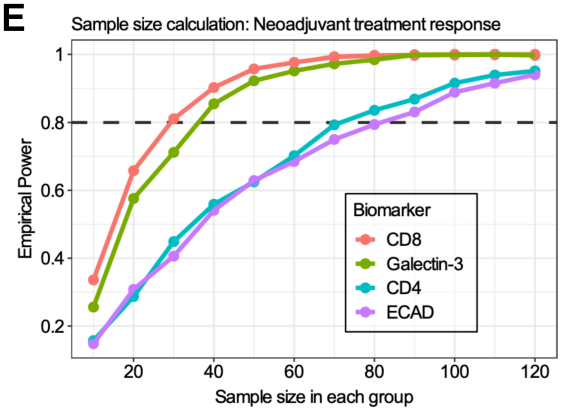


Figure 2-2 Association between biomarker expression and neoadjuvant FLOT4 chemotherapy. **A.** Paired boxplots for biomarker expression pre- and post-neoadjuvant chemotherapy. Each coloured point and line correspond to a single patient. Boxplots in grey represent the distribution of expression for all patients before and after chemotherapy. Paired Wilcoxon p-value is present in each plot. **B.** Boxplot comparison of expression for each respective biomarker between treatment response (purple) and incomplete response (green). The IHC biomarker is labeled on the heading of each graph. The y-axis represents IHC score, which is the percent of positive stained cells for Galectin-3, CD4, CD8 and ECAD and the H-score for ECAD H-score plot. **C.** Forest plot for metrics of 7 predictive models. The plot represents the out-of-sample accuracy and 95% confidence intervals for models estimated from 1000 bootstraps with replacement. The no information rate is shown by the solid and dotted vertical lines for univariable and multivariable models, respectively. **D.** Barplot of model coefficients from multivariable glmnet model. The y-axis represents model covariates and the x-axis the coefficient value. Treatment response is related to increasing covariate values or decreasing covariate values for positive and negative coefficients, respectively. The absolute importance of the coefficient is shown in blue according to the scale legend. **E.** Lineplot illustrating monte-carlo simulations for two-sample Wilcoxon sample size calculations. The y-axis is the empirical power and the x-axis is the sample size in each group. Each coloured line corresponds to a biomarker labelled according to the legend.

2.3.2 High CD4/CD8 ratio is associated with treatment response

Figure 2-2B outlines the relationship of biomarker expression to treatment response between pre- and post-treatment cancer specimens. For all analyses, we observed incomplete response in 14 patients (Partial = 9, Poor or No = 4, Progression to metastasis=1) and response in 4 patients (Complete = 1, Near Complete = 3). Statistically greater CD4/CD8 Ratios were observed in pre-treatment cancer biopsies compared to incomplete responders (Wilcoxon, p=0.025). Clinicopathologic characteristics were similar between treatment response groups (Table 2-2).

Next, we explored the utility of individual biomarkers (Models 1-6) and the combination of all biomarkers (Model 7) in predicting treatment response scores (Figure 2-2C). Given the small sample size and events per variable, we used out-of-sample estimates from 1000 bootstraps to limit bias by favouring pessimistic estimates of model accuracy. In this dataset, all biomarkers were effective at predicting incomplete tumour response (Sensitivity range 88-98%) but suffered from poor specificity (range 0-44%). CD4/CD8 Ratio was the only variable that provided significant model performance (Accuracy > NIR, one-sided binomial, $p<0.001$). The ElasticNet model using CD4/CD8 Ratio, CD4, CD8, Galectin-3 and

E-cadherin H-score as independent variables provided a mean accuracy greater than the NIR but failed to achieve statistical significance (p=0.26).

The optimal glmnet model provided coefficients for all variables despite tuning parameters allowing for L2 regularization (alpha = 0). To guide future studies, we evaluated the contribution of all biomarker variables to the predictive model using the final regularized ElasticNet coefficients (Figure 2-2D). The absolute value of coefficients found CD4/CD8 Ratio and Galectin-3 to provide the greatest influence in predicting favourable tumour response. Specifically, tumour response was associated with increasing CD4/CD8 Ratio and decreasing Galectin-3, respectively.

2.3.1 Sample size calculations

To inform future studies we performed sample size calculations using our pilot study sample distributions. In particular, we were interested in identifying the sample sizes required to evaluate the utility of biomarkers in explaining tumour response using a two-sample Wilcoxon test. In, Figure 2-2E we observe that CD8 and Galectin-3 require similar sample sizes of 30 and ~35 in each treatment response group to achieve adequate power. The relationship between sample size and empirical power was nearly identical for CD4 and E-cadherin, which were calculated to require ~70 and 80 samples in each group, respectively.

2.4 Discussion

Neoadjuvant chemotherapy is the standard of care for curative-intent treatment in western populations. Advantages of neoadjuvant chemotherapy include improved survival compared to surgery alone, greater R0 resection and reduction in nodal stage.[110,213,229,230] Irrespective of efficacy, cytotoxic chemotherapy is also associated with adverse events including peripheral neuropathy, neutropenia, infection or death.[112] Previous study has demonstrated that pathologic complete response (pCR) or treatment response defined as Tumour Regression Grade 1-3 is significantly associated with prognosis.[223,231] However, pCR occurs in only 3-15% of cases and complete or partial response in approximately 40% of patients.[213,223,232] In the absence of treatment response, neoadjuvant chemotherapy provides no additional benefit to surgical resection alone and still exposes the patient to adverse events.[232] Thus, to improve outcomes it is of paramount importance to identify clinicopathologic or molecular biomarkers to identify treatment responders.

Several potentially useful approaches for predicting treatment response have been recognized. Clinical or pathologic factors including age, tumour grade, signet cell pathology, serum carcinoembryonic antigen, various circulating lymphocyte populations and tumour size are significant predictors of tumour response.[233–235] The majority of predictive tumour biomarker research in gastric cancer has focused on identifying molecules associated with adjuvant chemotherapy response. For example, a multivariable model utilizing the measurement of several TIL populations in 879 patients provided 3-year survival prediction accuracies of 79 and 84% for surgery alone and adjuvant chemotherapy populations, respectively.[139] In the neoadjuvant setting, a post-hoc analysis of 83 patients in the COMPASS trial identified several candidate gene expression based-biomarkers such as *TIMP1* and *DSG2* using qRT-PCR.[236] Other studies to identify treatment response have used microRNAs, exosomes, inflammatory markers or medical imaging data.[237] Although predictive and prognostic factors identified in these studies show promise, there is limited external validity of these studies and clinical implementation is yet to be achieved.

In this pilot study, we present the utility of IHC-based expression of Galectin-3, E-cadherin, CD4 and CD8 in predicting treatment response to the neoadjuvant chemotherapy regimen FLOT4. First, we establish that Galectin-3, CD4, E-cadherin and the CD4/CD8 Ratio expression are significantly different between cancer and normal adjacent tissue. These findings suggest that these markers are intrinsic to the tumour or tumour microenvironment and thus may provide prognostic or predictive yield. Next, we establish that the CD4/CD8 Ratio is significantly greater in tumours with complete or partial response to neoadjuvant chemotherapy. In preliminary univariate and multivariate machine learning models, the CD4/CD8 Ratio was the only significant predictive marker of treatment response with an accuracy of 86%. Finally, we demonstrate that the tumour-specific expression of CD4, CD8 and E-cadherin is significantly modified in paired tumour samples before and after chemotherapy.

Previous characterization of the role of Galectin-3 in gastric cancer conflicts with our pilot study data. In a meta-analysis comprised of eight retrospective case-control studies, low Galectin-3 was associated with poor prognosis.[165] Potential reasons for this disparity may be due to the evaluation of adjuvant chemotherapy treatment only or the heterogeneous study population consisting of Asian, Turkish and South American patients. Our findings are consistent with the mechanistic role of Galectin-3 in *in-vitro* cell lines and mouse models which found that increasing Galectin-3 contributes to chemoresistance and M2 dominated immunosuppressive tumour microenvironments.[214–216] Furthermore, low Galectin-3

expression in bladder cancer is also associated with treatment response following neoadjuvant platinum based chemotherapy.[236]

This is the first study to evaluate the role of tumour-associated CD4/CD8 Ratio in gastric cancer neoadjuvant chemotherapy response. Increasing evidence has demonstrated the coordinated role of $CD4^+$ and $CD8^+$ T cells in mediating tumour immune surveillance, immunotherapy response and cancer prognosis.[165] Increasing CD4/CD8 Ratio of the tumour-host interface in triple negative breast cancer is associated with improved overall and recurrence-free survival.[165] In gastric cancer, Liu et al. found that increasing cytotoxic $CD8^+$ T cells relative to immunosuppressive $FOXP3^+$ T_{reg} cells is associated with improved prognosis.[222] In contrast to this study, we focused on $CD4^+$ T cell abundance without sub-stratification into T_h1, T_h2 and T_{reg} populations.

The dynamic increase in TIL expression following neoadjuvant chemotherapy in our pilot study also replicates previous findings. Significant work in breast cancer has implicated the pattern of TIL changes following chemotherapy to treatment response. In particular, greater $CD4^+$ T cell expression is associated with pathologic complete response.[238] Also, decreased immune infiltration is a notable characteristic of residual tumours following neoadjuvant chemotherapy relative to pre-treatment biopsy.[238] Continued evaluation of the relationship of dynamic changes in CD4 and CD8 populations in gastric cancer are required to fully leverage these biomarkers.

Our study design is intended to provide a reproducible and externally valid method of biomarker analysis. Using IHC allows for easier clinical implementation given that common pathology workflows already include IHC analysis. Our use of open-source digital pathology software such as QuPath also provides a standardized basis to internally and externally validate our method in future studies. Digital pathology allows measurement of annotated regions of interest within the software thus eliminating the need for complex physical microdissection utilized in other biomarker studies.

The main limitation of this study is the low enrollment of curative intent patients. This is likely due to low disease incidence in our population but also may be related to the COVID-19 pandemic. Given our rate of patient enrollment, future study should prioritize increasing sample size by using a retrospective design in order to provide more accurate estimates for future multi-centre prospective study. Our sample size calculation suggests that a limited

retrospective study with approximately eighty-five patients in each group will provide adequate power to assess these relationships.

2.5 Conclusion

The CD4/CD8 Ratio is a promising IHC-based biomarker with therapeutic implications for response to neoadjuvant chemotherapy in locally advanced gastric cancer. Future inquiry should focus on evaluating the prognostic value of these markers and the generation of a sufficient sample size to establish a predictive model for potential future clinical use.

Chapter 3: Murine gastric tissue establishes and maintains organoid viability after 48 hours of cold ischemia time in a mock-shipment simulation

3.1 Introduction

Organoid culture is an intriguing tool to advance translational medicine. Patient-derived organoids have been demonstrated to recapitulate parent tumour histology and molecular characteristics with higher fidelity compared to traditional 2-dimensional cell culture.[176] Organoids may be utilized longitudinally by creating biobanks and can be perpetuated *in-vitro* or *in-vivo* as xenografts in immunodeficient mice.[193] In oncology, *in-vitro* patient-derived organoid dose-response to anti-cancer therapy has been shown to correlate with clinical recurrence-free survival.[239]

Organoid culture is a specialized technique and resource-intensive.[240] Organoid development requires the procurement of fresh tissue, often from rare diseases, via biopsy, needle biopsy or surgical specimens.[174,178] Even with modern protocols, the successful establishment of tumour organoid culture ranges from 16% to 100% depending on tissue origin.[241,242] In contrast to biologically homogenous immortal cell lines, organoid research requires larger sample sizes to demonstrate a meaningful biological signal due to increased heterogeneity from patient-derived samples. Thus, tissue resources for organoid culture must be optimized for efficient and reliable testing to advance our understanding of organoid models and their applicability to the clinic.

We propose that tissue procurement for organoid culture could be increased by developing a shipment protocol for fresh surgical tissues or biopsies. In this scenario, increased organoid sample sizes could be achieved and more elaborate or controlled experimentation could occur if tissue for organoid growth was routed to fewer specialized labs. In our study, we investigate the feasibility of shipping fresh gastric tissue on ice over 24 or 48 hours for the purpose of establishing organoid culture. We assess the effects of various shipping media (Hank's Balanced Salts, University of Wisconsin solution and HTK solution) and dissociation times (Fresh, 24 hours and 48 hours) on organoid viability, growth rate, $LGR5^+$ and $TROY^+$ stem cell populations and anti-cancer drug response.

3.2 Methods

3.2.1 Study design

The study design is illustrated in Figure 1. Briefly, we performed gastrectomy in sixty C57BL/6J mice divided across 5 experimental cohorts. Twelve mice were included in each cohort. Three mouse stomachs were used for *in-vivo* controls and immediately processed for molecular and histology analysis. The nine remaining stomachs were cleaned and one each was placed in HBSS, UW or HTK solution and dissociated immediately as fresh tissue or mock shipped on ice in Styrofoam coolers and dissociated after 24 or 48 hours. Organoids were maintained for a total of 10 passages. Specimens for gene expression and histology were collected from control, initial dissociation, Passage 1 and 5 samples. A dose-response assay was conducted on all nine gastric organoids in cohort 4 from passage 2 and 6. Two human gastric cancer biopsies were mock-shipped for 24 and 48 hours respectively and established as human organoid culture (53C = 24 hours, 54C = 48 hours).

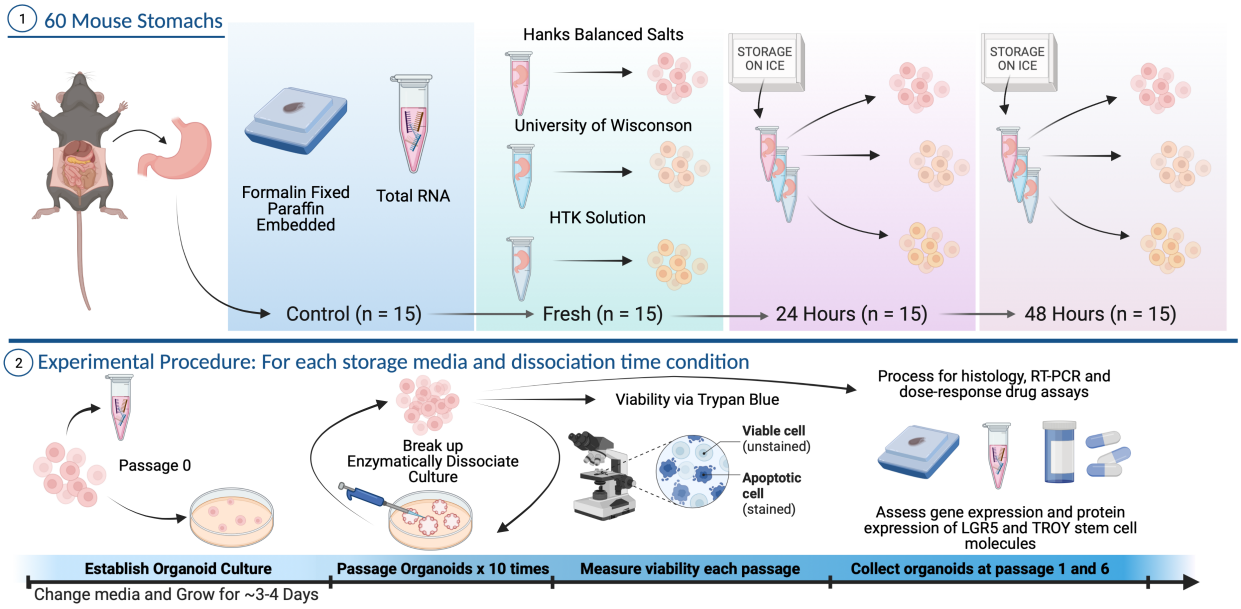


Figure 3-1 Study design.

3.2.2 Mice

Jackson C57BL/6J mice were retrieved from breeding colonies maintained by Health Sciences Laboratory Animal Services at the University of Alberta. Prior to gastrectomy, these mice were cared for according to guidelines established by the Canadian Council on

Animal Care. Ethics approval was granted by the University of Alberta Research Ethics Office under a Category A Exemption based on the principle of Reduction.

3.2.3 Organoid culture

Our organoid protocol closely approximates previously characterized intestinal organoid protocols from Bartfeld et al and Mowat et al.[179,243] Specific reagent information and catalogue numbers can be found in Appendix .C.1. Briefly, stomachs were resected from mice, washed in 1X sterile PBS and placed in 50mL conical tubes containing HBSS with 2X penicillin/streptomycin, UW or HTK solution (Figure 3-1). Tissues were minced with sterile scissors and washed three times in sterile 1XPBS. Minced tissues were enzymatically and mechanically digested in 20 mL digestion buffer (Advanced DMEM/F12, 1X penicillin/streptomycin, 2.5 ug/mL amphotericin B, 2.5% FBS, 75 units/ml Collagenase XI and 125 ug/ml Dispase II) and placed a mechanical water bath at 37 degrees Celsius for 1 hour. Next, tissue was pipetted aggressively up and down 10 times with a 10 mL pipette to dissociate cells and subsequently filtered through a 70µm strainer coated in 10%FBS/PBS. Cells were washed and suspended in DMEM. A 15 µL aliquot was mixed 1:1 with Trypan Blue (Gibco, 15250061) and counted on a hemocytometer. Cells were resuspended in ice-cold 70% Matrigel (Corning, 356253) in Advanced DMEM/F12 at a concentration of 1000 cells/µL and 35 µL Matrigel domes were placed in a prewarmed 24 well tissue culture treated plate. Organoids were cultured in 500 µL organoid culture medium at 37 degrees Celsius and 5% CO_2 until mature for splitting or downstream analysis. Organoid growth and media were assessed daily with media changes occurring every other day unless significant cellular debris or colorimetric change was present.

Organoid culture media contained 1:1 basal culture medium and conditioned L-WRN cell supernatant enriched in R-spondin, noggin and Wnt (ATCC, CRL-3276). Conditioned L-cell supernatant was prepared according to Miyoshi and Stappenbeck.[244] Organoid culture media contained Advanced DMEM/F12, 2mM L-Glut, 10mM HEPES, 1X penicillin/streptomycin, 2.5 ug/mL amphotericin B, 1X N2, 1X B27, 1 mM N-acetylcysteine, 1nM gastrin, 10 mM nicotinamide, 500 nM A83-01, 10 µM SB202190 and 50 ng/mL mouse EGF. Human organoid culture used human EGF and, for the first plating only, 10 µM Y-27632.

Mature organoids were passage approximately every 5 days depending on the abundance of cystic organoids and cell sloughing (Figure 3-1). Briefly, ice-cold Advanced DMEM/F12 was added to Matrigel domes and organoids were mechanically lifted using a pipette tip. Organoids were dissociated using TrypLE Express (Gibco, 12604013) and washed three

times and counted with a hemocytometer as above. Organoid cells were then cultured or allocated for downstream analysis. Organoid viability and relative growth rate were measured using a hemocytometer with Trypan Blue for each passage. The relative growth rate was calculated assuming exponential growth.

3.2.4 Annexin V-Cy3 Apoptosis Assay

We used an Annexin V-Cy3 immunocytochemistry detection kit (Sigma-Aldrich, APOAC) to assess the proportion of live, dead or apoptotic cells for different tissue dissociation conditions. Briefly, 50,000 cells were retrieved following the organoid culture filtration step above and placed on Poly-L-Lysine coated slides. The Annexin assay was performed per manufacturers protocol and imaged with an AxioCam HRc camera. Cells were counted using ImageJ.[245]

3.2.5 Immunofluorescence and immunohistochemistry

Tissue specimens for hematoxylin and eosin (H+E) staining and immunofluorescence (IFC) of whole mouse stomach or organoid containing Matrigel domes were fixed in zinc-formalin (Z-Fixx) for 24 hours, washed three times and stored in 70% ethanol prior to preservation in paraffin. Fixed organoids were suspended in agar prior to paraffin embedding. Specific antibody combinations, dilutions, incubation times and antigen retrieval buffers are listed in Appendix .C.2. Briefly, 5 µm tissue sections were deparaffinized in Histoclear (National Diagnostics) and rehydrated. Microwave heat-induced epitope retrieval was performed using Sodium Citrate (pH 6, heated to 94 degrees Celsius in 1-minute intervals followed by 9 minutes continuous heat). Permeabilization was performed with 0.5% Triton X-100. Non-specific epitopes were blocked using 10% normal goat serum. Tissue sections were stained with primary antibodies (rabbit anti-pan cytokeratin/mouse anti-MUC5AC, rabbit anti-pan cytokeratin/mouse anti-TROY, and rabbit anti-pan cytokeratin/mouse anti-LGR5) and secondary antibodies (anti-rabbit IgG Alexa Fluor 488 (green) and anti-mouse IgG Alexa Fluor 568 (red)). Autofluorescence was diminished using a TrueView Quenching kit per manufacturer's protocol (Vector, SP-8400-15). Nuclear counterstaining was performed with DAPI (blue) followed by cover slipping with Vectashield Vibrance Antifade mounting media (Vector, H-1700). Images were captured on an AxioCam HRc camera and processed using ImageJ.[245]

Immunohistochemistry was performed on human gastric organoids. Tissues were rehydrated and antigen retrieval was performed as above. Endogenous peroxidases were

blocked with 3% hydrogen peroxide in methanol. Non-specific binding was mediated by blocking with 20% normal goat serum and avidin/biotin blocker per manufacturers protocol (Vector Laboratories-2001). MUC5AC primary antibody was diluted 1:250 and incubated at 4 degrees Celsius overnight. Biotinylated IgG secondary antibody was incubated at 1:200 dilution for 30 minutes. Antibody detection was performed using avidin-biotin complex/horseradish peroxidase (Vector Laboratories) and 3,3-diaminobenzidine tetrahydrochloride (DAB, Abcam, ab64238) per manufacturer's protocol. Tissues were counterstained with hematoxylin.

3.2.6 qRT-PCR

RNA was extracted from organoids or whole mouse stomach using TRIzol Reagent according to the manufacturer's protocol (Invitrogen, 15596026). Whole mouse stomach was disrupted using a bead mill homogenizer. RNA concentration and purity were assessed using a NanoDrop 1000 (Thermo Scientific). Only samples with an A260/280 greater than 1.8 were included for analysis. Isolated RNA was stored at minus 80 degrees Celsius with 0.2 Units/μL SUPERASE-In RNase inhibitor (Invitrogen, AM2694). Total RNA was reverse transcribed into cDNA using 100 ng RNA with the High-Capacity RNA-to-cDNA Kit (Applied Biosystems, 4387406). Quantitative real-time PCR was performed in a 96 well CFX Connect Bio-Rad RT-PCR Detection System using 10 ng input cDNA, TaqMan Fast Advanced Master Mix (Applied Biosystems, 4444964) and stock TaqMan PCR primers from ThermoFisher Scientific [Tnfrsf19/TROY (Mm00443506_m1), LGR5 (Mm00438890_m1), and beta-actin/ACTB (Mm01205647_g1)].

PCR Amplification efficiency was assessed to establish optimal input cDNA concentration and to validate the assay for our selected transcripts (Appendix .C.3; Appendix Figure C3-2). We conducted a validation experiment to confirm equal efficiency of the target and housekeeping genes (Appendix .C.3; Appendix Figure C3-1).[246] Relative gene expression levels were calculated using the Comparative CT Method (delta delta/ΔΔCT) using pooled passage 8 mouse organoids as reference. Gene expression values were subsequently scaled to the mean of whole stomach controls prior to downstream analysis.

3.2.7 in-vitro dose-response assay

Our combination FOLFOX (5-fluorouracil, oxaliplatin and leucovorin) dose-response assay was validated using the human gastric cancer cell line AGS (ATCC CRL-1739). AGS cells were cultured in Ham's F-12K (Gibco), 10% FBS (Gibco) and 1X penicillin and streptomycin

(Gibco). Briefly, 5000 cells were plated in 96 well plates and grown for 24 hours. Anti-cancer drugs 5-fluorouracil (Tocris, 3257) and oxaliplatin (Tocris, 2623) were added in triplicate over 8 half-log dilutions with 800 µM and 2400 µM initial concentrations, respectively. A single 500 µM dose of leucovorin (Toronto Research, L330400) was added to each treatment well. Cells were treated for 48 hours followed by a CCK-8 viability assay (Abcam, ab228554), which was performed according to the manufacturer's protocol. We replicated our assay in three independent trials to assess reproducibility.

For organoid dose-response assays, organoids were passaged and dissociated according to our protocol. As above, 5000 cells were plated in 96 well plates and grown in organoid media for 24 hours followed by 48 hours of FOLFOX treatment and viability assessment using a CCK-8 assay.

3.2.8 Statistical analysis

Statistical analyses were completed using R version 4.1.2 and Prism 9.3.1.[226] Differences between three or more groups were assessed using Kruskal-Wallis test with post-hoc Dunn's test for multiple comparisons when applicable. Multiple comparison corrections were performed using the Bonferroni method. A paired Wilcoxon test was used to assess paired dose-response data. We used Pearson's correlation to assess the relationship between gene expression and dose-response data. A one-way analysis of variance (ANOVA) was used to compare viability and growth rate. Multivariable generalized additive models (GAMs) of growth rate, viability and stem cell gene expression were created to determine the independent effects of media, dissociation time, cohort and passage number.[247] A series of models were analyzed and compared using likelihood ratio tests of nested models. We used Variable Inflation Factor to assess for collinearity. GAM fits were tested relative to multiple linear regression and linear regression with relevant spline terms for continuous variables. Predictions from these models were used to create adjusted response plots using ggplot2.

Dose response data was processed and analyzed in GraphPad. First, baseline media control absorbance was subtracted from all experimental wells. Absorbance values were then normalized between 0 and 100% given that adequate minimum and maximum treatment effect controls were present. Mean drug concentrations from 5-fluorouracil, oxaliplatin and leucovorin were log10 transformed. Next, outliers were removed using Q = 1% and least squares variable slope non-linear regression estimated dose-response curves and half-maximal inhibitory concentration (IC50).

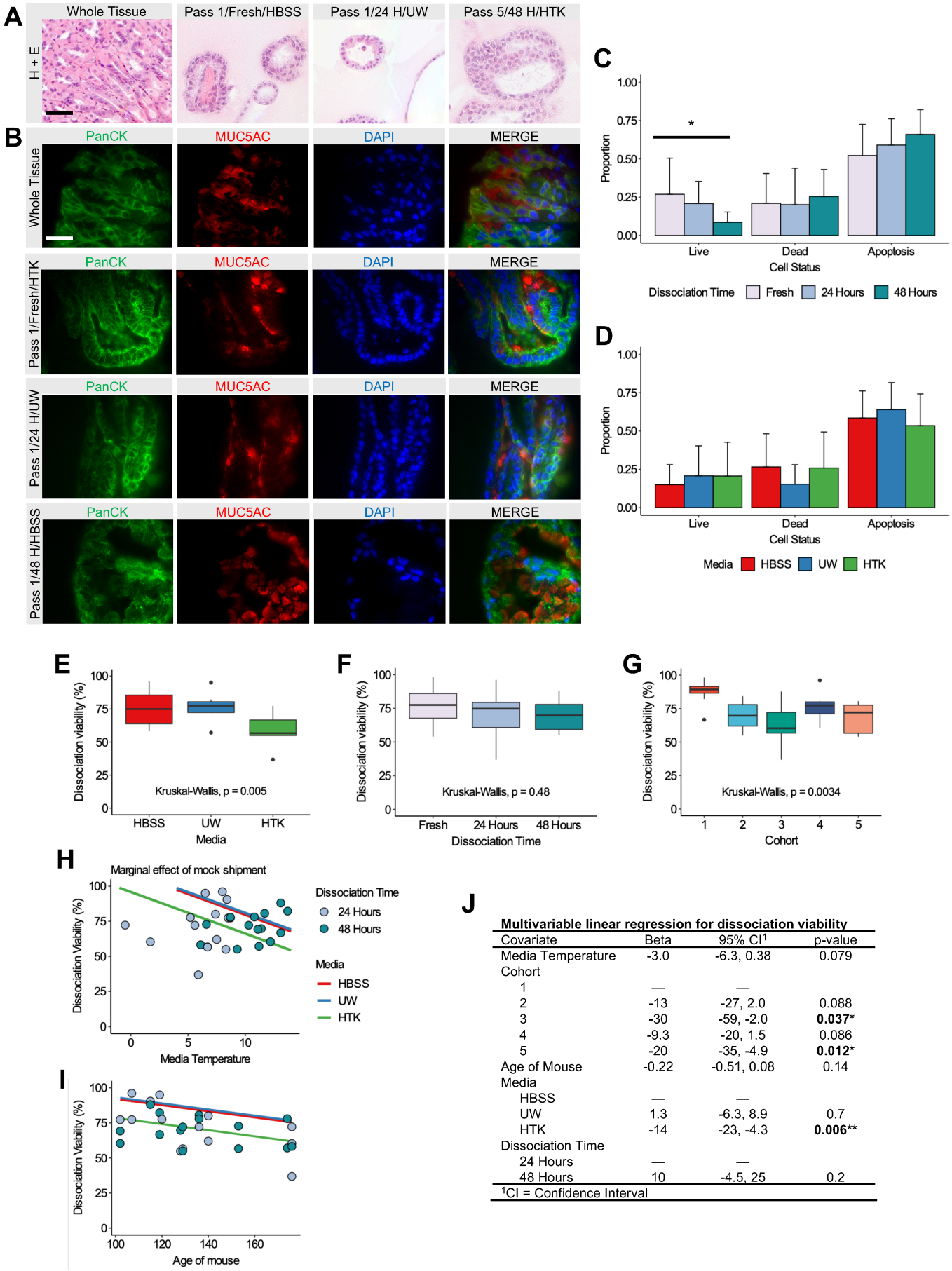
3.3 Results

3.3.1 Organoids recapitulate murine gastric tissue and morphology

We successfully established gastric organoids from 45 mice (34 male, 26 female) with a median age of 129 days (IQR 119-140). The median number of days between passages was 5 (3,6) and the median total time spent in culture was 54 (41,56) days. We assessed the morphology of gastric organoids from each media condition and dissociation time relative to whole mouse stomach using hematoxylin and eosin staining. Figure 3-2A shows representative images that demonstrate the recapitulation of tissue morphology. Irrespective of media or dissociation time, organoids from passages 1 and 6 formed circular hollow structures with cells bound by a basement membrane reminiscent of gastric epithelial tissue. We confirmed that our organoids were of gastric origin by assessing the presence of MUC5AC, which is exclusively present in normal gastric tissue (Figure 3-2B).[248] Once again, structures reminiscent of gastric glandular tissue were observed with MUC5AC expression localized to the interior of hollow structures bound by epithelial cells identified by pan cytokeratin.

3.3.1 Dissociation time is associated with decreased viable cells when accounting for early apoptosis

We used an Annexin V-Cy3 apoptosis assay to assess the effect of shipment media and dissociation times on isolated stomach cells prior to organoid culture. A significant difference in healthy viable cells was demonstrated between dissociation times (Kruskal-Wallis, $p=0.049$). Increasing shipment times were associated with decreasing live cells. Minimal difference was observed in the proportion of necrotic/dead cells between dissociation times but slight increases in cells undergoing early apoptosis occurred with increasing shipment time. There was no significant differences or appreciable trends regarding cell status between shipment media conditions.



J

Multivariable linear regression for dissociation viability

Covariate	Beta	95% CI[1]	p-value
Media Temperature	-3.0	-6.3, 0.38	0.079
Cohort			
1	—	—	
2	-13	-27, 2.0	0.088
3	-30	-59, -2.0	**0.037***
4	-9.3	-20, 1.5	0.086
5	-20	-35, -4.9	**0.012***
Age of Mouse	-0.22	-0.51, 0.08	0.14
Media			
HBSS	—	—	
UW	1.3	-6.3, 8.9	0.7
HTK	-14	-23, -4.3	**0.006****
Dissociation Time			
24 Hours	—	—	
48 Hours	10	-4.5, 25	0.2

[1]CI = Confidence Interval

Figure 3-2 Gastric organoid histology and dissociation viability. **A.** Representative hematoxylin and eosin images of whole mouse stomach and gastric organoids. The passage number, media conditions and dissociation times are noted in the image headings. Scale bar represents 50 μm. **B.** Immunofluorescent images of whole mouse stomach and gastric organoids recapitulating epithelial pan cytokeratin (green) and gastric MUC5AC (red). Nuclei is stained with DAPI (blue). Scale bar represents 20 μm. **C and D.** Barplots of Annexin V-Cy3 apoptosis assay for dissociation time (C) and media conditions (D). Experimental conditions are represented by colours identified in the plot legend. Statistical analysis was conducted with Kruskal-Wallis test (*p<0.05). Error bars represent standard deviation. **E, F and G**. Boxplots representing dissociation viability (percent of living cells) measured by trypan blue dye for media (E), dissociation time (F) and cohort (G) conditions. The significance of a Kruskal-Wallis test is annotated in each respective plot. **H and I.** Marginal effect plots for dissociation viability as measured by trypan blue for media temperature and age of mouse given the multivariable model (J). The effects are stratified by shipment media and coloured according to the legend plot. Individual data points are coloured according to dissociation time as specified in the legend plot. **J.** Regression table of linear regression model for dissociation viability. Covariates are listed with accompanying beta coefficient values, 95% confidence intervals and p-values. Hyphens denote reference category for discrete variables. Significant p values are bold.

3.3.2 Dissociation viability is affected by shipment media and is cohort dependent

Initial cell viability following dissociation was also assessed using trypan blue dye and a hemocytometer. In this assay, live cells comprise viable cells and cells undergoing apoptosis with an intact cell membrane. Univariable analysis of dissociation viability is demonstrated in Figure 3-2E, F and G with respect to media, storage time and each dissociation procedure/cohort. Cell viability was significantly decreased in tissues shipped in HTK solution relative to HBSS and UW solutions (Dunn's Test, HBSS-HTK adjusted p=0.03, UW-HTK adjusted p=0.01). Significant variability in dissociation viability was identified between cohorts (Kruskal-Wallis, p<0.01).

We constructed a multivariable linear regression model to evaluate the effect of measured confounders and our experimental groups on dissociation viability following shipment for 24 and 48 hours (Figure 3-2J). Of note, no difference in cell viability was found between cells dissociated after 24 or 48 hours. HTK solution and cohorts 3 and 5 were significantly associated with decreased cell viability relative to HBSS and cohort 1, respectively. Mouse age and the temperature of the media measured at the time of dissociation were not significantly associated with dissociation time. In Figure 3-2H, we observe that each one-degree Celsius increase in the temperature of shipment media corresponds to a 3%

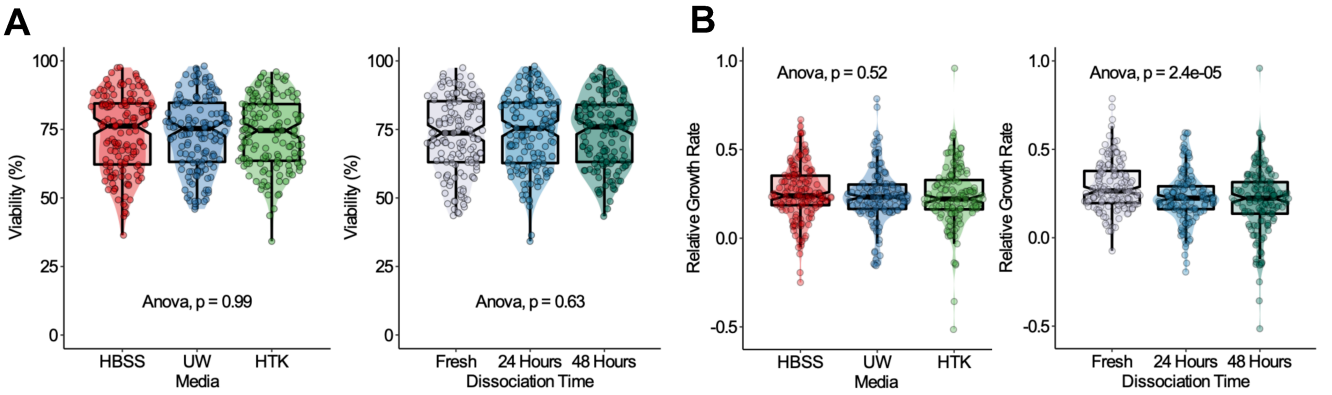
decrease in dissociation viability. A small decrease in viability was also associated with increasing mouse age (Figure 3-2I).

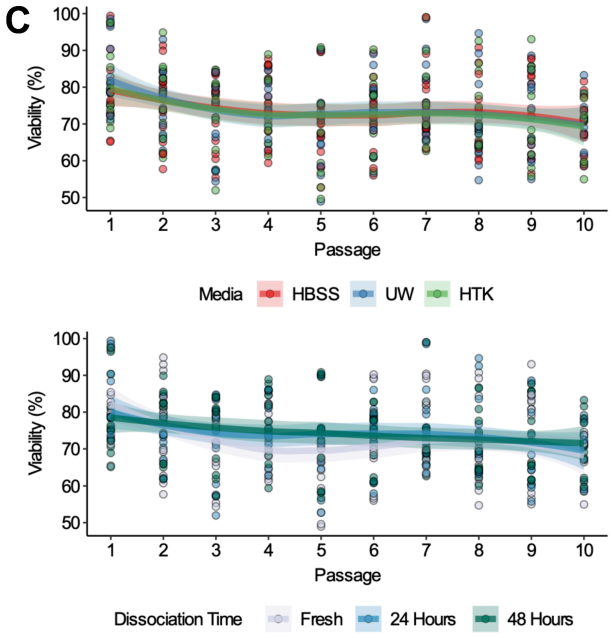
3.3.1 Long-term organoid viability is not affected by shipment media and dissociation time

Organoid cell viability was measured during each passage to assess the overall health of organoids following mock-shipment. Of note, we established viable organoid cultures in 100% of experimental conditions and all organoids were propagated for a total of 10 passages. Univariable analysis found no significant difference in pooled viability with respect to shipment media and dissociation time (Figure 3-3A Left and Right, respectively).

Next, we assessed the relationship of cell viability over the course of 10 organoid passages. Given the significant cohort-dependent effects related to dissociation viability we examined the trend of viability for each cohort (Appendix Figure C5-4). We observed heterogenous nonlinear variation in viability between cohorts, however, there were no obvious effects related to shipment media or dissociation time (Appendix Figure C5-5).

Using likelihood ratio tests, we developed a GAM to best approximate the relationship of viability with our experimental conditions and relevant confounding variables (Figure 3-3D). Penalized cubic regression splines were fit using generalized cross validation for passage number, days in culture and the relative growth rate. We visually present the adjusted effects of shipment media and dissociation time on viability in Figure 3-3C. Of note, despite differences in shipment media and dissociation time, organoid viability exists within a fairly narrow range between ~65 and 85% with tight 95% confidence intervals of the loess smooth functions. We found no significant relationship in viability between UW or HTK solutions relative to HBSS solution. We identified a small, yet statistically significant increase in cell viability for organoids dissociated after 24 and 48 hours of mock-shipment relative to fresh organoids. In Figure 3-3C Bottom, we observe that 24- and 48-hour organoids have a relatively linear relationship with viability over successive passages whereas fresh organoids exhibit a sinusoidal pattern.



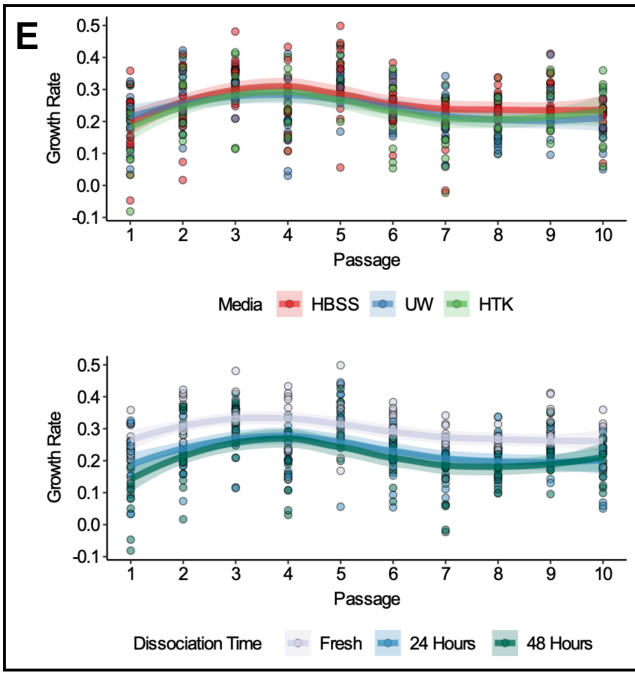


D

Multivariable model for organoid viability

Covariate	Beta	95% CI[1]	p-value
Media			
HBSS	—	—	
UW	0.55	-1.5, 2.6	0.6
HTK	0.77	-1.3, 2.9	0.5
Dissociation Time			
Fresh	—	—	
24 Hours	4.0	1.9, 6.1	**<0.001**
48 Hours	4.5	2.3, 6.6	**<0.001**
Cohort			
1	—	—	
2	-6.2	-9.2, -3.1	**<0.001**
3	-9.1	-12, -6.2	**<0.001**
4	-10	-13, -7.3	**<0.001**
5	-14	-18, -11	**<0.001**
s(Passage number)			**<0.001**
s(Days in culture)			**<0.001**
s(Relative growth rate)			**<0.001**

[1]CI = Confidence Interval



F

Multivariable model for organoid relative growth rate

Covariate	Beta	95% CI[1]	p-value
Media			
HBSS	—	—	
UW	-0.02	-0.05, 0.01	0.2
HTK	-0.02	-0.05, 0.01	0.2
Dissociation Time			
Fresh	—	—	
24 Hours	-0.08	-0.11, -0.04	**<0.001**
48 Hours	-0.09	-0.12, -0.06	**<0.001**
Cohort			
1	—	—	
2	0.01	-0.04, 0.05	0.8
3	0.01	-0.04, 0.05	0.8
4	0.05	0.01, 0.10	**0.025**
5	0.06	0.00, 0.11	**0.046**
Viability	0.01	0.00, 0.01	**<0.001**
s(Passage number)			**<0.001**
s(Days in culture)			**<0.001**

[1]CI = Confidence Interval

Figure 3-3 Gastric organoid viability and growth rate over 10 *in-vitro* passages. **A.** Boxplot-scatter-violin plots of percent viability versus media (Left) and dissociation time (Right). **B.** Boxplot-scatter-violin plots of relative growth rate versus media (Left) and dissociation time (Right). For (A) and (B), the significance of a one-way ANOVA test is annotated in each plot. Colours in plot legend indicate experimental conditions. Boxplot notches approximate the 95% confidence interval of the median. Each data point represents a single measurement. The distribution of data is represented by superimposed violin plots. **C and E.** Adjusted response plot for viability (C) and growth rate (D) versus passage number stratified by media and dissociation time. Adjusted viability estimates were predicted from the multivariable models (D and F). Each data point represents a single measurement. The trend in data is estimated using a loess smooth bound by semi-transparent 95% confidence intervals. **D and F.** Regression table of generalized additive models for viability (D) and growth rate (F). Covariates are listed with accompanying beta coefficient values, 95% confidence intervals and p-values. Covariates listed with s() were fit using penalized cubic regression splines optimized by generalized cross validation. Hyphens denote reference category for discrete variables. Significant p values are bold.

Once again, we found that specific cohorts exerted significant batch effects on cell viability Figure 3-3D). The graphical representation of nonlinear covariates presented in Appendix Figure C5-6 provides insight into the relationship of the significant variables passage number, days in culture and growth rate on cell viability. For example, passage number exerts a complex nonlinear effect on viability, whereas days in culture and growth rate exert relatively near-monotonic effects on viability.

3.3.2 Mock-shipment for 24 or 48 hours decreases organoid growth rate

Univariate analysis of shipment media found no significant effect on pooled growth rate (Figure 3-3A Right). However, organoids derived from fresh tissues were associated with significantly greater relative growth rate (Figure 3-3B Right) compared to organoids generated after 24-hour and 48-hour mock-shipment (Dunn's Test, Fresh-24 Hours adjusted $p < 0.001$, Fresh-48 Hours adjusted $p = <0.001$).

Compared to viability, growth rates over successive passages were similar between cohorts (Appendix Figure C6-7). We fit a GAM to model growth rate in keeping with the procedure outlined in the previous section. After adjusting for confounders, dissociation after 24 or 48 hours remained significantly related to decreased growth rate (Figure 3-3F; Appendix Figure C6-8). In Figure 3E Bottom, we observed this consistent effect across all 10 passages. Once again, shipment media did not exert significant effects on growth rate (Figure 3-3E Top and Figure 3-3F).

Variability in growth rate was significantly related to cohort 4 and 5, viability, passage number and the number of days in culture. In our model, viability was best interpreted as a linear relationship with growth rate in which increasing viability was associated with increasing growth rate. The complex nonlinear relationships for passage number and the number of days in culture are observed in Appendix Figure C6-9.

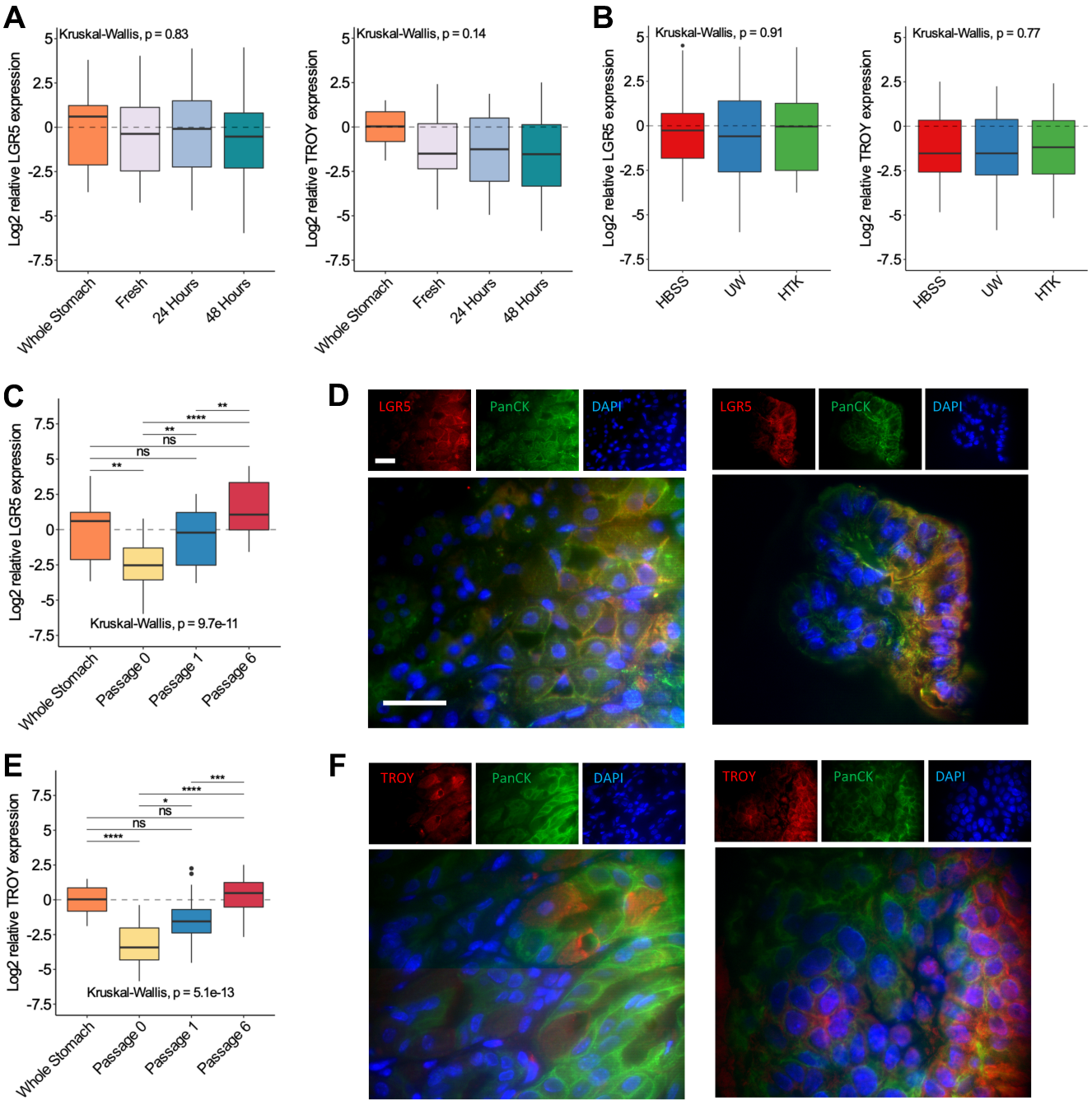


Figure 3-4 Expression of LGR5 and TROY stem cell markers in gastric organoids. **A.** Boxplots of Log2 relative LGR5 gene expression (Left) and Log2 relative TROY gene expression (Right) versus dissociation time. Whole stomach indicates control endogenous gene expression levels. **B.** Boxplots of Log2 relative LGR5 gene expression (Left) and Log2 relative TROY gene expression (Right) versus media. **C.** Boxplot of Log2 LGR5 gene expression versus passage number. The significance of a Kruskal-Wallis and Dunn's post-hoc test is annotated in the plot. (****p<0.0001, ***p<0.001, **p<0.01, *p<0.05, ns = not significant) **D.** Immunofluorescence expression of LGR5 (red) and pan cytokeratin (green) in whole mouse stomach and gastric organoids. Nuclei is stained with DAPI (blue). **E.** Boxplots of Log2 TROY gene expression versus passage number. The significance of a Kruskal-Wallis and Dunn's post-hoc test is annotated in the plot. **F.** Immunofluorescence expression of TROY (red) and pan cytokeratin (green) in whole mouse stomach and gastric organoids. Nuclei is stained with DAPI (blue). For C and E, gene expression from control endogenous tissue and cells isolated from the initial dissociation are denoted by Whole Stomach and Passage 0, respectively. Gene expression for Passage 1 and 6 is measured from established organoids. For D and F, scale bar represents 20 μm.

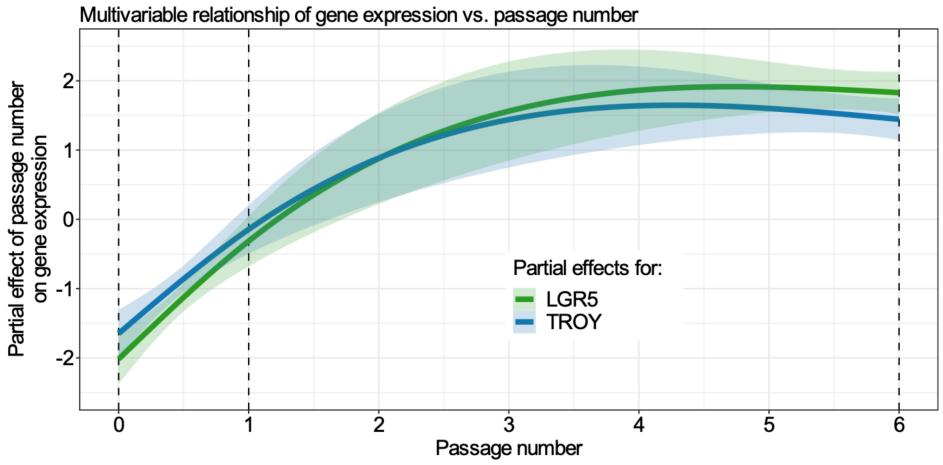
3.3.3 *in-vitro* gene expression of gastric stem cell markers LGR5 and TROY re-establishes baseline endogenous levels after 6 passages and is unaffected by mock-shipment conditions

We assessed the gene expression of essential gastric stem cell molecules LGR5 and TROY in whole stomach tissue, upon dissociation (passage 0) and after passages 1 and 6. No significant relationship was identified in LGR5 and TROY gene expression between organoid dissociation time or shipment media (Figure 3-4A and B). In Figure 3-4C and E, we identified significant differences between LGR5 and TROY gene expression between whole stomach tissue, dissociated tissues and subsequent organoid passages. We examined the significance of differences between each condition with Dunn's test. In both cases, LGR5 and TROY gene expression significantly decreased upon dissociation of tissues relative to endogenous whole stomach levels. Stem cell gene expression significantly increased in successive passages and was statistically similar to endogenous levels following passage 1 and 6 (Figure 3-4C and E).

To confirm the translation of LGR5 and TROY, we assessed protein expression using immunofluorescence. In Figure 3-4D, LGR5 expression is noted in representative images of whole stomach tissue (Left) and gastric organoid (Right). LGR5 was predominately co-localized with pan cytokeratin expressed on cell membranes, which is consistent with prior studies.[249,250] TROY expression was also identified in whole stomach tissue and gastric organoids (Figure 3-4F Left and Right, respectively). In both tissues, TROY is expressed

mainly in the cytoplasm and cell membrane, although some expression is present in nuclei.[250]

A



B

Multivariable model for LGR5 gene expression

Covariate	Beta	95% CI[1]	p-value
Cohort			
1	—	—	
2	-0.55	-1.4, 0.34	0.2
3	-0.29	-1.2, 0.66	0.5
4	-0.78	-1.7, 0.11	0.090
5	2.4	1.5, 3.3	**<0.001**
Media			
HBSS	—	—	
UW	-0.23	-0.82, 0.35	0.4
HTK	-0.07	-0.66, 0.52	0.8
Dissociation Time			
Fresh	—	—	
24 Hours	0.20	-0.39, 0.78	0.5
48 Hours	-0.31	-0.89, 0.28	0.3
s(passage)			**<0.001**

[1]CI = Confidence Interval

C

Multivariable model for TROY gene expression

Covariate	Beta	95% CI[1]	p-value
Cohort			
1	—	—	
2	-2.2	-3.1, -1.3	**<0.001**
3	-1.1	-2.1, -0.21	**0.018**
4	-2.4	-3.3, -1.5	**<0.001**
5	-0.76	-1.6, 0.10	0.086
Media			
HBSS	—	—	
UW	-0.37	-0.94, 0.20	0.2
HTK	0.02	-0.56, 0.59	>0.9
Dissociation Time			
Fresh	—	—	
24 Hours	-0.09	-0.65, 0.48	0.8
48 Hours	-0.37	-0.94, 0.20	0.2
s(passage)			**<0.001**

[1]CI = Confidence Interval

Figure 3-5 Multivariable analysis of LGR5 and TROY gene expression in gastric organoid culture. **A.** Partial effects plot demonstrating the significant nonlinear effect of passage on LGR5 (green) and TROY (blue) gene expression. The y-axis represents the change in gene expression for a given passage number as identified by the smoothed function. The solid coloured lines are penalized cubic regression splines for passage number derived from multivariable models in B and C. Semi-transparent ribbons represent 95% confidence intervals. The dotted vertical lines represent times at which gene expression was measured. **B and C.** Regression tables of generalized additive models for LGR5 (C) and TROY (D) gene expression. Covariates are listed with accompanying beta coefficient values, 95% confidence intervals and p-values where applicable. Covariates listed with s() were fit using penalized cubic regression splines optimized by generalized cross validation. Hyphens denote reference category for discrete variables. Significant p values are bold.

Multivariable GAMs were established to assess TROY and LGR5 gene expression using mock-shipment covariates, cohort and passage number. We modelled passage number as a continuous variable to evaluate the nature of the relationship between organoid culture and gene expression. In Figure 3-5A, we observed a linear increase in TROY and LGR5 expression among dissociated tissues and organoids after passage 1. However, the rate of increase in stem cell gene expression subsides and appears to reach a steady-state by passage 4. Overall, passage number contributed significantly ($p<0.001$) to stem cell gene expression and the relationship between LGR5 and TROY was nearly identical. Again, significant cohort batch effects were identified for LGR5 and TROY expression (Figure 3-5B and C). Mock shipment conditions related to media and dissociation time remained insignificant.

3.3.4 *in-vitro* organoid dose-response to cytotoxic therapy is significantly associated with passage number and stem cell gene expression

Ideally, the shipment of tissues would increase organoid procurement and not affect downstream analyses such as dose-response assays. We performed a dose-response assay using FOLFOX cytotoxic therapy to assess possible effects related to mock-shipment. Using the AGS cell line, we validated the reproducibility of our drug assay. In three independent drug assays our dose-response curves were statistically similar (Appendix Figure C4-3).

Dose-response curves for a single cohort of organoids at passage 1 and 6 are shown in Figure 3-6A. We achieved reliable nonlinear estimates of all drug assays with adjusted goodness of fits (R^2) ranging from 0.82 to 0.99. Greater efficacy of cytotoxic therapy was observed in paired organoid lineages from passage 1 relative to passage 6 (Figure 3-6B, paired Wilcoxon, $p <0.01$). No significant difference in dose-response IC50 was identified between dissociation times or shipment media variables (Figure 3-6C and D).

Next, we assessed the Pearson's correlation of dose-response IC50 values with LGR5 and TROY gene expression measured from the same organoid lineages. Both LGR5 and TROY gene expression was strongly correlated with IC50 (Pearson's $R=0.7$, $p<0.01$). In Figure 3-6E and F we observe that this association also corresponds to clear differences in expression between organoid passages.

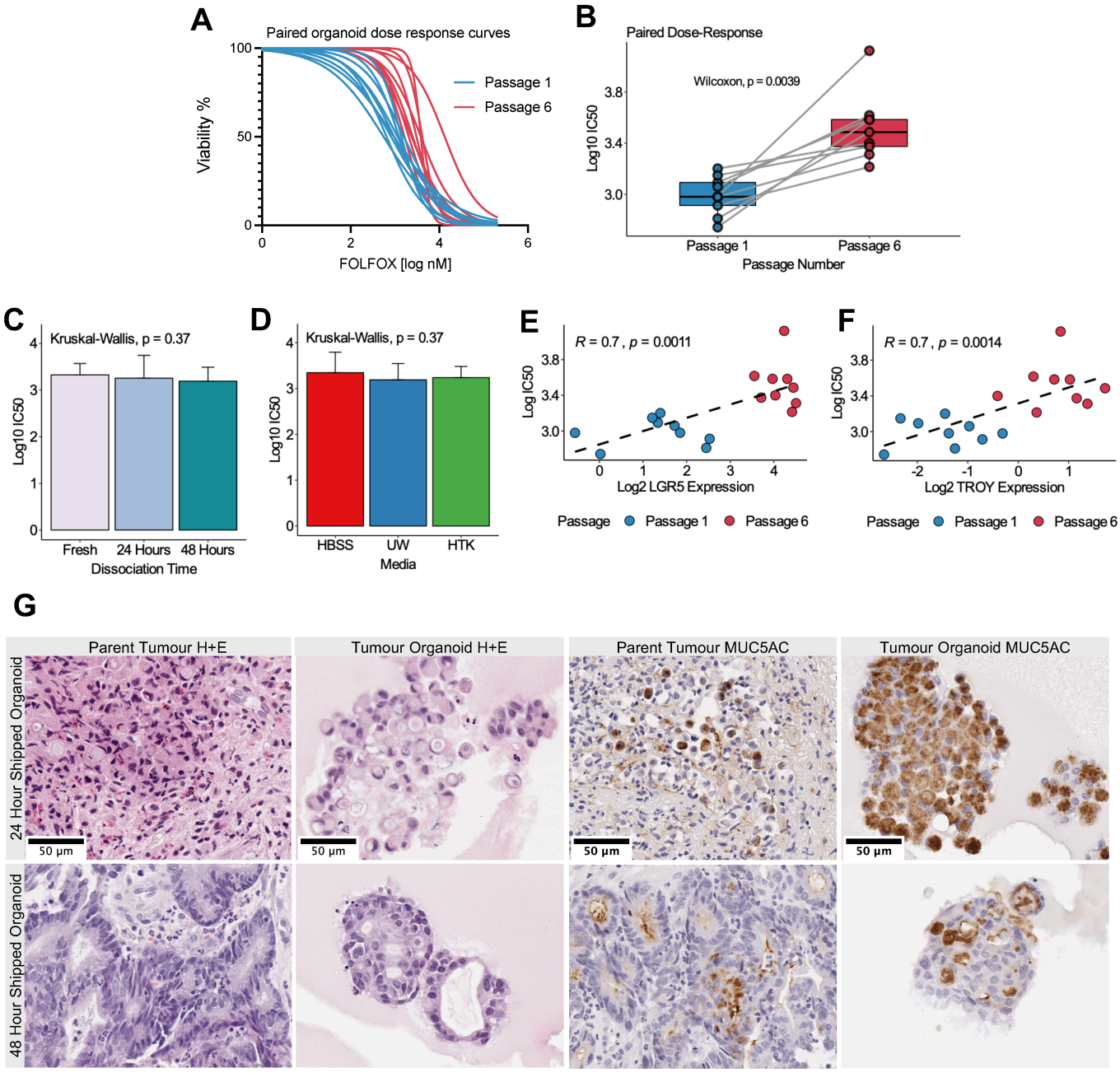


Figure 3-6 Effect of mock-shipment on dose-response to FOLFOX therapy and growth of human gastric cancer organoids. **A.** Dose-response curves for 9 organoid lineages in passage 1 (blue) and passage 6 (red). Cell viability measured by a CCK8 assay is presented on the y-axis and Log10 FOLFOX concentration on the x-axis. Least squares nonlinear models with variable slope were fit for each half-log 8 dilution series. Half-maximal inhibitory concentration (IC50) was estimated from these models. **B.** Paired boxplots for Log10 IC50 values versus passage number. Each coloured point and line correspond to a single organoid lineage. Organoids were derived from each experimental condition (n=9). Boxplots represent the distribution of Log10 IC50 for all organoid lineages. Paired Wilcoxon p-value is annotated in

the plot**. C and D.** Barplots of Log10 IC50 values versus dissociation time (C) and media conditions (D). Experimental conditions are represented by colours identified in the plot legend. Kruskal-Wallis p values are annotated in the plot. Error bars represent standard deviation. **E and F.** Pearson's correlation between Log10 IC50 values and Log2 relative gene expression for LGR5 (E) and TROY (F). Pearson's R and p value is annotated in each plot. Points represent one organoid lineage and are coloured according to passage as indicated in plot legend below. The dotted black line represents the fit of a simple linear regression. **G.** Representative brightfield images of human gastric cancer tumours and patient-derived organoids following 24- (Top) and 48-hours (Bottom) of mock-shipment in HBSS solution. The image identification is presented in the plot heading. The first two columns show hematoxylin and eosin images. The third and fourth columns demonstrates the recapitulation of MUC5AC (brown) expression in tumour organoids. Scale bar represents 50 μm.

3.3.5 Human tumour organoids can be established following 24- and 48-hour mock-shipment in HBSS solution

Endoscopic biopsies were retrieved from gastric cancer in 2 patients and mock-shipped in HBSS solution for 24 and 48 hours, respectively. All patients underwent informed consent according to our approved ethics protocol from the Health Research Ethics Board of Alberta. Organoids were successfully established and propagated following mock-shipment. Tumour-derived organoids recapitulated the morphology of parent tumour tissue in hematoxylin and eosin stained sections (Figure 3-6G). Furthermore, MUC5AC was expressed in organoid tissues, which confirmed the growth of gastric epithelium. Of note, organoids from 24-hour shipped tissue recapitulated diffuse type histology with signet ring cells noted in the parent tumour.

3.4 Discussion

Using a mouse gastric organoid model, we demonstrated that gastric organoids are capable of growth following cold shipment for 24 or 48 hours. Although HTK solution was associated with decreased viability following dissociation, no significant difference in viability or growth rate was found across 10 passages between HBSS, UW and HTK solutions. Decreased growth rate was identified in organoids established after 24- and 48-hours of mock-shipment, however, this did not affect cell viability, LGR5 and TROY stem cell gene expression or dose-response assays. These data support the concept that cold shipment of fresh tissue on ice is a feasible approach to improve organoid tissue procurement.

Intestinal organoid models are thought to provide a high-fidelity model that will likely form the cornerstone of future translational oncology research.[193,251] Several pre-clinical organoid

studies have identified a correlation between *in-vitro* organoid dose-response and clinical tumour response.[182,183,195,196] Despite the optimism and promising preliminary findings, prospective evidence supporting the use of organoids as a translational model for personalized medicine is lacking. The first clinical trial assessing whether colon cancer organoids could allocate clinical therapy failed to establish feasibility.[174] To improve organoid models additional research is required. Thus, enhancing tissue procurement eliminates a practical bottleneck preventing the expansion of organoid research.

Previous research has attempted to improve tissue procurement and preservation. Walsh et al. compared the effectiveness of flash freezing fresh tissue in liquid nitrogen to cryopreservation in dimethyl sulfoxide (DMSO) at -80 degrees Celsius prior to organoid development.[252] In comparison to organoids derived from fresh unfrozen tissue, organoids developed from cryopreserved tissues in DMSO most accurately recapitulated dose-response to various anticancer drugs. Another option to distribute organoid tissue involves shipping established organoid cultures from patient-derived tissue or induced pluripotent stem cells. Using live retinal organoid cultures, shipment protocols at room temperature or 37 degrees Celsius have proven to maintain organoid viability.[253,254] We argue that both of these methods are not feasible for the widespread collection and distribution of organoid tissues. The first method requires the availability of liquid nitrogen, adequate -80-degree Celsius freezer capacity and DMSO within surgical facilities. Further, the shipment of these frozen tissues would require expensive and elaborate shipping conditions. The second method relies on establishing organoid culture prior to transport. Thus, this requires the availability of a nearby organoid research lab. Furthermore, any established cultures will be prone to greater batch variability between separate organoid labs. Our method only requires the availability of affordable HBSS solution and ice packs.

In this study, we identified a number of significant confounders related to organoid viability, growth rate and dose-response. In all multivariable models, cohort batch effects were significant. We identified cohort-dependent effects on initial dissociation viability and among viability and growth rate over 10 organoid passages. Cohort effects upon tissue dissociation could be attributed to variation in reagents, subtle differences in the surgical resection procedures or variability in enzymatic dissociation. However, cohort batch effects occurring across 10 passages suggests that unknown variability is intrinsic to the establishment of each organoid culture. The growth and behaviour of gastric organoids was also found to associate with passage number and days in culture. Complex nonlinear relationships were

identified for viability and growth rate with respect to passage number. Whether organoid growth beyond 10 passages would establish a reliable steady-state is unknown.

To our knowledge, we are the first to characterize the gene expression of stem cell markers LGR5 and TROY in gastric organoids starting from tissue dissociation through to subsequent passages. Given that organoid media purposefully drives stem cell growth it is possible that organoids could become enriched in stem cells over successive passages. Our findings suggest that drug assay response is associated with passage number and/or stem cell gene expression. In either case, any effects on dose-response efficacy would need to be considered in translational organoid models to better approximate true *in-vivo* efficacy. Indeed, $LGR5^+$ and $TROY^+$ cells may give rise to cancer stem cells.[255,256] Increased LGR5 in human gastric cancer is associated with worse overall survival and decreased tumour response grade following neoadjuvant chemotherapy.[257] Inhibition of LGR5 expression in the AGS cell line was also associated with enhanced *in-vitro* dose response. Future characterization of the relationship between increasing organoid passages, stem cell gene expression and cytotoxic therapy efficacy is required.

3.5 Conclusions

In this study, we establish that shipment of fresh gastric tissue from mouse and human is a feasible and reliable method to increase procurement of primary organoid tissue. Our shipment method uses ice packs, standard shipping coolers and HBSS solution supplemented with 2% penicillin/streptomycin. Without the need for liquid nitrogen or minus 80-degree Celsius fridges, our method is relatively simple to implement in a clinical setting. In addition to feasibility, we show that shipment conditions do not affect the reliability of downstream dose-response assays and organoids maintain viability for at least 10 passages.

Chapter 4: Integration of independent comprehensive molecular classification frameworks provides novel insight into personalized medicine in gastric cancer

4.1 Introduction

Gastric cancer is the fifth most common cancer and the third most common cause of cancer death worldwide.[11–13] Over the previous three decades, gastric cancer diagnosis, treatment, and surgical management have significantly improved. However gastric cancer remains a deadly disease with a mortality-to-incidence ratio of 0.65 in developed countries. Comparatively few targeted therapy options are available for gastric cancer.[117,258,259] Questions remain regarding the implication of microsatellite instability on neoadjuvant treatment allocation.[7,260–263] Ongoing trials intend to extend immunotherapy options to the neoadjuvant setting.[1,2] One strategy to enhance the application of personalized medicine in the setting of increasingly complex therapeutic and clinical challenges is to leverage whole-transcriptome or multi-omics molecular classification.

The utility of omics technology in capturing relevant molecular subtypes in a variety of cancers has grown. Over the previous decade, multiple molecular classification systems on a cancer-specific and pan-cancer basis have been proposed using transcriptome, whole genome, epigenome and proteome data.[198–202,264,265] In gastric cancer, prominent classification systems include those proposed by The Cancer Genome Atlas project (TCGA), the Asian Cancer Research Group (ACRG) and the Tumour Microenvironment Score (TME).[26,82,207] These classification systems have been demonstrated to carry prognostic utility and may provide guidance for personalized therapy, especially in the case of microsatellite instability (MSI) and Epstein-Barr virus type (EBV) tumours.[83,206]

Clinical implementation of these classifiers faces multiple challenges. There has been ongoing development of novel classification systems without consideration of comparative analysis between them. Evaluating the true clinical relevance of these classification schemes is difficult given the small sample sizes due to the immense cost and logistics of transcriptome-related research. Translation to more feasible tests or existing cohorts is hindered by the curse of dimensionality, cross-platform heterogeneity of genomic measurements and fragmented nucleic acids related to clinical formalin-fixed paraffin-embedded samples.[266,267]

Even if these challenges are overcome we currently struggle to implement truly personalized research. To enhance the interpretability of complex cancer biology and mathematics, classification schemes often reduce molecular analyses to categorical subtypes. However, cancer biology is not simple or discrete. Worse, diminishing continuous biological processes to discrete categories is known to limit statistical power and conceal nuance.[268–270] Current models also fail to interpret estimates of prognosis or treatment effects at the individual level. In order to advance communication and discovery in personalized medicine novel models designed for personalization are desirable.

This study provides the largest comprehensive and integrated analysis of TCGA, ACRG and TME molecular classification systems for gastric cancer. We explore modelling techniques to enhance clinical translation of multi-omics classifiers, investigate methods to evaluate heterogeneity between classifiers and provide a translational driven analysis with a focus on personalized medicine models in gastric cancer. Principal among our findings is the beneficial effects of an active tumour immune microenvironment in prognosis and enhancing survival in patients who also receive chemotherapy.

4.2 Methods

4.2.1 Dataset descriptions

We used 11 publicly available gastric cancer datasets that contained RNA-seq or microarray whole transcriptome data.[26,57,208,271–276] Detailed information pertaining to all datasets including specific preprocessing procedures is available on GitHub (https://github.com/skubleny/Thesis-Supplemental-Files). Briefly, we identified datasets that were previously used in gastric cancer molecular classification studies.[26,82,207,208] Gene expression data was imported from Gene Expression Omnibus using GEOquery, cBioPortal or downloaded from published supplementary data.[26,277–279] Phenotype data containing clinicopathologic characteristics, including survival data and treatment status were retrieved using GEOquery ,TCGAbiolinks , cBioPortal or supplementary data files.[280–282]

Whole-transcriptome data was retrieved in a variety of formats. RNA-seq data from TCGA was quantified using RNA-Seq by Expectation-Maximization (RSEM) and batch normalized.[283] Affymetrix microarray data was retrieved after Robust Multi-array Average (RMA) normalization or was processed from raw CEL files followed by RMA normalization.[283] Illumina gene expression data was mainly retrieved as Log2 transformed and quantile normalized. Where applicable, normal or non-gastric adenocarcinoma samples were

excluded from subsequent data analysis and normalization steps. For GSE26899 and GSE26901, gene expression values with less than 20% missing data were imputed using k-nearest neighbours. Gene symbol names were assigned using AnnotationDBI.

For all datasets, we performed exploratory analysis using Bland-Altman plots, qqplots, boxplots and probability density functions to confirm proper normalization was achieved. We filtered low expression and low variance genes for the purpose of future translation to a Nanostring codeset.[284,285] A cut-off of Log2 gene count <1 was used for RNA-seq data. For microarray data histograms of gene expression row median values were visualized and low-intensity probes were filtered.[286] Following this procedure we achieved a common set of 9824 genes across 11 studies.

4.2.2 Supervised machine learning models

4.2.2.1 Model selection

We created a series of supervised machine learning classification models for TCGA, ACRG and TME subtypes (Figure 4-1). The training data was comprised of 376 TCGA, 300 ACRG and 1797 TME patients with molecular subtypes assigned by the primary publication. Models were generated using caret or glmnet in R.[227,228] We structured our experiment around three feature selection algorithms: LASSO, Gradient Boosting Machine (GBM) and Support Vector Machine with a linear kernel (SVM). After feature selection we implemented ElasticNet, naïve bayes (NB), random forest (RF), GBM, k-nearest neighbours (KNN), SVM and single layer neural network (NNET) models in R.[227,228,287–292]

Models were constructed in three nested layers. To mitigate estimation bias we used nested stratified 10-fold cross-validation (CV) in the outermost folds.[293] Given the presence of imbalanced classes in our data, stratification ensured that a consistent proportion of classes were present in each training and testing fold.[294] From the training data in the outermost layer, ten LASSO, GBM and SVM models were constructed using all 9824 gene features. Model performance was estimated on each respective hold-out dataset. Model selection and hyperparameter optimization for these LASSO, GBM and SVM models occurred within the second layer of 10-fold CV. We then performed feature selection using each of these models to decrease the number of genes required to classify subtypes. The lambda for LASSO feature selection was chosen using the minimum mean cross validated error for ACRG and TME classifiers and the lambda within one standard error of the minimum error for TCGA. For GBM models, the top 50 features were selected as defined by variable importance within

caret. For SVM, the top 15 features from each subtype class were selected as defined by variable importance in caret. Duplicate selected genes were removed. Next, the innermost layer was used to construct models using the reduced number of genes to predict subtypes for TCGA, ACRG and TME classes, respectively. This innermost fold used 5-fold CV with 3 resampling repeats to optimize hyperparameters. Performance of the selected models from this innermost layer was then estimated using the outermost hold-out test set.

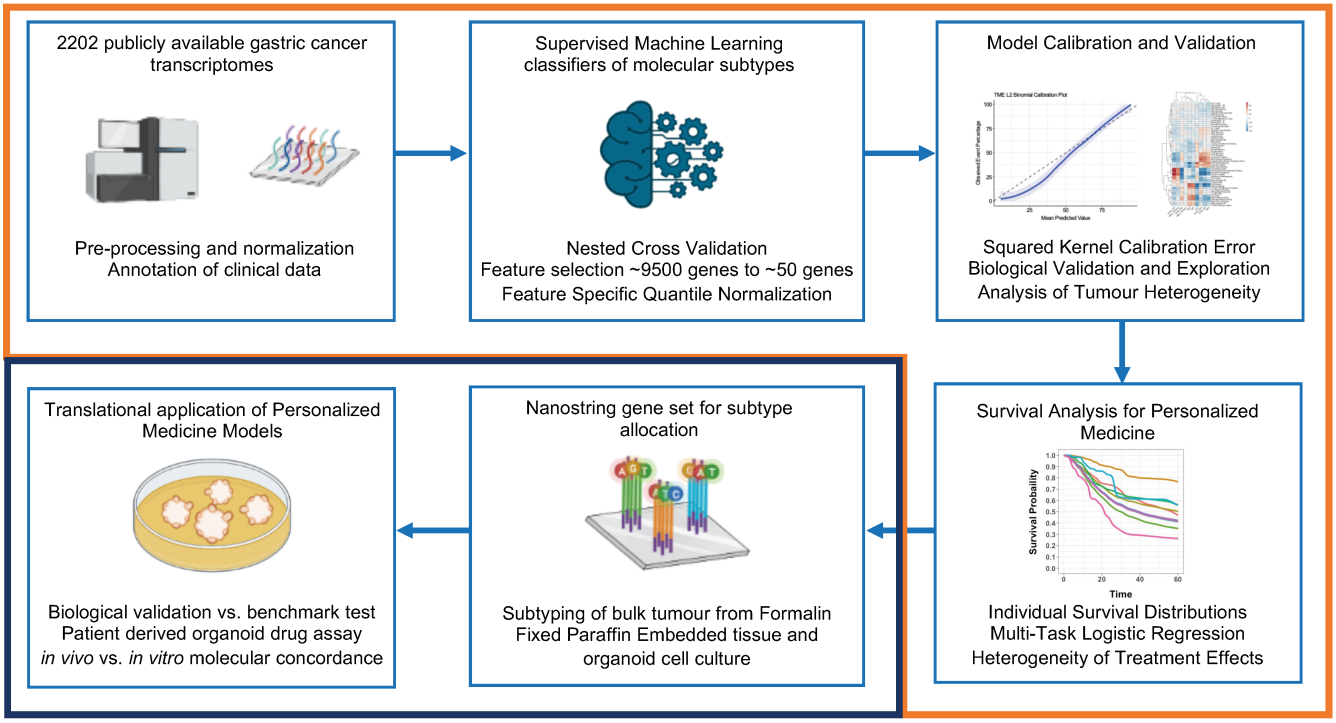


Figure 4-1 Study overview for Chapters 4 and 5. Each box illustrates a fundamental step to develop integrated molecular classification in gastric cancer with a Nanostring test. Components within Chapter 4 and 5 are denoted by the orange and blue box, respectively.

Final model performance was always estimated on unseen hold out data from the outermost layer. Allocation of subtype classes was assessed using confusion matrices and accuracy was used to select the final model. Given the presence of imbalanced subtype classes we used Cohen's Kappa to ensure model predictions were not due to random chance.[295] Brier score was also calculated to assess calibration using a proper scoring rule.[296]

4.2.2.2 Cross-platform normalization

We used Feature Specific Quantile Normalization (FSQN) to account for confounding related to differences in gene expression technologies.[297] For TME classification, FSQN was used to match training data from Illumina and RNA-seq platforms to the Affymetrix distribution in the ACRG cohort. FSQN was subsequently used to match gene expression distributions to the training distribution of each model (ie. Target distribution). For example, our TCGA classifier was trained using only RNA-seq data from the TCGA cohort. Prior to learning TCGA subtypes for all other datasets, gene expression values had to first be normalized to the TCGA distribution using FSQN.

4.2.3 Calibration Analysis

Calibration is a desirable feature for medical decision making. For every patient, each of our selected final models for TCGA, ACRG and TME classification produced categorical subtype assignments and a score corresponding to each subtype. These model scores may approximate calibrated probabilities, in which the frequency of actual events/subtypes corresponds to the score given by a model. We visually assessed the uncalibrated model scores using calibration plots.[298] We then applied post-hoc calibration methods to the model scores with the intent to improve calibration. For TCGA and ACRG model scores, we applied a novel multiclass calibration method called Dirichlet calibration.[299] Dirichlet calibration was implemented in R using log transformed model scores as input into L2 multinomial regression and penalized multinomial regression in the glmnet and nnet libraries, respectively.[228,291] For binary TME scores, we applied Platt scaling, which produced model singularities.[300] To mitigate this problem, we applied a modified version of Platt scaling using L2 regularized binomial regression on model scores. We used the same stratified 10-fold CV folds to asses calibration metrics and optimize hyperparameters.

To quantitatively assess calibration, we used various calibration metrics. For TME models we used Estimated Calibration Error (ECE) and Maximum Calibration Error (MCE), which are defined in Guo et al.[301] For multiclass models, we used classwise-ECE (cw-ECE), classwise-MCE (cw-MCE), confidence-ECE (conf-ECE) and confidence-MCE (conf-MCE) as defined by Kull et al.[299,301] Finally, we calculated a novel metric proposed by Widmann et al. called Squared Kernel Calibration Error (SKCE), which provides an unbiased and consistent metric to evaluate multiclass and binary calibration.[302] Optimally calibrated models were selected as final models.

4.2.4 Variance Heterogeneity

We developed a novel measurement of molecular classification heterogeneity called variance heterogeneity (varHet). Variance heterogeneity is a continuous variable and may be applied to one or more classification systems as single classification variance heterogeneity (sc-varHet) or integrated variance heterogeneity (int-varHet), respectively.

For the model scores derived from a single classifier model, *j*, we define a m x n matrix where m rows represent individual patients and $n_1 \ldots n_i$ columns represent score probabilities for *i* subtypes in a given classifier. For example, scores for a single patient, m, assigned by a classifier with 4 subtypes is represented by the matrix:

$$[0.054 \quad 0.098 \quad 0.126 \quad 0.722]$$

For reference, an example of scores derived from our models is given in Appendix .D.1. Next, we define *rowvar* as the variance (s^2) for a single row. Then class corrected *rowvar* (ccRowVar) for a single patient, m, is:

Equation 4-1

$$ccRowVar = rowvar \times i$$

For a single patient, m, single classification variance heterogeneity (sc-varHet) is:

Equation 4-2

$$sc - varHet = -log(ccRowVar)$$

For a single patient, m, integrated variance heterogeneity sums individual ccRowVar values, as defined in Equation 4-1, for *j* independent classifiers. Thus, integrated variance heterogeneity is defined by:

Equation 4-3

$$int - varHet = -log\left(\frac{1}{j}\sum ccRowVar_j\right)$$

Given this framework the theoretical minimum of sc-varHet and int-varHet is 0, which is given by high variance in assigned subtype scores, and the maximum heterogeneity approaches infinity with decreasing variance in subtype scores. Simply, a classifier that confidently predicts a subtype in a given tumour with a high score (ie. For a binary setting, $Subtype_A$ = 0.99 and $Subtype_B$ = 0.01) has low variance heterogeneity, whereas a tumour with nearly equal scores has high variance heterogeneity (ie. For a binary setting, $Subtype_A$ = 0.49 and $Subtype_B$ = 0.51)

4.2.5 Model validation in reference to gold-standard features

For the purposes of this paper, internal validation of our models was performed using publicly available whole-transcriptome data. Within our definition we still assess the ability of classifiers to capture gold-standard reference features on seen and unseen data. We reserve the term external validation for future development of a targeted stand-alone test that does not incorporate whole-transcriptome data.

Briefly, gold-standard features were available to validate TCGA EBV, TCGA MSI and ACRG MSI subtypes. A feature developed by the ACRG for evaluating genomic instability termed CNV GI was used to approximate aneuploidy, which should be captured by TCGA Chromosomal instability (CIN) and ACRG microsatellite stable TP53 negative (MSS TP53-) subtypes. CNV-GI is a binary variable defined as "high" if the number of deleted or amplified chromosomes is >2.3 x the population's median average deviation in each chromosome.[18] Using subtype scores derived from our supervised classifiers we performed logistic regression to predict the allocation of gold-standard features. Area under the receiver operating characteristic (AUROC) was calculated and receiver operating characteristic (ROC) curves were plotted from pROC in R.[303]

4.2.6 Propensity score matched analysis

We performed a propensity score matched analysis to infer the causal treatment effects of chemotherapy on overall survival among our selected molecular subtypes. Briefly, we used logistic regression to calculate the propensity score of receiving chemotherapy with confounding variables including molecular subtypes, age, stage, tumour location, Lauren classification and study identification.[304] To optimize matching we tested non-linear relationships of confounding variables as a function over time using likelihood ratio tests (Appendix Table D2-1; Appendix Figure D2-10). Age was fit as a natural cubic spline with

two degrees of freedom. Nearest matching at a 4:1 ratio with a caliper was performed using the average treatment effect on the controls as the target estimand.

Marginal treatment effects on overall survival were estimated with the survival package using robust standard errors. Sampling weights and cluster identification was appropriately included. We performed sensitivity analyses using separate matched cohorts comprised of various stages of disease. Furthermore, sensitivity of model effects to unmeasured confounding was assessed using the E-Value.[305]

4.2.7 Statistical and Survival Analysis

Statistical analyses were completed using R version 4.1.2.[226] Squared Kernel Calibration Error estimates and tests were executed in Julia version 1.5.4, but implemented in R using the rcalibration package.[302] Differences between groups were assessed with a Wilcoxon two-sample test or Kruskal-Wallis test. When applicable, post-hoc tests were completed using Dunn's test. Multiple comparisons corrections were made using either Benjamini-Hochberg or Bonferroni methods as stipulated in the results or figures. Statistical significance was defined at alpha = 0.05.

Overall survival of independent and integrated molecular classification was assessed using Cox Proportional Hazards model with the survival package.[306] Treatment effects in propensity score matched survival analysis were estimated by accounting for cluster stratification and robust standard errors. Interaction effects were visualized using the simPH package.[307] Non-nested Cox models were assessed using corrected Akaike Information Criterion (AICc), with a significant improvement defined as a decrease in AICc greater than 10.

Individual survival distribution models were implemented as suggested by Haider et al (https://github.com/haiderstats/ISDEvaluation).[308] Briefly, we tested the performance of Cox with Kalbfleisch-Prentice extensions (Cox-KP), ElasticNet Cox (CoxEN-KP), Random Survival Forest , Accelerated Failure Time and Multi-task logistic regression (MTLR) models for our survival prediction task. Model performance and optimization was assessed using 5-fold CV. Models were selected in consideration of the Concordance Index (C-Index), 1-calibration and D-calibration.

4.3 Results

4.3.1 Performance and selection of supervised TCGA, ACRG and TME molecular classification models

Figure 4-2A, C and E show the accuracy of all twenty-four supervised machine learning models constructed for TCGA, ACRG and TME classification systems, respectively. Our optimal TCGA classifier achieved an accuracy of 89.5% ± 0.04 standard deviation using 57 gene features selected by LASSO within an ElasticNet model. Classification accuracy of ACRG subtypes was maximized at 84.7% ± 0.04 using a GBM model with 39 gene features selected from an SVM model. We attained an accuracy of 89.3% ± 0.02 for binary TME classification using the top 50 GBM features in a random forest model. Overall, most model combinations performed well across the three classification tasks and little variance occurred in model performance between folds. In all cases, optimal feature selected models achieved near equivalent or superior accuracy, kappa and Brier scores compared to models using all 9824 genes. Selected genes and numerical metrics for each model combination are provided in the online supplement (https://github.com/skubleny/Thesis-Supplemental-Files).

Next, we statistically assessed critical differences in accuracy among model combinations for each prediction task (Appendix .D.3). Poor model combinations were dominated by KNN and NB models using GBM and SVM features for TCGA and ACRG multiclass predictions. Superior model accuracy for TME classification was dominated by feature selection using GBM and LASSO in combination with RF, GBM, KNN and NNET models.

4.3.2 Calibration of molecular classification models

We applied Dirichlet calibration and Platt scaling to our chosen classification models. In Figure 4-2B, D and F we present the model performance and calibration metrics for the uncalibrated and calibrated models averaged over 10-fold CV. We also calculated the unbiased Square Kernel Calibration Error for each model (Figure 4-2G).

In keeping with Kull et al., Dirichlet calibration improved class-wise ECE definitions for multiclass ACRG and TCGA models (Kruskal-Wallis, $p=0.004$ and $p<0.001$, respectively).[299] For ACRG models, calibration methods significantly decreased class-wise MCE but also increased ECE of the dominant class (Figure 4-2D). Calibration plots for ACRG models show that L2 Dirichlet calibration resulted in less variance in mean calibration error for all classes and modified the calibration curves for individual classes (Appendix Figure D4-20).

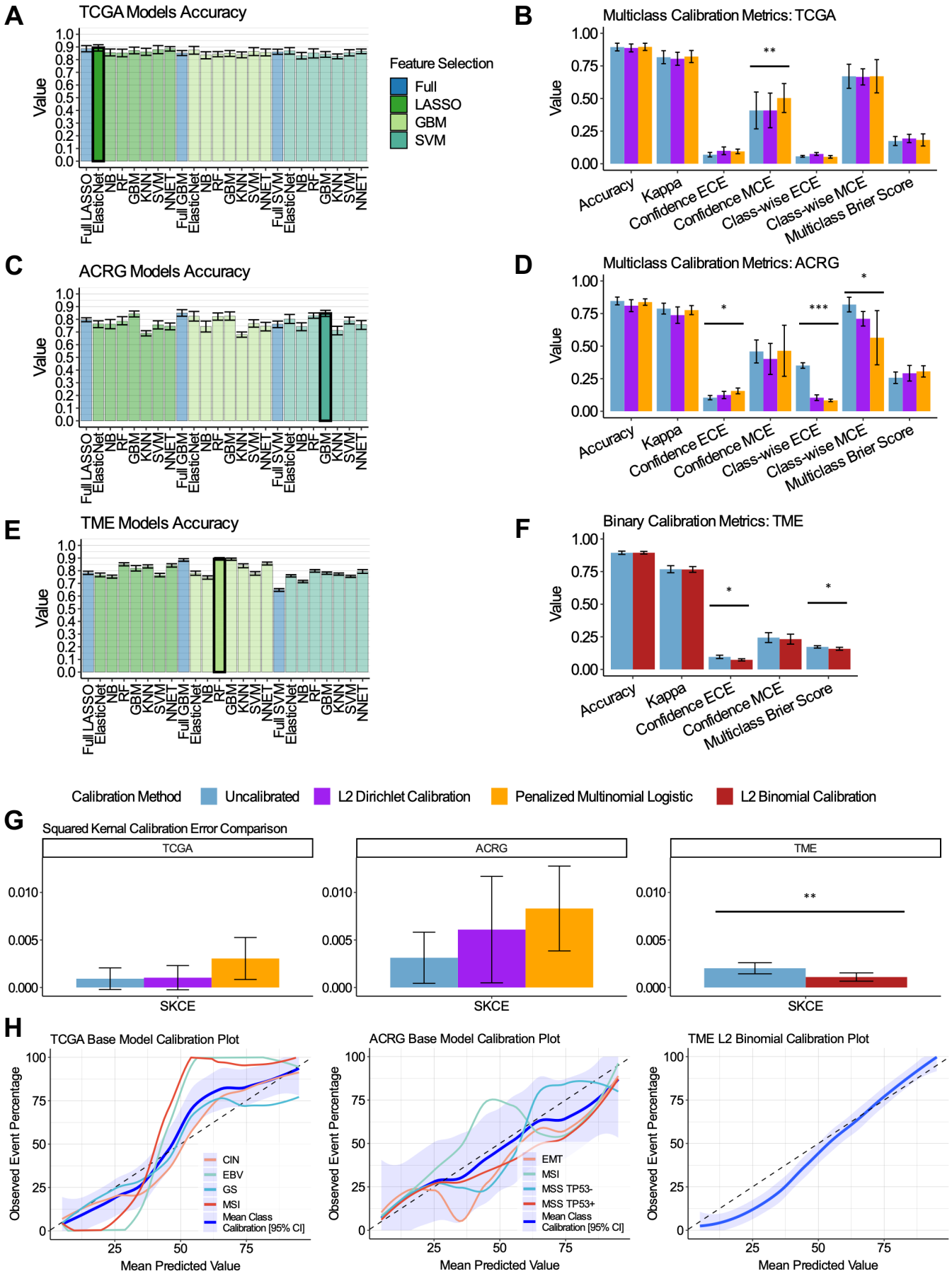

A
TCGA Models Accuracy
Value
Feature Selection
Full
LASSO
GBM
SVM
B
Multiclass Calibration Metrics: TCGA
Accuracy
Kappa
Confidence ECE
Confidence MCE
Class-wise ECE
Class-wise MCE
Multiclass Brier Score
C
ACRG Models Accuracy
D
Multiclass Calibration Metrics: ACRG
E
TME Models Accuracy
F
Binary Calibration Metrics: TME
Calibration Method
Uncalibrated
L2 Dirichlet Calibration
Penalized Multinomial Logistic
L2 Binomial Calibration
G
Squared Kernal Calibration Error Comparison
TCGA
ACRG
TME
SKCE
H
TCGA Base Model Calibration Plot
ACRG Base Model Calibration Plot
TME L2 Binomial Calibration Plot
Observed Event Percentage
Mean Predicted Value
CIN
EBV
GS
MSI
EMT
MSS TP53-
MSS TP53+
Mean Class Calibration [95% CI]

Figure 4-2 Performance and calibration of supervised machine learning classifiers. **A, C and E.** Mean accuracy estimated using 10-fold nested cross-validation for TCGA (A), ACRG (C) and TME (E) models. Bolded bars represent chosen models. Colours denoted by the plot legend show Full models (9824 genes) and models after feature selection using LASSO, GBM and SVM. **B, D and F.** Mean calibration metrics as well as accuracy and Cohen's kappa values of uncalibrated and calibrated models for TCGA (B), ACRG (D) and TME (F) classification. **G.** Mean Squared Kernel Calibration Error of uncalibrated and calibrated models. Calibration methods denoted by colours in the plot legend. For A-G, mean accuracy, Cohen's kappa and calibration was calculated using 10-fold nested cross validation. Error bars represent 95% confidence intervals computed using 1000 bootstraps. Horizontal bars demonstrate statistical significance using a Kruskal-Wallis test (***p<0.001, **p<0.01, *p<0.05). **H.** Calibration plots for final selected models including uncalibrated base TCGA model (Left), uncalibrated base ACRG model (Middle) and modified Platt calibrated TME model (Right). Calibration was assessed using deciles and visualized using a loess smooth with span = 0.75. The mean calibration for all subtypes is represented in blue with 95% confidence intervals in light blue. Where applicable, the figure legends describe colour identification for each molecular subtype.

Table 4-1 Patient demographics

Variable	**n/N (Missing %)**	**N = 2,202[1]**
Age	2,128 / 2,202 (3.4%)	61 (52, 69)
Stage	1,678 / 2,202 (24%)	
I		258 (15%)
II		484 (29%)
III		590 (35%)
IV		346 (21%)
Sex	2,132 / 2,202 (3.2%)	
Female		714 (33%)
Male		1,418 (67%)
TCGA Subtype	2,202 / 2,202 (0%)	
CIN		1,631 (74%)
EBV		124 (5.6%)
GS		189 (8.6%)
TCGA MSI		258 (12%)
ACRG Subtype	2,202 / 2,202 (0%)	
EMT		278 (13%)
MSI		372 (17%)
MSS TP53-		850 (39%)
MSS TP53+		702 (32%)
TME Subtype	2,202 / 2,202 (0%)	
High		731 (33%)
Low		1,471 (67%)
Study	2,202 / 2,202 (0%)	
ACRG		300 (14%)
Kosin		109 (5.0%)
KUGH		93 (4.2%)
MDACC		40 (1.8%)
Samsung		432 (20%)
Shanghai		70 (3.2%)
Singapore		248 (11%)
TCGA		412 (19%)
Yonsei		433 (20%)
Yonsei MDACC		65 (3.0%)
Grade	513 / 2,202 (77%)	
G1		32 (6.2%)
G2		166 (32%)
G3		286 (56%)
Other		29 (5.7%)
Lauren Class	1,471 / 2,202 (33%)	
Intestinal		760 (52%)
Diffuse		642 (44%)
Mixed		69 (4.7%)
Signet Ring	510 / 2,202 (77%)	
Signet Ring		83 (16%)
No		381 (75%)
Other		46 (9.0%)
Tumour Location	1,389 / 2,202 (37%)	
Proximal		692 (50%)
Distal		673 (48%)
Whole		24 (1.7%)
Chemotherapy	1,371 / 2,202 (38%)	957 (70%)
Radiation	654 / 2,202 (70%)	152 (23%)

[1]Median (IQR); n (%)

Penalized multinomial regression using nnet severely distorted calibration curves. In contrast to traditional histogram-based metrics such as ECE and MCE, the unbiased SKCE demonstrated that calibration methods

for multiclass models actually worsens calibration, albeit non-significantly (Figure 4-2G Left and Middle). Given these findings we selected the uncalibrated TCGA and ACRG models.

Regularized Platt scaling of our TME model significantly improved ECE and Brier score compared to the uncalibrated model (Wilcox, p=0.011 and p=0.043, respectively) (Figure 4-2F). The calibrated TME model also resulted in significantly less SKCE (Wilcox, p = 0.007) (Figure 4-2G Right). Calibration plots show that modified Platt scaling straightened the sigmoidal uncalibrated curve to better approximate perfect calibration, especially for underconfident prediction scores greater than 50 (Appendix Figure D4-21). L2 binomial calibration also improved model accuracy to 89.4% ± 0.01, and was thus selected as our final model.

Next, we assessed whether our models were actually calibrated using a statistical test. We performed an Asymptotic SKCE test under the null hypothesis that our model is perfectly calibrated (ie. SCKE = 0). All configurations of model calibration for each classification model was statistically significant and thus our models were not calibrated.

Calibration methods maintained similar accuracy and kappa performance compared to uncalibrated TCGA, ACRG and TME models. Calibration plots generated over 10 histogram bins for our selected final models are illustrated in Figure 4-2H. The TCGA model achieved the lowest SKCE value relative to other models despite greater observed class-wise calibration error and mean class calibration error.

4.3.3 Gastric cancer possesses significant molecular subtype heterogeneity

Next, we used our selected models to learn molecular subtypes for all 2,202 patients included in this study. Feature Specific Quantile Normalization was performed to the target distribution of each model. Demographic data for all patients is provided in Table 4-1 and detailed demographic data for each molecular subtype is in Appendix Table D5-2, Table D5-3 and Table D5-4.

The median patient age was 60 years (52,69) and 67% were male. Stage had greater than 20% missing data and grade, Lauren class, signet ring and tumour location contained at least 30% missing data. Age and stage were significantly different between subtypes within TCGA, ACRG and TME classification systems (Kruskal-Wallis and Chi-squared, $p<0.05$). We found that sex, grade, Lauren class and tumour location was significantly different between TCGA and ACRG subtypes but similar between high immune microenvironment scores (TME

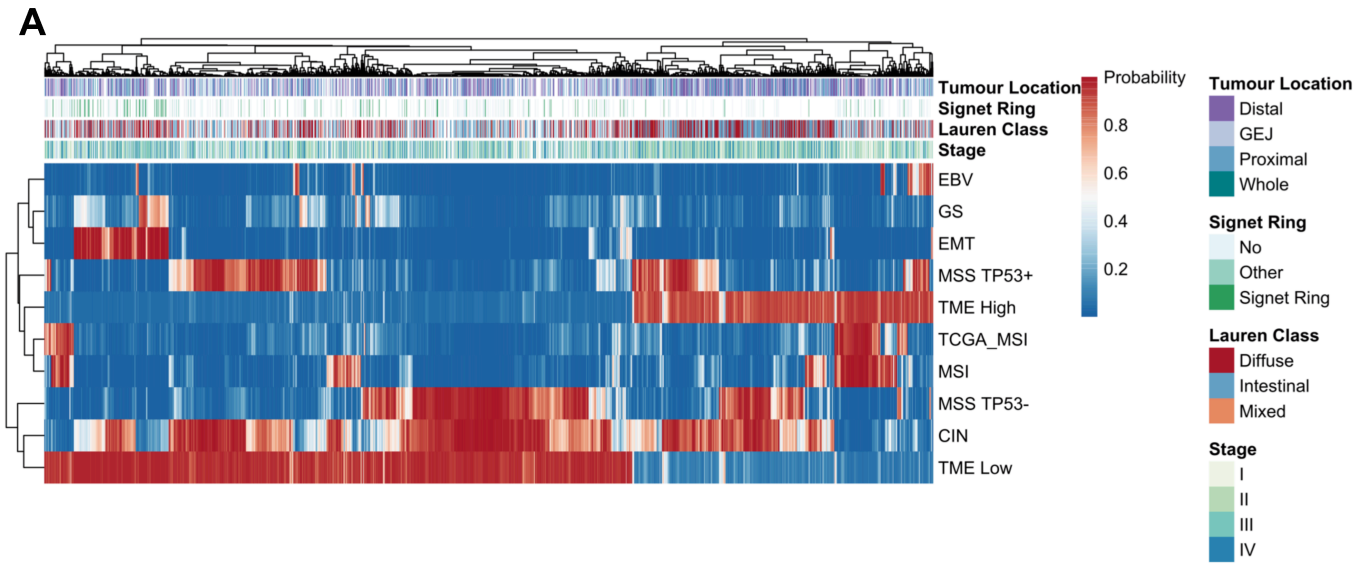
High) and low immune microenvironment score (TME Low) tumours (Chi-squared and Fisher's Exact Tests). These findings suggest that immune based classification is agnostic to traditional clinicopathologic factors, whereas the ACRG and TCGA scores capture a greater degree of epithelial derived signal.

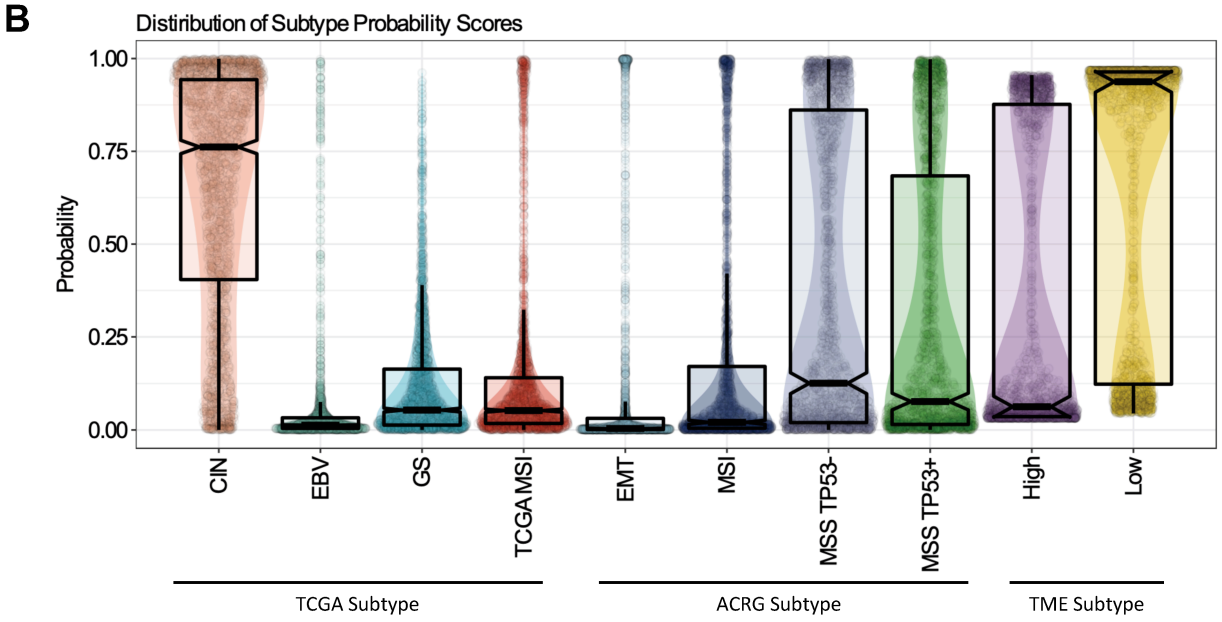
Table 4-2 Contingency table for TCGA and ACRG subtypes stratified by TME status

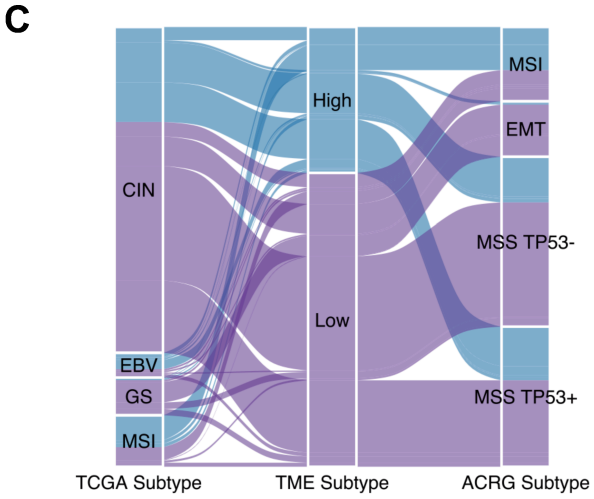
TME High (n=731)					
	EMT	MSI	MSS TP53-	MSS TP53+	TCGA Sum
CIN	12	66	199	199	476
EBV	5	22	9	46	82
GS	0	3	3	7	13
TCGA MSI	0	125	18	17	160
ACRG Sum	17	216	229	269	

TME Low (n=1471)					
	EMT	MSI	MSS TP53-	MSS TP53+	TCGA Sum
CIN	148	72	572	363	1155
EBV	8	9	6	19	42
GS	102	10	33	31	176
TCGA MSI	3	65	10	20	98
ACRG Sum	261	156	621	433	

At a categorical level, CIN was the most prevalent TCGA subtype (1,631 patients, 67%). Comparatively, 124 (5.6%) EBV, 189 (8.6%) Genomically stable (GS) and 258 (12%) TCGA MSI tumours were present. In decreasing prevalence, 850 (39%) MSS TP53-, 702 (32%) Microsatellite Stable TP53 positive (MSS TP53+), 372 (17%) MSI and 278 (13%) Epithelial to Mesenchymal Transition (EMT) ACRG subtypes were assigned. TME High and TME Low tumours comprised 731 (33%) and 1471 (67%) of patients, respectively. In Figure 4-3C, we illustrate the heterogeneity of assigned subtype categories using an alluvial plot. We found that all possible combinations of molecular subtypes exist among 2,202 patients. Prominent trends include the strong association of CIN tumours with MSS TP53+ and MSS TP53- tumours, as well as the paucity of TME High tumours in GS and EMT subtypes. The discrete heterogeneity of TCGA and ACRG subtypes stratified by TME High and TME Low is demonstrated in Table 4-2.


A
Tumour Location
Signet Ring
Lauren Class
Stage
EBV
GS
EMT
MSS TP53+
TME High
TCGA_MSI
MSI
MSS TP53-
CIN
TME Low
Probability
0.8
0.6
0.4
0.2
Tumour Location
Distal
GEJ
Proximal
Whole
Signet Ring
No
Other
Signet Ring
Lauren Class
Diffuse
Intestinal
Mixed
Stage
I
II
III
IV


B
Distiribution of Subtype Probability Scores
Probability
1.00
0.75
0.50
0.25
0.00
CIN
EBV
GS
TCGA MSI
EMT
MSI
MSS TP53-
MSS TP53+
High
Low
TCGA Subtype
ACRG Subtype
TME Subtype


C
CIN
EBV
GS
MSI
High
Low
MSI
EMT
MSS TP53-
MSS TP53+
TCGA Subtype
TME Subtype
ACRG Subtype

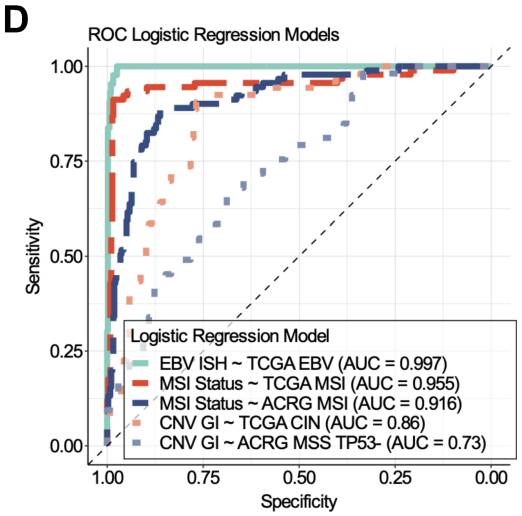

D
ROC Logistic Regression Models
Sensitivity
1.00
0.75
0.50
0.25
0.00
Specificity
Logistic Regression Model
EBV ISH ~ TCGA EBV (AUC = 0.997)
MSI Status ~ TCGA MSI (AUC = 0.955)
MSI Status ~ ACRG MSI (AUC = 0.916)
CNV GI ~ TCGA CIN (AUC = 0.86)
CNV GI ~ ACRG MSS TP53- (AUC = 0.73)

Figure 4-3 Molecular subtype model scores and categorical subtypes possess heterogeneity. **A.** Heatmap of optimal calibration probability scores for each patient. Hierarchical clustering was performed using complete linkage on Euclidean distance. Select clinicopathologic characteristics are annotated above the heatmap. The legend illustrates categorical representation of each characteristic. Magnitude of class probability is denoted by colour scale legend. **B.** Boxplot-scatter-violin plots for the distribution of subtype probability scores for each molecular subtype in TCGA, ACRG and TME classification systems. Boxplot notches approximate the 95% confidence interval of the median. Each data point represents a single measurement. The distribution of data is represented by superimposed violin plots. **C.** Alluvial plot illustrating flow of discrete molecular classification labels between TCGA, TME and ACRG classification schemes for 2,202 patients. Blue and purple identifies TME High and TME Low tumours, respectively. Each column contains 2,202 patients. **D.** Receiver Operating Characteristic Curve of logistic regression models predicting biological features external to the transcriptome using class probability scores. Dependent and independent variables (ie. Dependent ~ independent) of the model with corresponding Area Under the Curve values (Accuracy) are denoted in the plot legend.

In keeping with our intent to represent biology as a continuous variable we plotted the model subtype probability scores in Figure 4-3A and B. The heatmap using complete clustering of probability scores on Euclidean distance revealed relationships between different classification systems. For example, MSS TP53-, CIN and TME Low tumours formed the most prominent cluster. We found that GS and EMT tumours were closely related but also clustered with high EBV subtype scores. Considering previously characterized poor prognosis of GS and EMT tumours relative to EBV tumours, this was a prominent finding.[26,83,206] No major patterns of clustering coincided with tumour location, stage or Lauren classification. However, a predominance of signet ring cell tumours was appreciated to cluster with greater GS and EMT subtype scores. In Figure 4-3B we illustrate that large variation exists between subtype scores and that the shape of subtype score distributions is heterogenous.

4.3.4 Molecular subtype scores emulate intended gold-standard molecular tests

Despite accurate model performance that was assessed using cross-validation we wanted to ensure our subtype probability scores appropriately captured their target phenotype. We extracted available gold-standard data for validation of EBV and MSI subtypes measured by Epstein-Barr encoded small RNAs in-situ hybridization (EBV ISH) and multiplex PCR according to Bethesda criteria (MSI Status).[309–311] We used the CNV GI score established in the ACRG study to capture the presence of aneuploidy because data were annotated to both

TCGA and ACRG cohorts. Importantly, the selected gold-standard references measured molecules external to the transcriptome to prevent any possible data leakage intrinsically related to gene expression data.

We created Logistic regression models for each intended target and our model subtype scores (Figure 4-3D). All subtype scores provided statistically significant value to their intended targeted phenotype. Subtype scores for EBV type cancer provided almost perfect assignment of positive EBV infection as defined by EBER (AUC 0.997 [95% CI 0.994, 1], p<0.001). Both TCGA MSI and ACRG MSI scores performed well in recognizing MSI status (TCGA MSI AUC 0.955 [95% CI 0.923,0.987], p<0.001; ACRG MSI AUC 0.916 [95% CI 0.884, 0.948], p<0.001). TCGA CIN scores better approximated CNV GI status compared to MSS TP53- scores (CIN AUC 0.860 [95% CI 0.805,0.915], p<0.001; MSS TP53- AUC 0.732 [95% CI 0.656, 0.808], p<0.001). These findings provide additional evidence that we developed effective and biologically consistent classifier models.

4.3.5 Variance heterogeneity provides a flexible and useful heterogeneity metric for probabilistic classifier models

Given the heterogeneity identified between molecular subtypes with both discrete and continuous variables, we developed and applied a novel measurement of molecular classification heterogeneity. First, we replicated the Intratumor Heterogeneity (ITH) score proposed by Marisa et al. to the TCGA and ACRG scores (Appendix .D.6).[312] This metric reduces heterogeneity identified in continuous model probability scores to an ordered categorical representation of heterogeneity. In a framework comprised of 4 subtype classes, the ITH score for a given patient is derived from the number of subtypes with a probability score > 0.2.

We found that ITH failed to provide significant prognostic yield in our patients (Appendix Figure D4-21A and C). In contrast to Marisa et al., increasing ITH scores trended towards improved overall survival for TCGA and ACRG subtypes. Furthermore, the ITH score was designed to be implemented to a single classification system and was based on arbitrary cut-off values. This motivated us to develop a continuous metric of heterogeneity based on the same principle – that decreasing variance in the scores assigned to a given patient reflects the presence of multiple molecular subtype signatures within a single tumour.

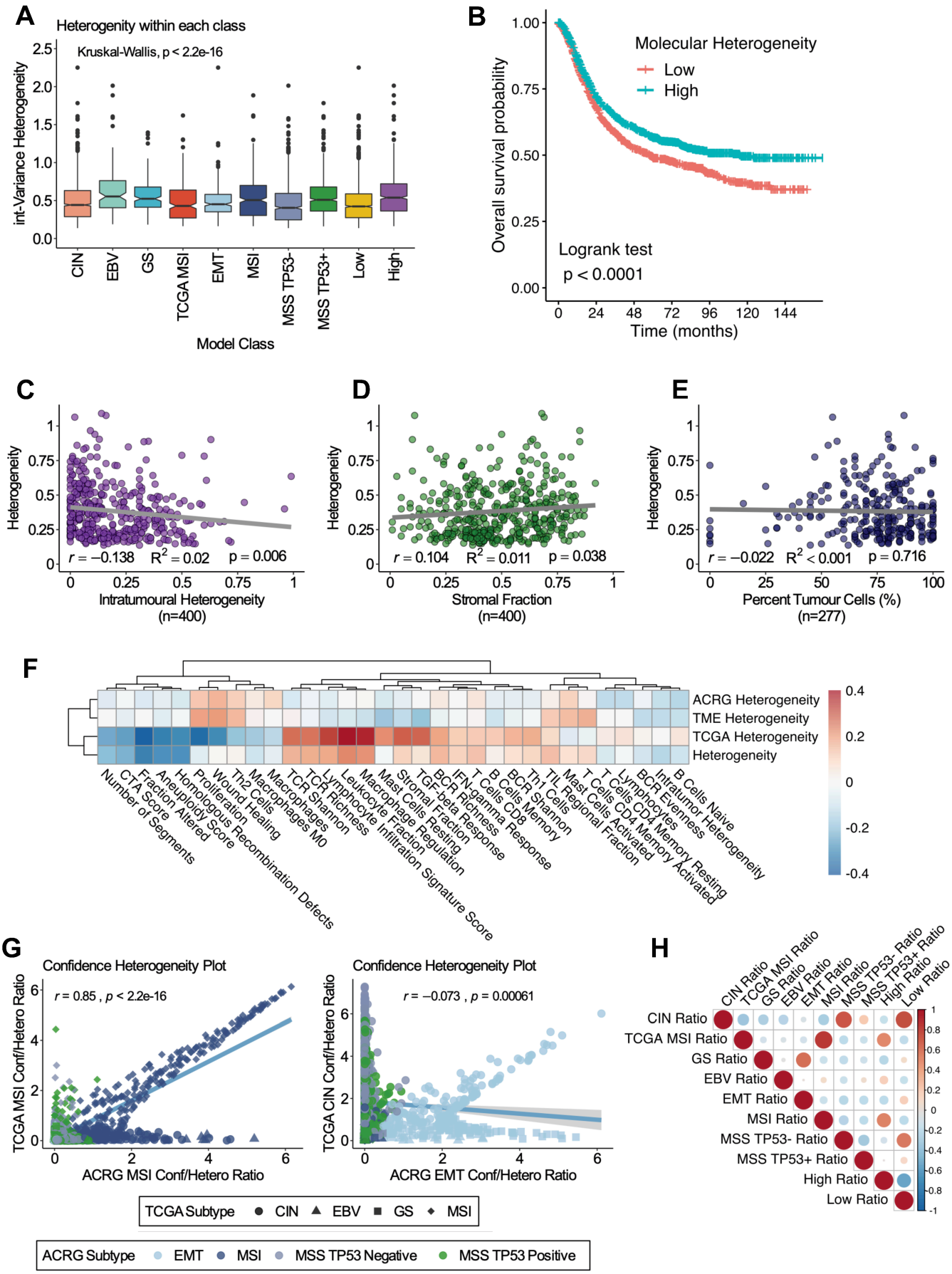

A
Heterogenity within each class
Kruskal-Wallis, p < 2.2e-16
int-Variance Heterogeneity
CIN
EBV
GS
TCGA MSI
EMT
MSI
MSS TP53-
MSS TP53+
Low
High
Model Class
B
Molecular Heterogeneity
Low
High
Overall survival probability
Logrank test
p < 0.0001
Time (months)
C
Heterogeneity
r = −0.138 R2 = 0.02 p = 0.006
Intratumoural Heterogeneity
(n=400)
D
Heterogeneity
r = 0.104 R2 = 0.011 p = 0.038
Stromal Fraction
(n=400)
E
Heterogeneity
r = −0.022 R2 < 0.001 p = 0.716
Percent Tumour Cells (%)
(n=277)
F
ACRG Heterogeneity
TME Heterogeneity
TCGA Heterogeneity
Heterogeneity
Number of Segments
CTA Score
Fraction Altered
Aneuploidy Score
Homologous Recombination Defects
Proliferation
Wound Healing
Th2 Cells
Macrophages
Macrophages M0
TCR Shannon
TCR Richness
Lymphocyte Infiltration Signature Score
Leukocyte Fraction
Macrophage Regulation
Mast Cells Resting
Stromal Fraction
TGF-beta Response
BCR Richness
IFN-gamma Response
T Cells CD8
B Cells Memory
BCR Shannon
Th1 Cells
TIL Regional Fraction
T Cells CD4 Memory Activated
Mast Cells Activated
T Cells CD4 Memory Resting
Lymphocytes
BCR Evenness
Intratumor Heterogeneity
B Cells Naive
G
Confidence Heterogeneity Plot
r = 0.85 , p < 2.2e-16
TCGA MSI Conf/Hetero Ratio
ACRG MSI Conf/Hetero Ratio
Confidence Heterogeneity Plot
r = −0.073 , p = 0.00061
TCGA CIN Conf/Hetero Ratio
ACRG EMT Conf/Hetero Ratio
TCGA Subtype CIN EBV GS MSI
ACRG Subtype EMT MSI MSS TP53 Negative MSS TP53 Positive
H
CIN Ratio
TCGA MSI Ratio
GS Ratio
EBV Ratio
EMT Ratio
MSI Ratio
MSS TP53- Ratio
MSS TP53+ Ratio
High Ratio
Low Ratio

Figure 4-4 Variance based molecular class heterogeneity is biologically diverse and an independent prognostic factor. **A**. Boxplot demonstrating that molecular class heterogeneity varies between molecular subtypes. Boxplot notches approximate the 95% confidence interval of the median. **B.** Kaplan-Meier survival probability curve from a Cox Proportional Hazards model comparing High and Low Heterogeneity tumours. Heterogeneity (int-varHet) was divided at the median heterogeneity score. A Log Rank test was performed with $p < 0.0001$. **C.** Heterogeneity (int-varHet) versus Intratumoral Heterogeneity (subclonal heterogeneity) calculated by ABSOLUTE. **D.** Heterogeneity (int-varHet) versus Stromal Fraction defined as 1 minus tumour purity calculated with ABSOLUTE. **E.** Heterogeneity (int-varHet) versus percent of tumour cells observed on hematoxylin and eosin stained tissue sections. **F.** Heatmap of Pearson's correlation coefficients of heterogeneity scores versus molecular and immune feature scores of TCGA samples established by The Immune Landscape of Cancer. Hierarchical clustering was performed using complete linkage. **G.** Similarity of molecular classes can be evaluated using the ratio of confidence (class probability) and heterogeneity (variance heterogeneity). For example, TCGA MSI vs. ACRG MSI (Left) and TCGA CIN vs. ACRG EMT (Right). The figure legends below denote the shape and colour corresponding to TCGA and ACRG subtypes, respectively. **H.** Pearson's Correlation plot of confidence heterogeneity ratios for all classes, which demonstrates biological similarity of GS and EMT, MSI and MSI, CIN and MSS TP53 negative, CIN and TME low and MSI and TME High. For C, D and G the gray/blue line represents the slope of simple linear regression corresponding to Pearson's correlation coefficient (*r).* R^2 describes the proportion of variance in Heterogeneity explained by Intratumoral Heterogeneity. P value denotes statistical significance with alpha level = 0.05. For F and H, the magnitude and direction of Pearson's correlation coefficient is denoted by the colour scale legend.

We calculated the sc-varHet and int-varHet scores. Our sc-varHet scores for TCGA and ACRG closely approximated ITH scores (Appendix Figure D4-21B and D). The distribution of int-varHet were significantly different between molecular subtypes, but generally approximated a median value of 0.5 (Figure 4-4A). Consistent with our findings for ITH scores, patients with int-varHet scores greater than the median value of 0.4574 had significantly greater overall survival (Figure 4-4B).

Our definition of heterogeneity is based on the probabilistic scores derived from supervised learning models. We assessed our definition of heterogeneity relative to previously defined perspectives which included tumour purity/sub-clonal heterogeneity, stromal tumour fraction and the relative of abundance of tumour cells present.[313] We observed minimal Spearman correlation values relative to int-varHet for all three metrics (Figure 4-4C, D and E). Despite statistical significance, the magnitude of effect was small, suggesting that int-varHet provides a distinct perspective into intratumour heterogeneity.

Next, we characterized the association of variance heterogeneity metrics for each molecular classification system and integrated heterogeneity (Figure 4-4F). Pearson's correlation with tumour molecular characteristics specific by the PanCancer Atlas and Immune Landscape of Cancer projects revealed that individual classification heterogeneity corresponds to unique biological phenomenon. For example, increasing TCGA based sc-varHet was associated with increasing macrophage regulation and leucocyte fractions, whereas ACRG and TME sc-varHet corresponded to increased proliferation and wound healing signatures. These findings suggested that we cannot assume a unidirectional effect of probabilistic model-based heterogeneity. Instead, heterogeneity within each classification system must be considered independently.

We continued to assess the utility of variance-based heterogeneity by considering its relationship to subtype probability scores between separate subtype classes. In Figure 4-4D, we demonstrate that the ratio between the model derived probability score and the int-varHet score provides insight into the degree of similarity in independent molecular subtypes. We tested this hypothesis using TCGA MSI and ACRG MSI subtype scores given their close recapitulation of true MSI status. A strong visual and statistical correlation was found between these similar molecular subtypes. In contrast, TCGA CIN, which is characterized by tumour somatic copy number alterations and ACRG EMT, which closely resembles TCGA-GS, had near zero correlation. The strength of similarity between all independent molecular subtypes is demonstrated in Figure 4-4H. These findings were similar to heatmap-based clustering in Figure 4-3A, however, using correlations between Confidence/Heterogeneity ratios provides a common strength of association and statistical metric.

4.3.6 Survival analysis

We expanded on prior prognostic interpretations of our selected molecular classification systems using our cohort of 2197 patients with overall survival data. Our cohort contained 992 events and 54.8% of patients were censored (Appendix Figure D7-23) We performed univariate analysis for each molecular subtype (Table 4-3, Table 4-4 and Table 4-5). Consistent with prior studies, EBV and TCGA MSI tumours possessed significantly greater overall survival compared to CIN and GS (Figure 4-5A). Compared to the original ACRG cohort, we found that MSS TP53+ and MSS TP53- tumours were more closely related in terms of prognosis. Indeed, pairwise comparisons with Benjamini-Hochberg corrections failed to provide statistically significant prognostic differences (Figure 4-5B). In Figure 4-5C,

we found that TME High tumours possessed large favourable prognostic effects relative to TME Low.

Next, we compared the predictive value of using continuous molecular subtype scores versus the traditional categorical interpretations (Figure 4-5D). We assessed non-nested models for each molecular classification and a multivariable Cox model containing all molecular subtypes using corrected Akaike's Information Criterion. Models using continuous interpretations of molecular subtypes achieved significant (defined as AICc > 10) improvement in model performance for TCGA, TME and integrated molecular classification models. Evaluation of ACRG subtype overall survival favoured continuous values to a lesser degree.

Table 4-3 Cox proportional hazards model for discrete TME subtypes

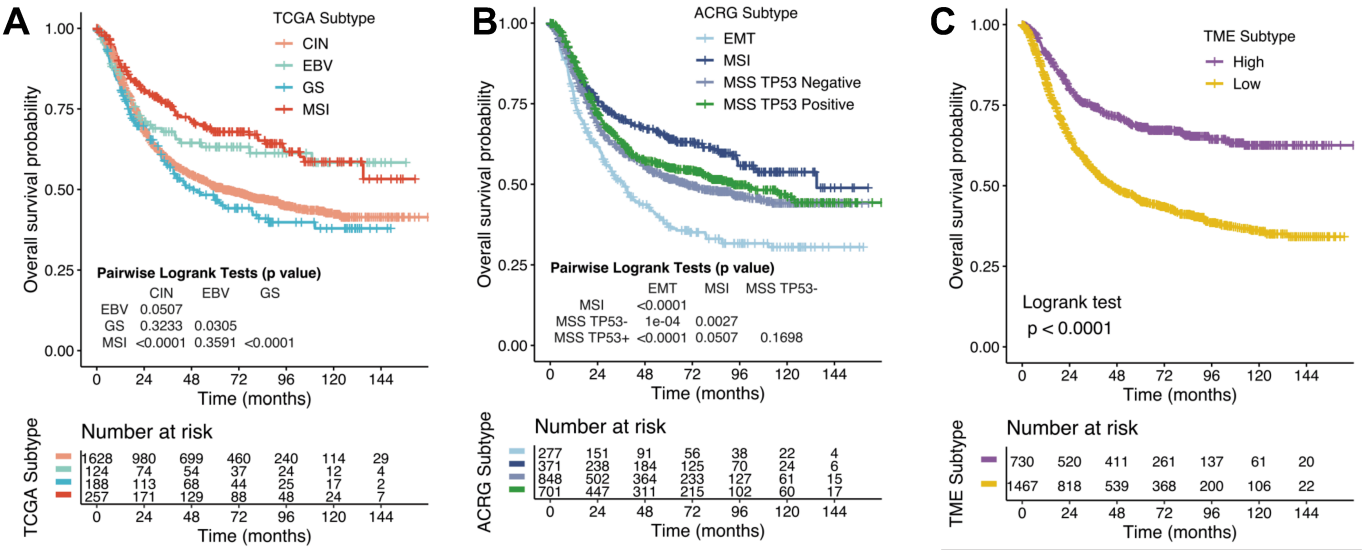
Variable	HR[1]	95% CI[1]	p-value
TME Subtype			
Low	—	—	
High	0.47	0.41, 0.55	<0.001
Global Logrank test = p <0.001, n=2197			
[1]HR = Hazard Ratio, CI = Confidence Interval			

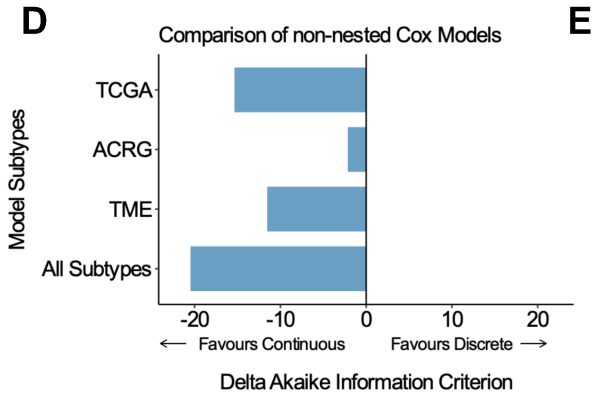
Table 4-4 Cox proportional hazards model for discrete TCGA subtypes

Variable	HR[1]	95% CI[1]	p-value
TCGA Subtype			
CIN	—	—	
EBV	0.71	0.52, 0.97	0.034
GS	1.13	0.91, 1.39	0.3
TCGA MSI	0.59	0.47, 0.74	<0.001
Global Logrank test = p <0.001, n=2197			
[1]HR = Hazard Ratio, CI = Confidence Interval			

Table 4-5 Cox proportional hazards model for discrete ACRG subtypes

Variable	HR[1]	95% CI[1]	p-value
ACRG_subtype			
EMT	—	—	
MSI	0.50	0.40, 0.63	<0.001
MSS TP53n-	0.69	0.58, 0.83	<0.001
MSS TP53+	0.62	0.52, 0.75	<0.001
Global Logrank test = p <0.001, n=2197			
[1]HR = Hazard Ratio, CI = Confidence Interval			

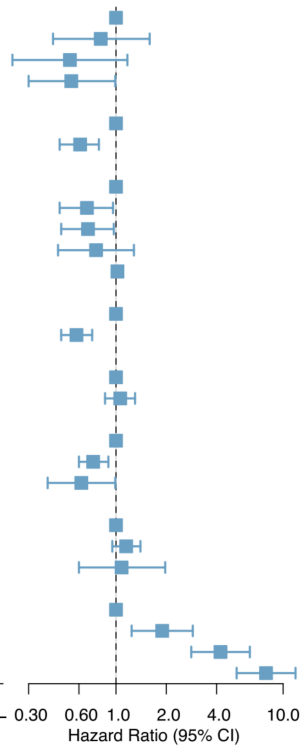


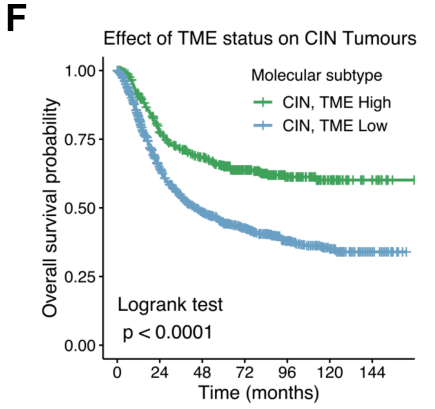


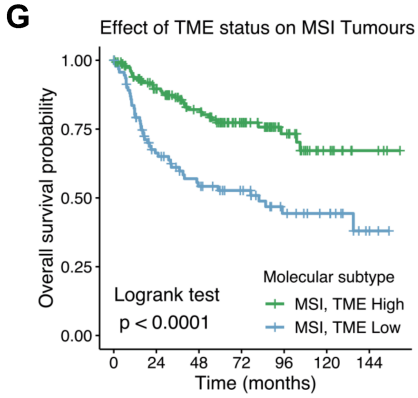
E

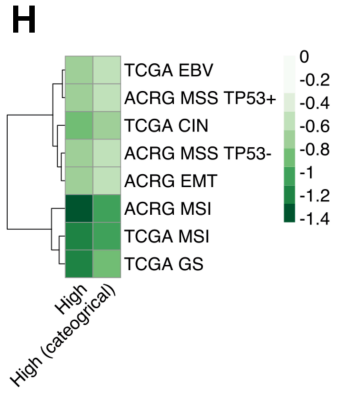
Variable	HR (95% CI)	Global Wald Test[1]
TCGA Subtype		**0.281**
CIN	reference	
GS	0.81 (0.42, 1.59)	
EBV	0.53 (0.24, 1.17)	
TCGA MSI	0.54 (0.30, 0.99)*	
TME Subtype		**<0.001*** **
Low	reference	
High	0.61 (0.46, 0.79)***	
ACRG Subtype		**0.146**
EMT	reference	
MSS TP53+	0.67 (0.46, 0.96)*	
MSS TP53-	0.68 (0.47, 0.97)*	
MSI	0.76 (0.45, 1.28)	
Age	1.02 (1.01, 1.03)***	**<0.001*** **
Chemotherapy		**<0.001*** **
No	reference	
Yes	0.58 (0.47, 0.72)***	
Sex		**0.566**
Female	reference	
Male	1.06 (0.86, 1.30)	
Lauren Class		**0.011***
Diffuse	reference	
Intestinal	0.73 (0.60, 0.90)**	
Mixed	0.62 (0.39,0.99)*	
Tumour Location		**0.391**
Distal	reference	
Proximal	1.15 (0.95, 1.40)	
Whole	1.08 (0.60, 1.97)	
Stage		**<0.001*** **
Stage I	reference	
Stage II	1.89 (1.24, 2.88)**	
Stage III	4.22 (2.82, 6.32)***	
Stage IV	7.90 (5.27, 11.86)***	

[1]Adjusted p value by Benjamini-Hochberg









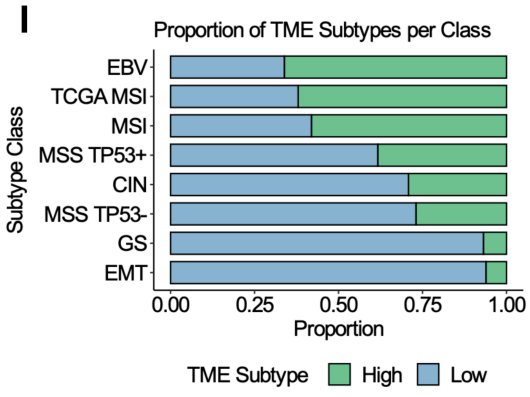


Figure 4-5 Cox proportional hazards model survival analysis of integrated molecular classification in gastric cancer. **A, B and C.** Kaplan Meier curves for univariate analysis of TCGA (A), ACRG (B) and TME (C) subtypes. Log Rank p value or pairwise comparison p values are shown for TME and TCGA and ACRG, respectively. **D.** Barplot for comparison of non-nested Cox models using corrected Akaike's information criterion. A negative value beyond -10 was deemed a significantly improved model. Negative values corresponded to continuous subtype scores. **E.** Forest plot for multivariable Cox model hazard ratios and 95% confidence intervals. **F and G.** Kaplan-Meier curves for TCGA CIN (F) and ACRG MSI (G) tumours stratified by TME High (green) and TME low (blue). **H.** Heatmap of Cox model coefficients for each molecular subtype stratified by TME subtype (High vs. Low) using both continuous and categorical subtype interpretations. A more negative coefficient (darker green) corresponds to improved survival. **I.** Proportional barplot of TME subtypes for each molecular subtype class. TME subtype is denoted by the plot legend.

Given that molecular subtypes are best interpreted as continuous variables, we constructed a comprehensive multivariable Cox model to assess the prognostic utility of molecular subtypes relative to prominent clinicopathologic factors (Figure 4-5AE). According to a Global Wald's Test, TME subtype was the only statistically significant independent prognostic molecular classification system (Hazard Ratio (HR) 0.61 [95% CI 0.46, 0.79], Benjamini-Hochberg adjusted $p<0.001$). The beneficial effect of receiving chemotherapy was nearly identical to TME High tumours (HR 0.58 [95% CI 0.47, 0.72], Benjamini-Hochberg adjusted $p<0.001$). Survival was also favoured by younger age, intestinal and mixed Lauren Class tumours, and lower stage tumours. Stage III and IV tumours carried the largest effect on overall survival.

After characterizing the survival advantage associated with greater TME High tumour scores we conducted a subgroup analysis of TME heterogeneity within each molecular subtype. Regardless of prognostic differences between CIN and TCGA MSI tumours, nearly identical separation of Kaplan-Meier curves was appreciated after stratification by TME tumour status (Figure 4-5F and G). In Figure 4-5H, we present the stratified Cox model coefficients for each molecular subtype in relation to TME High tumours interpreted as categorical or continuous variables. Regardless of the reference molecular subtype, negative Cox model coefficients identified improved overall survival in tumours with greater TME High scores/status. When we considered the proportion of TME High tumours within each molecular subtype relative to the univariate overall survival estimates we appreciated that tumours with a greater degree of TME High tumours generally possessed greater overall survival (Figure 4-5I).

4.3.7 Propensity score matched analysis

We applied traditional methods of causal inference using propensity score matched analysis to delineate potential treatment effects among molecular subtypes. The improvement in absolute standardized mean difference in unmatched and matched covariates and the distribution of matched propensity scores is presented in Appendix Figure D2-11 and Figure D2-12. The balanced cohort demographics accounting for matching weights is presented in Table 4-6.

The overall marginal effect in our cohort favoured chemotherapy for Stage II-IV gastric adenocarcinoma, which is consistent with previous studies (HR 0.66 [95% CI 0.50,0.87]).[314,315] Our model also statistically satisfied proportional hazard assumptions (Appendix Figure D2-13). We found significant effects favouring chemotherapy treatment in patients with high TME probability scores (HR 0.47 [95% CI 0.29, 0.74, p=0.04). In Figure 4-6A we illustrate the effects derived from our Cox model. We observed that patients with the lowest TME High scores have nearly similar effects if they receive chemotherapy or not. As TME High score increases some minimal survival benefit is identified for patients without chemotherapy but proportionally greater survival benefit is seen in those who receive chemotherapy. We also identified significant differences in the effect of treatment depending on the study cohort. Besides TME classification, all other molecular classification subtypes possessed nearly identical treatment effects to that of the population marginal effect (Appendix Figure D2-14).

We performed sensitivity analyses for a variety of alternative scenarios to assess the strength of our findings (Appendix Figure D2-15 and Figure D2-16). Matched cohorts were established for populations containing Stage I-IV and Stage II-III tumours. In all cases, we matched cohorts and fit models for categorical and continuous interpretations of TME High scores. In general, the direction of our subgroup effect remained similar but statistical significance was only achieved in Stage II-IV cancers for continuous and discrete TME variables. For our main analysis we calculated an E-value of HR 2.24. With reference to our multivariable model we concluded that unmeasured confounding roughly equivalent to the difference between Stage II and Stage III cancer would be required to nullify our effect.

Table 4-6 Propensity score matched cohort demographics

Variable	Overall, N = 395[1]	Chemotherapy, N = 237[1]	No Chemotherapy, N = 158[1]	p-value[2]
Age	64 (56, 70)	64 (56, 70)	63 (55, 70)	>0.9
Stage				>0.9
II	136 (34%)	81 (34%)	55 (35%)	
III	162 (41%)	96 (41%)	66 (42%)	
IV	97 (25%)	60 (25%)	37 (23%)	
Sex				0.7
Female	137 (35%)	84 (35%)	53 (34%)	
Male	258 (65%)	153 (65%)	105 (66%)	
TME Subtype				0.8
High	76 (19%)	45 (19%)	31 (20%)	
Low	320 (81%)	193 (81%)	127 (80%)	
TCGA Subtype				>0.9
CIN	282 (71%)	167 (70%)	115 (73%)	
EBV	22 (5.5%)	14 (5.9%)	8 (5.1%)	
GS	40 (10%)	26 (11%)	14 (8.9%)	
TCGA MSI	52 (13%)	31 (13%)	21 (13%)	
ACRG Subtype				>0.9
EMT	65 (16%)	39 (16%)	26 (16%)	
MSI	66 (17%)	39 (17%)	27 (17%)	
MSS_TP53neg	147 (37%)	90 (38%)	57 (36%)	
MSS_TP53pos	117 (30%)	69 (29%)	48 (30%)	
Study				0.9
ACRG	163 (41%)	96 (41%)	67 (42%)	
Kosin	55 (14%)	32 (14%)	23 (15%)	
KUGH	38 (9.7%)	25 (11%)	13 (8.2%)	
TCGA	126 (32%)	76 (32%)	50 (32%)	
Yonsei MDACC	12 (3.1%)	7 (3.1%)	5 (3.2%)	
Tumour Location				0.7
Distal	209 (53%)	127 (54%)	82 (52%)	
Proximal	186 (47%)	110 (46%)	76 (48%)	
Lauren Class				0.8
Diffuse	143 (36%)	89 (37%)	54 (34%)	
Intestinal	237 (60%)	139 (59%)	98 (62%)	
Mixed	15 (3.8%)	9 (3.8%)	6 (3.8%)	

[1]Median (IQR); n (%)
[2]Wilcoxon rank-sum test for complex survey samples; chi-squared test with Rao & Scott's second-order correction

4.3.8 Individual survival distribution

We constructed Individual Survival Distributions to expand on the utility of our integrated molecular classification models using continuous model probability scores. Using 1043 patients with available clinicopathologic characteristics, we evaluated the performance of several ISD models for our prediction task (Online Supplement). Multi-task logistic regression provided a superior calibrated model with nearly identical C-Index compared to CoxEN-KP (MTLR C-Index = 72.1% ± 3.3 versus CoxKP-EN = 72.2% ± 2.9). The MTLR

model was significantly D-calibrated and 1-calibrated for all bins except the 50% percentile (Appendix Figure D7-24).

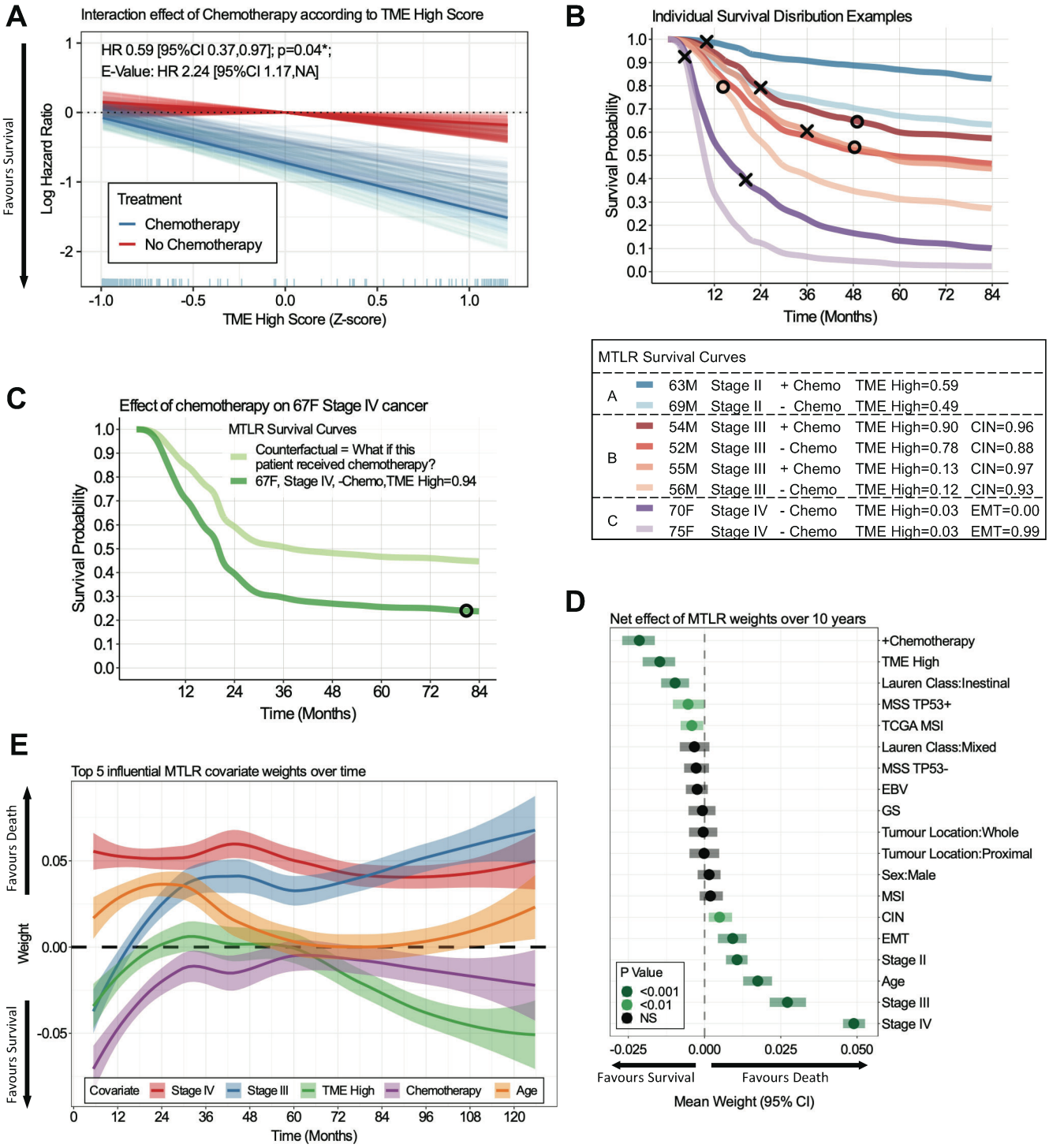


Figure 4-6 Propensity score matched survival analysis and individual survival distributions using multi-task logistic regression. **A.** Visualization of the subgroup analysis interaction effect between chemotherapy and TME High score from a Cox model using simPH in R. The z-score for TME High is on the x-axis and the log hazard ratio (Cox beta coefficient) is on the y-axis. A negative log hazard ratio favours survival. We performed 1000 simulations by drawing samples from a multivariate normal distribution based on parameters and covariances derived from the Cox model. The solid line represents the median simulation effect and each additional semi-transparent line is one model simulation. **B.** Individual survival curves for 8 patients using an MTLR model. The x-axis represents time in months and the y axis is survival probability. The patient characteristics for each colour is denoted in the plot legend below. **C.** Example of a counterfactual scenario illustrating the effect of chemotherapy for a given patient. For B and C, uncensored patients (ie. death event) and censored patients are denoted by a "X" and "O", respectively. **D.** Mean MTLR weights averaged over all time points for each covariate in our model. The point represents the mean and the semi-transparent box represents the 95% confidence interval estimate from 1000 bootstraps. Weights less than zero favour survival. A one-sample t-test was performed to assess if a covariate was significantly greater than zero and p values were corrected using Benjamini Hochberg. The p value significance is denoted in the plot legend. **E.** Loess smooth curves for the top 5 most influential covariate MTLR weights generated using 5-fold cross validation. Positive weights (y-axis) correspond to worse survival. The semi-transparent line signifies the 95% confidence interval.

We evaluated the capability of our MTLR model to provide insights into personalized medicine for gastric cancer. Three scenarios are presented in Figure 4-6B. Scenario A presents the effect of chemotherapy in two Stage II males in their sixties with similar TME scores. Here, the patient who receives chemotherapy has a 13.8% and 18.7% greater probability of survival at 24 and 48 months, respectively. Scenario B illustrates the relationship of TME score and chemotherapy in Stage III CIN tumours. We observed that survival curves for a patient who does not receive chemotherapy is nearly identical to one who receives chemotherapy with a TME low tumour. Consistent with our subgroup analysis in the previous section, MTLR also demonstrated additional survival benefit for TME High tumours in combination with chemotherapy. MTLR provided additional insight into individual survival effects of other molecular subtypes. In Scenario C we found that high EMT score profoundly affects overall survival in comparable Stage IV, chemo naïve and TME low tumours (Median survival High EMT = 9.4 months versus Low EMT = 14.5 months). We also investigated potential counterfactual applications of ISD models to facilitate communication of personalized medicine. In Figure 4-6C, we demonstrated the survival benefit, as interpreted by our MTLR model, of administering chemotherapy to a 67-year-old female with a Stage IV, high TME score gastric cancer. Although additional research is required,

counterfactual scenarios could be presented to patients as a visual interpretation of otherwise foreign and abstract statistical estimates of treatment benefits/harms.

Next, we assessed whether estimates from MTLR models can be translated to reference "gold-standard" Cox regression in an effort to enhance familiarity with ISD models. In Figure 4-6D, negative weights derived from the MTLR model correspond to improved survival. Here, we present the mean weights for all predicted time points with their respective 95% confidence intervals. Assuming a null-hypothesis weight of zero provides no survival effect, we assessed the significance of each covariate using a one-sample Wilcoxon test. In keeping with our Cox multivariable model, Stage III and IV provided the greatest effect on survival. Survival was also significantly decreased by age, stage II, and increasing EMT and CIN scores. The most beneficial survival effects were observed for chemotherapy and increasing TME high scores. Microsatellite instability as defined by our TCGA classifier significantly improved survival whereas ACRG MSI was not significant. Given this appealing presentation of survival effects we were satisfied that MTLR can provide familiar and concordant interpretations of survival relative to traditional Cox models.

In contrast to Cox, MTLR is not bound by the proportional hazard's assumption. We evaluated whether MTLR could provide unique insight into survival effects by modeling the weights derived from 5-fold CV. In Figure 4-6E, loess smooths and their 95% confidence intervals illustrate the relationship of the top 5 most influential covariates to survival over the entire time course of our model. As opposed to presenting a consistent risk over time, this approach suggested that certain covariates present more prominent effects at different times from disease presentation. For example, early beneficial effects are observed for chemotherapy which taper off after 24 months. Stage IV disease exerted a fairly constant negative survival effect over 10 years, but Stage III disease did not provide increased death until after one year.

4.4 Discussion

In this study we provide a comprehensive framework for developing, validating and applying multiple independent molecular classification systems to advance understanding of personalized medicine. We demonstrate that the integration of several molecular subtypes provides enhanced prognostic understanding and potential therapeutic implications. Our individual supervised machine learning classifiers predicted TCGA, ACRG and TME subtypes with accuracies ranging from 84.7% to 89.5% using only 39-57 genes per model. Using these models, we learned molecular subtypes in 2,202 gastric cancer patients. We

confirmed the superiority of interpreting molecular subtypes as continuous variables in Cox survival models. Using this perspective, we established that only TME classification provides significant prognostic utility in a multivariable Cox model. Furthermore, the prognosis of each molecular subtype was tightly related to the proportion of TME High tumours present. We characterized the beneficial relationship between greater TME High scores and chemotherapy using Cox regression with a propensity score matched cohort and individual survival distributions.

Previous research has demonstrated the benefit of a favourable tumour immune microenvironment for cancer prognosis and treatment. Indeed, the TME score provided predictive utility in identifying complete and partial responses to immunotherapy in melanoma and urothelial bladder cancer cohorts.[207] The abundance of TME High tumours in MSI and EBV tumours follows previous work which demonstrates the high degree of tumour immune infiltration and favourable response to immunotherapy in these tumours.[206] EBV tumours originate in the presence of viral infection, which may serve to provide increased interactions with the immune system.[316,317] Microsatellite instability also increases the propensity for a tumour to interact with the immune system as hyper-mutable cancers create an abundance of neo-antigens.[85,318] The problem is that chemotherapy and immunotherapy are simply improving survival outcomes for patients who already possess a superior prognosis. Thus, future investigation is required to find targeted therapies for other molecular subtypes or enhance the TME High signature in immunosuppressed tumours. Potential avenues to achieve this worthy goal includes the investigation of chemokine therapy or cancer vaccines.[125,243,319,320]

Intratumour heterogeneity is a significant barrier to personalized medicine in gastric cancer that obscures tumour classification and contributes to therapeutic resistance.[321–324] Probabilistic scores within our classifier models provides a measure of intratumor heterogeneity. For any supervised molecular classifier model, a score for each subtype will be produced for each patient. The sum of these scores equals one. Thus, for a given binary classifier, a patient will be allocated to a subtype if they possess a score of 0.51 or 0.99. To evaluate the effect of variability between model scores, prior research has proposed representing different combinations of scores as discrete heterogeneity categories.[312] Weighted *in-silico* pathology (WISP) is another method which interprets ITH using gene expression values as opposed to model scores.[325] The former strategy reduces biology to categorical data and has not been extended to the presence of heterogeneity across multiple independent classification systems. The latter approach provides a continuous

metric but is independent of the classifier model scores. Our novel approach called *Variance Heterogeneity* is derived directly from model scores and provides a continuous measure of ITH within and between molecular classifiers. Although additional work is required we demonstrated that varHet carries prognostic implications, is largely independent of sub clonal heterogeneity and can be used to quantitatively infer the biological concordance between subtypes.

Fundamental to our findings is the use of unbiased model construction techniques such as stratified folds and nested cross-validation for supervised machine learning.[293,294] We performed cross-platform normalization to account for confounding related to differences in gene expression technologies. Care must be taken in evaluating studies utilizing cross-platform data. As demonstrated by Franks et al., failure to account for cross-platform batch effects can result in poor model performance. For example, despite using identical patient samples in breast and colon cancer patients, the failure to normalize microarray and RNA-seq data resulted in subtype concordances of only 17-53% for unnormalized data versus 91-98% percent for Feature Specific Quantile Normalized data.[297] A variety of other methods have been proposed to perform cross-platform normalization.[326–328] Compared to non-paranormal transformation, quantile normalization and training distribution matching, FSQN performs well with real-world data and small sample sizes.

Attention must be delivered to model calibration to optimize the development of personalized medicine.[329,330] A calibrated model provides estimates that approximate real-world event rates. For example, in a calibrated model, a 30% probability of rain will translate to rain actually occurring 30% of the time. The calibration of binary classification models is well-defined and can be optimized with Platt scaling or isotonic regression.[300] However, multiclass calibration is fraught with a lack of standardized calibration metrics and methods.[302] Unfortunately, poor calibration accompanies the excellent accuracy of modern machine learning models such as random forest or neural networks.[299,301] In our study, none of our molecular classification models were statistically calibrated according to Asymptotic SKCE tests but our MTLR survival model passed D-calibration and 4 out of 5 1-calibration tests. We also found large heterogeneity in calibration metric values and the relative change in calibration error following the application of calibration methods. For example, Dirichlet calibration improved cw-ECE but worsened SKCE values. Heterogenous definitions of calibration further obscure the state of calibration in artificial intelligence, especially for multiclass models. Advancements in calibration methods for neural networks such as temperature scaling do not provide solutions for other multiclass models.[331] We

conclude that issues surrounding model calibration are a significant limitation to advancing personalized medicine.

Several limitations of our research exist. These findings are based on observational data and thus implications of TME High subtypes derived from propensity score matched analysis must be evaluated in prospective studies. The use of multiple public cohorts contributes significant heterogeneity to our patient population, especially in regard to regional differences in surgical management and chemotherapy regimens. Although public data drives additional research, there is a significant degree of missing data for important clinicopathologic variables. For example, in this study we were only able to use approximately ~50-60% of patients in multivariable analyses due to missing data. In many cases, imputation methods are not an option because missing data for a given variable is greater than 10-20%. Complete data for patient age, stage, pathology and treatment could further enhance the utility of public databases.

4.5 Conclusion

To date, this is the largest integrative molecular classification analysis for gastric cancer using TCGA, ACRG and TME molecular subtypes. We demonstrate that integration of independent molecular classification frameworks is feasible and provides novel biological understanding. To enhance understanding of personalized medicine we advocate for use of continuous interpretations of molecular subtypes and adoption of personalized survival models such as MTLR. Our novel metric called *Variance Heterogeneity* allows the assessment of intratumour molecular class heterogeneity at a single classifier or multi-classifier level. Principal among our findings is that prognosis and treatment effects using current chemo- or chemoradiotherapy regimens is largely driven by the tumour immune microenvironment, regardless of TCGA or ACRG subtype. Future applications of our approach to integrated molecular classification should be extended to external patient cohorts using a cost-effective translational test.

4.6 Data availability

Data used to perform all analyses and generate plots is provided in online supplement material at https://github.com/skubleny/Thesis-Supplemental-Files.

Chapter 5: A 107 gene Nanostring assay effectively characterizes complex multi-omics gastric cancer molecular classification in a translational patient-derived organoid model

5.1 Introduction

Patient-derived organoids and molecular classification are two promising methods to advance personalized medicine in gastric cancer. Three-dimensional cell culture to create "mini-organs" called organoids may be used to evaluate molecular cancer biology, assess pre-clinical therapeutic efficacy or potentially guide personalized treatment.[182,186,239] Given these findings some have suggested that 2-dimensional cell culture models could be displaced by organoids.[172] Molecular classification frameworks, such as those proposed by The Cancer Genome Atlas (TCGA) and Asian Cancer Research Group (ACRG) provide robust classification systems that define intertumoral heterogeneity.[26,82] Although understanding of tumour heterogeneity may help guide therapy it also represents a significant barrier to achieving personalized medicine.[332]

The utility of organoid models and molecular classification is yet to be fully realized by the patient. Logistics and cost inhibit widespread implementation of molecular classification, which is mainly predicated on whole-genome measurements consisting of tens of thousands of genes.[201,333] Furthermore, the first prospective trial evaluating whether patient-derived colon cancer organoids could provide relevant therapeutic information to the patient failed to show feasibility.[174] How should we approach the future of personalized medicine in gastric cancer? To maximize the advancement of personalized medicine in gastric cancer we argue that there is a need to better understand the advantages and disadvantages of each pre-clinical model and improve accessibility to molecular classification.

In previous work (Chapter 4), we developed supervised machine learning classifiers for TCGA, ACRG and TME subtypes. These models were shown to accurately reproduce subtype assignment through cross-validated estimates and in reference to external gold-standard features. In this prospective study, we evaluate the feasibility of using a 107 gene Nanostring assay in assigning TCGA and TME molecular subtypes derived from our classifier models to clinical formalin-fixed paraffin-embedded (FFPE) specimens and patient-derived organoid cultures. We evaluate the validity of subtype classification and scores to gold-

standard reference features and assess the landscape of molecular subtypes in 2D and 3D cell culture models.

5.2 Methods

5.2.1 Study Design

We performed this single-center, prospective study at the University of Alberta in Edmonton, Alberta, Canada from June 2019 to January 2022. All human clinical participants consented according to the approved ethics protocol granted by the Health Research Ethics Board of Alberta (Study ID: HREBA.CC-17-0228). Treatment naïve Stage I-IV sporadic gastric adenocarcinoma patients aged greater than 18 years were included. Patients with a known inherited oncogenic germline mutation or hereditary syndrome (ie. Familial Adenomatous Polyposis) were excluded. The study overview is illustrated in Figure 4-1.

Specimens were retrieved via endoscopic biopsy at the time of diagnosis, staging laparoscopy or at the time of surgical resection at the Walter C Mackenzie Health Sciences Centre or Royal Alexandra Hospital. Four biopsies were collected for permanent pathology in Z-fixx (Sigma) for cancer and adjacent normal tissue, respectively. For the generation of patient-derived organoids four additional fresh cancer and adjacent normal biopsies were collected.

5.2.2 Organoid culture

Our previously described human organoid protocol resembles methods from Bartfeld et al. and Mowat et al.[179,243] Specific reagent information and catalogue numbers can be found in Appendix .E.2. Briefly, fresh gastric tissue from tumour and normal adjacent stomach were collected in Hank's Balanced Salt solution supplemented with 2X penicillin and streptomycin on ice. Tissues were minced with sterile scissors and washed three times in sterile PBS. Minced tissues were enzymatically and mechanically digested in 20 mL digestion buffer (Advanced DMEM/F12, 1X penicillin/streptomycin, 100 µg/mL Primocin, 2.5% FBS, 75 units/ml Collagenase XI and 125 µg/ml Dispase II) and placed a mechanical water bath at 37 degrees Celsius for 1 hour. Dissociated cells were filtered, washed and counted with a hemocytometer using 1:1 DMEM and Trypan Blue (Gibco, 15250061). Cells were resuspended in ice-cold 70% Matrigel (Corning, 356253) in Advanced DMEM/F12 at a concentration of 1000 cells/µL and 35 µL Matrigel domes were placed in a prewarmed 24 well tissue culture treated plate. Organoids were cultured in 500 µL organoid culture

medium at 37 degrees Celsius and 5% CO_2 until mature for passaging or downstream analysis.

Organoid culture media contained 1:1 basal culture medium and conditioned L-WRN cell supernatant enriched in R-spondin, noggin and Wnt (ATCC, CRL-3276). Conditioned L-cell supernatant was prepared according to Miyoshi and Stappenbeck.[244] Organoid culture media contained Advanced DMEM/F12, 2 mM L-Glut, 10 mM HEPES, 1X penicillin/streptomycin, 100 µg/mL Primocin, 1X N2, 1X B27, 1 mM N-acetylcysteine, 1 nM gastrin, 10 mM nicotinamide, 500 nM A83-01, 10 µM SB202190 and 50 ng/mL human EGF. Rho-associated kinase inhibitor (10 µM Y-27632) was added for the first plating and first media change after splitting.

Normal gastric organoids were passage approximately every 7 to 14 days. Cancer organoids were passaged every 7 to 21 days due to cancer specific variable growth rates. Briefly, ice-cold Advanced DMEM/F12 was added to Matrigel domes and organoids were mechanically lifted using a pipette tip. Organoids were dissociated using TrypLE Express (Gibco, 12604013), washed three times and counted with a hemocytometer as above. Dissociated organoid cells were either cultured or allocated for histology, molecular analysis or dose-response assays.

5.2.3 Nanostring assay

Organoid cells (500,000 to 1,000,000 cells) were dissociated using TrypLE express and washed in PBS three times. Nucleic acids were stored at -80 degrees Celsius after cells were lysed in 350 µL Buffer RLT Plus (Qiagen, 1053393) with 10 µL/ml 2-mercaptoethanol and homogenized using Qiashredder tubes (Qiagen, 79654). DNA and RNA were isolated from four 10 µm FFPE curls using an AllPrep FFPE DNA/RNA kit (Qiagen, 80234). We followed the manufacturer's protocol except we performed proteinase-k digestion for 2 hours. Organoid DNA and RNA was separated using an AllPrep DNA/RNA kit (Qiagen, 80204). In both FFPE and organoid derived samples, mRNA cleanup was performed using a RNA Clean and Concentrator kit. (Zymo, R1015). Quality and concentration of RNA was assessed using NanoDrop 1000 (Thermo Scientific). DNA and RNA were stored at -80 degrees Celsius prior to downstream analysis.

A custom Nanostring codeset was designed to allocate TCGA (57 genes) and TME subtypes (50 genes) using 107 genes chosen by previously developed supervised machine learning classifiers (Online Supplement). Measurement of gene expression using our custom codeset

was performed by the Lab Medical Pathology core service at the University of Alberta. Overnight hybridization of 100ng of total RNA was performed followed by measurement of gene counts using the Nanostring nCounter machine.

Normalization of Nanostring data was performed according to Molania et al.[334] Raw Nanostring counts were imported to nSolver software 4.0 and samples with quality control flags were excluded. All included samples were subsequently normalized using Removal of unwanted variation-III algorithm (RUV-III) with technical replicates spanning different Nanostring cartridges.[335,336] Following RUV-III normalization, Feature Specific Quantile Normalization was used to transform gene expression values to the target distribution of TCGA and TME classifiers.[297] Molecular subtypes were learned using previously characterized models. For technical replicates we used the mean value of subtype scores.

5.2.4 Immunofluorescence and Immunohistochemistry

Tissue biopsies and organoids were fixed in zinc-formalin (Z-Fixx, Sigma) for 24 hours, washed three times and stored in 70% ethanol prior to preservation in paraffin. Fixed organoids were suspended in agar prior to paraffin embedding. All tissue sections for microscopy were cut at 5 µm. Specific antibody combinations, dilutions, incubation times and antigen retrieval buffers are listed in Appendix .E.1.

Briefly, tissue sections were deparaffinized in Histoclear (National Diagnostics) and rehydrated. Microwave heat-induced epitope retrieval was performed using Sodium Citrate (pH 6, heated to 94 degrees Celsius in 1-minute intervals followed by 9 minutes continuous heat) or Tris-EDTA (pH 9, heated to 94 degrees Celsius in 1-minute intervals followed by 8 minutes 30 seconds continuous heat).

For immunofluorescence, permeabilization was performed with 0.5% Triton X-100. Non-specific epitopes were blocked using 10% normal goat serum. Tissue sections were stained with primary antibodies (mouse anti-pan cytokeratin, rabbit anti-vimentin) and secondary antibodies (Goat anti-mouse IgG Alexa Fluor 568, 1:50, Invitrogen, A-11004 and Goat anti-rabbit IgG Alexa Fluor 488, 1:200, Abcam, ab150077). Autofluorescence was diminished using a TrueView Quenching kit per manufacturers protocol (Vector, SP-8400-15). Nuclear counterstaining was performed with DAPI followed by cover slipping with Vectashield Vibrance Antifade mounting media (Vector, H-1700). Images were captured on an AxioCam HRc camera and processed using ImageJ.[245]

For immunohistochemistry, endogenous peroxidases were blocked with 3% hydrogen peroxide in methanol. Non-specific binding was mediated by blocking with 20% normal goat serum and avidin/biotin blocker per manufacturers protocol (Vector Laboratories-2001). Mouse anti-MUC5AC primary antibody or rabbit anti-pepsinogen II/PGC was diluted 1:250 and 1:100, and incubated at 4 degrees Celsius overnight, respectively. Goat anti-rabbit or goat anti-mouse biotinylated IgG secondary antibodies were incubated at 1:200 dilution for 30 minutes for anti-PGC and anti-MUC5AC, respectively. Antibody detection was performed using avidin-biotin complex/horseradish peroxidase (Vector Laboratories) and 3,3-diaminobenzidine tetrahydrochloride (DAB, Abcam, ab64238) per manufacturer's protocol. Tissues were counterstained with hematoxylin. Brightfield microscopy images were captured using a Leica Aperio CS2 digital slide scanner.

5.2.5 Epstein-Barr encoded early RNAs *in-situ* hybridization

To confirm that our Nanostring derived TCGA EBV score captured true EBV positive tumours we used gold-standard Epstein-Barr encoded early RNAs in-situ hybridization (EBER ISH) in parent tumour and patient-derived organoids.[309] Tissue biopsies and organoids were preserved in FFPE and prepared for microscopy as above. Detection of EBER ISH was performed using biotinylated alkaline phosphatase mediated *in-situ* hybridization with the Rembrandt detection kit according to manufacturer's protocol (ThermoFisher, A500K.0105). Tissue sections were counterstained using nuclear fast red.

5.2.6 Microsatellite instability

Pentaplex PCR is the gold-standard method for assessing tumour microsatellite instability (MSI).[310,311] Briefly, DNA was isolated from FFPE sections and PDOs in paired normal and cancer tissue as above. The concentration of mRNA free genomic DNA was quantitatively measured using QuBit (ThermoFisher). We tested for MSI status using the MSI Analysis System, Version 1.2 (Promega, MD1641), which assesses MSI using five fluorescent labeled mononucleotide markers (BAT-25, BAT-26, NR-21, NR-24 and MONO-27). Briefly, 2 ng of genomic DNA was amplified using GoTaq MDx Hot Start polymerase (Promega, D6001) on a BioRad T100 thermal cycler per manufacturer recommended settings. Capillary electrophoresis was performed by The Applied Genomics Core at the University of Alberta on an Applied Biosystems 3130XL Genetic Analyzer. Data was processed using Geneious Prime version 2021.2.2. Alleles were called using 2nd order least squares. Samples were deemed MSI-High, MSI-Low or microsatellite stable if allelic variations were found in $\geq$ 2, 1 or 0 microsatellite markers, respectfully.

5.2.7 *In-vitro* dose response assay

Our combination FLOT (5-fluorouracil, leucovorin, oxaliplatin and docetaxel) dose-response assay was validated using the AGS human gastric cancer cell line (ATCC CRL-1739). AGS cells were cultured in Ham's F-12K (Gibco), 10% FBS (Gibco) and 1X penicillin and streptomycin (Gibco). Briefly, 5000 cells were plated in 96 well plates and grown for 24 hours. Anti-cancer drugs 5-fluorouracil (Tocris, 3257), oxaliplatin (Tocris, 2623) and docetaxel (Tocris, 4056) were added in triplicate over 8 half-log dilutions (5-fluororacil and oxaliplatin) or 10-fold dilutions (docetaxel). Initial concentrations for 5-fluorouracil, oxaliplatin and docetaxel were 800 µM, 2400 µM and 2400 nM, respectively. A single 500 µM dose of leucovorin (Toronto Research, L330400) was added to each treatment well. Cells were treated for 48 hours followed by a CCK-8 viability assay (Abcam, ab228554), which was performed according to the manufacturer's protocol. We replicated our assay in three independent trials to assess reproducibility.

For organoid dose-response assays, organoids were passaged and dissociated according to our protocol. As above, 5000 cells were plated in 96 well plates and grown in organoid media for 24 hours followed by 48 hours of FLOT treatment and viability assessment using a CCK-8 assay.

Nonlinear regression of dose response data was performed using GraphPad Prism version 9. First, baseline media control absorbance was subtracted from all experimental wells. Absorbance values were then normalized between 0 and 100% given that adequate minimum and maximum treatment effect controls were present. Mean drug concentrations from 5-fluorouracil, oxaliplatin, docetaxel and leucovorin were log10 transformed. Next, outliers were removed using Q = 1% and least squares variable slope non-linear regression estimated dose-response curves and half-maximal inhibitory concentration (IC50). Drug Sensitivity Score (DSS) was calculated using estimated parameters from nonlinear regression using the DSS package.[337]

5.2.8 Cancer Cell Line Encyclopedia Analysis

We retrieved multi-omics data and dose-response data for gastric cancer cell lines present in the Cancer Cell Line Encyclopedia.[338] Data were downloaded from the Broad DepMap Portal (https://depmap.org/portal/download/). A complete list of datasets used for this analysis is available in the online supplement (https://github.com/skubleny/Thesis-Supplemental-Files).

RNA-seq data were retrieved as log2 transformed RNA-Seq by Expectation-Maximization (RSEM) counts.[283] We filtered out gene counts < 1. Data were assessed using qqplots, boxplots and probability density functions to ensure adequate normalization was achieved. Molecular subtypes for TCGA, ACRG and TME classification were assigned using our previously characterized supervised classifiers. Prior to assigning subtypes, gene expression values were transformed to target distributions of each classifier using FSQN.

We compared the association of molecular subtype scores from each classification model to dose-response data and validated the intended molecular phenotypes of each scores using aneuploidy scores, p53 mutation status and microsatellite instability status .[338–340]

Area under the curve drug sensitivity values were retrieved from the Sanger Genomics of Drug Sensitivity in Cancer (Sanger GDSC2).[341] We filtered out drugs with greater than 25% missing data. Pearson's correlation of subtype score and drug sensitivity was performed. Pearson's correlation values were then pooled according to drug targets as specified by GDSC. A Kruskal-Wallis test evaluated drug target categories with respect to the Pearson's correlation of drug efficacy and molecular subtype score.

5.2.9 Statistical Analysis

Statistical analyses were completed using R version 4.1.2.[226] Differences between groups were assessed with a Wilcoxon two-sample test or Kruskal-Wallis test. Differences between paired samples was assessed using two-tailed paired Wilcoxon test. When applicable post-hoc tests were completed using Dunn's test. Pearson's correlation or Spearman's Rho was used to assess associations between continuous variables. Multiple comparisons corrections were made using Benjamini-Hochberg or Bonferroni method, as specified. Statistical significance was defined at alpha = 0.05.

5.3 Results

5.3.1 Demographics of prospective cohort, molecular subtypes

Demographics data for 39 patients in our prospective cohort are found in Table 5-1. The median age was 65 (60, 74) with a predominance of male patients (67%). All stages were equally present but the majority of tumours were poorly differentiated (G3 = 69%). The majority of tumours were also proximal, diffuse and exhibited signet ring cell features.

The assigned molecular subtypes of our patient population using Nanostring gene counts measured from FFPE specimens are provided in Table 5-2. Following appropriate normalization procedures, we used our previously characterized supervised classifiers to assign TCGA and TME classes and scores. The proportion of TCGA classes was largely concordant to subtypes assigned in a public cohort of 2,202 patients. Chromosomal instability was observed in 74.4% of patients (Public cohort = 74.1%), EBV in 5.1% (Public cohort = 5.6%), GS in 7.7% (Public cohort = 8.6%) and MSI in 12.8% (Public cohort = 11.7%). However, the prevalence of TME High tumours was less in our population (12.8% vs. 33.2%). Among TME High tumours we observed the expected presence of EBV and MSI classes but the prevalence of CIN tumours was lower than the reference population (2.6% vs 21.6%).

Table 5-1 Patient demographics

Variable	n/N (Missing %)	N = 39[1]
Age	39 / 39 (0%)	65 (60, 74)
Sex	39 / 39 (0%)	
F		13 (33%)
M		26 (67%)
Stage	39 / 39 (0%)	
I		11 (28%)
II		10 (26%)
III		8 (21%)
IV		10 (26%)
Grade	39 / 39 (0%)	
G1		1 (2.6%)
G2		10 (26%)
G3		27 (69%)
Gx		1 (2.6%)
Tumour Location	39 / 39 (0%)	
Distal		13 (33%)
Proximal		23 (59%)
Whole stomach		3 (7.7%)
Lauren Classification	38 / 39 (2.6%)	
Diffuse		23 (61%)
Intestinal		13 (34%)
Mixed		2 (5.3%)
Signet Ring Cell (Positive)	38 / 39 (2.6%)	23 (61%)

[1]Median (IQR); n (%)

5.3.2 Organoid model

We established twelve patient derived gastric cancer organoids derived from endoscopic biopsy or from resected surgical specimens. Our success rate of establishing organoid culture was 75%, with two samples failing to grow and two others experiencing fungal contamination. Histology of parent cancer and tumour-organoid pairs largely revealed consistent tissue morphology. In Figure 5-1B representative images demonstrate morphological recapitulation of the parent tumour in patient-derived organoids.

Table 5-2 Categorical assignment of molecular subtypes in bulk-tumour samples

	CIN	EBV	GS	MSI	TME Sum
TME High	1	2	0	2	5
TME Low	28	0	3	3	34
TCGA Sum	29	2	3	5	

To ensure recapitulation of gastric epithelium we assessed the expression of PGC and MUC5AC in parent tissue and patient derived organoids (Figure 5-1A). Concordance was observed between all organoid pairs. The presence of epithelial and stromal tissue was confirmed using immunofluorescent staining of pan cytokeratin and vimentin. Minimal vimentin expression was observed in diffuse type organoids relative to differentiated intestinal-like organoids.

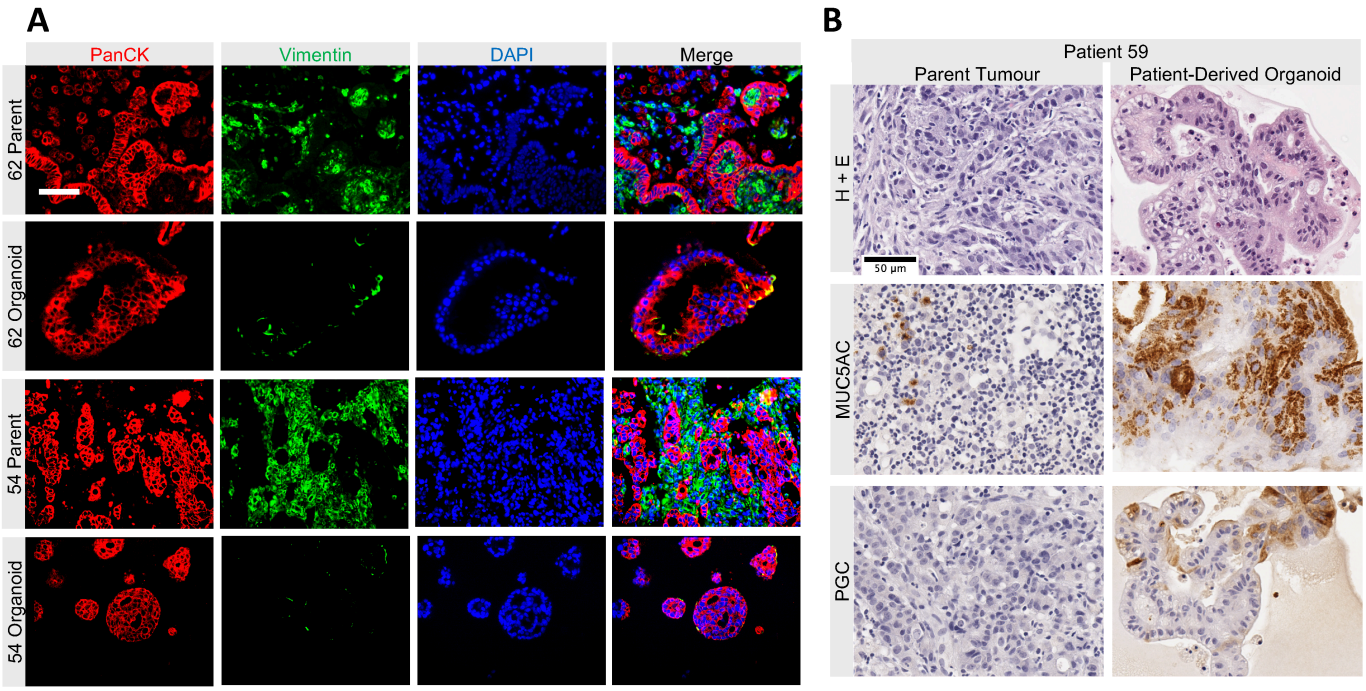


Figure 5-1 Representative images of human gastric cancers and patient-derived organoids demonstrating recapitulation of stromal microenvironment and gastric tissue. **A.** Immunofluorescence demonstrating recapitulation of stromal cells within organoids. Pan cytokeratin (red) is present on epithelial cell membranes, Vimentin (green) is a stromal cell specific marker and DAPI (blue) was used to stain nuclei. **B.** Tumours and patient-derived organoids demonstrated similar morphology on hematoxylin and eosin stains. The gastric specific markers MUC5AC and PGC were present in both parent and organoid tissues. Scale bar represents 50 µm.

5.3.3 Normalization technique

Normalization according to methods proposed by Molania et al. provided reliable elimination of sample and batch differences (Figure 5-2A and B). We observed close approximation of technical replicates following normalization with the median log differences values approaching zero (Figure 5-2C). Feature Specific Quantile Normalization provided excellent approximation of normalized Nanostring gene expression data to target distributions to learn molecular subtypes from our previously established supervised machine learning model (Figure 5-2D).

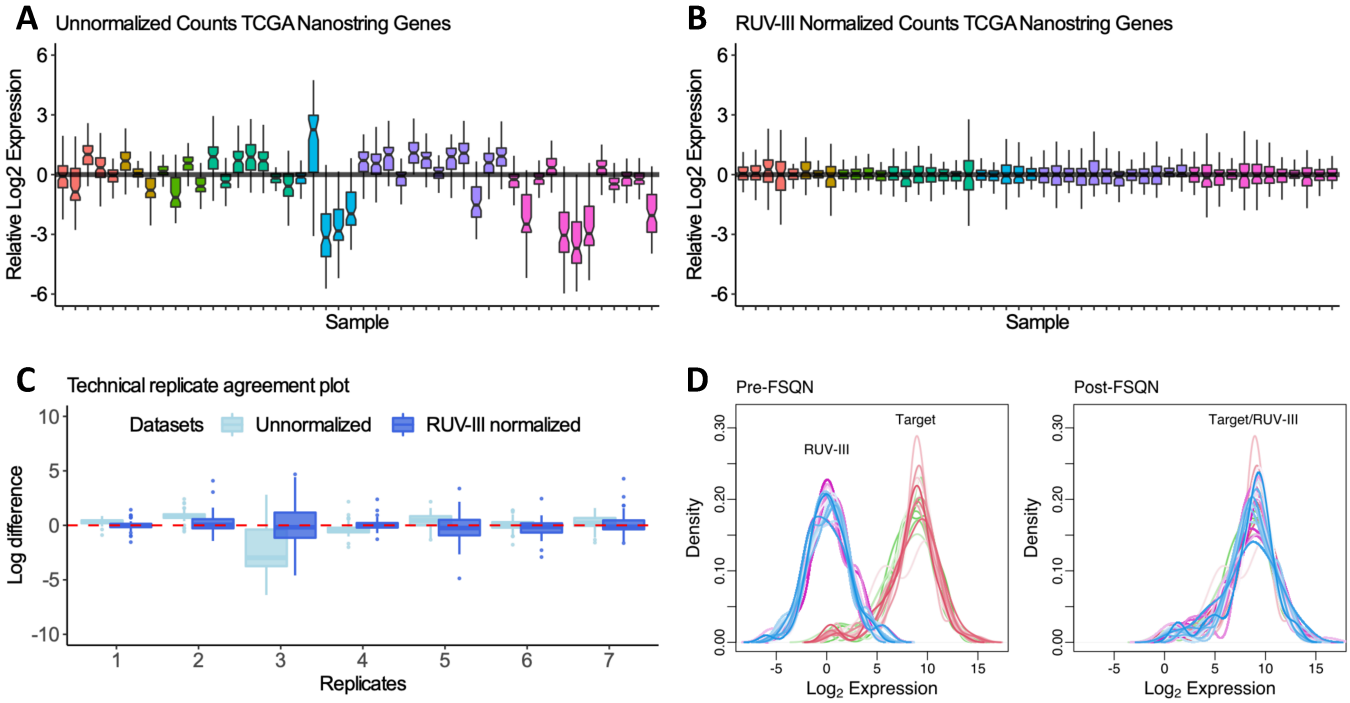


Figure 5-2 Nanostring normalization and Feature Specific Quantile Normalization example. **A.** Boxplots for 39 formalin-fixed paraffin-embedded specimens demonstrating the distribution of unnormalized Log2 expression for 57 TCGA genes. Boxplot notches approximate the 95% confidence interval of the median. **B.** Boxplots showing the same samples from (A) following RUV-III normalization. **C.** Boxplot showing comparison of unnormalized and normalized log difference between 7 samples that we used as technical replicates for RUV-III normalization of patients in (A) and (B). Optimal technical agreement should approach zero and improvement relative to unnormalized samples should be observed. **D.** Example of Feature Specific Quantile Normalization (FSQN) for same patients from (A), (B) and (C). FSQN normalizes the RUV-III distribution developed using our Nanostring assay to the target distribution of the TCGA RNA-seq data used to train our machine learning classifier.

5.3.4 Our Nanostring assay is externally valid in reference to gold standard features

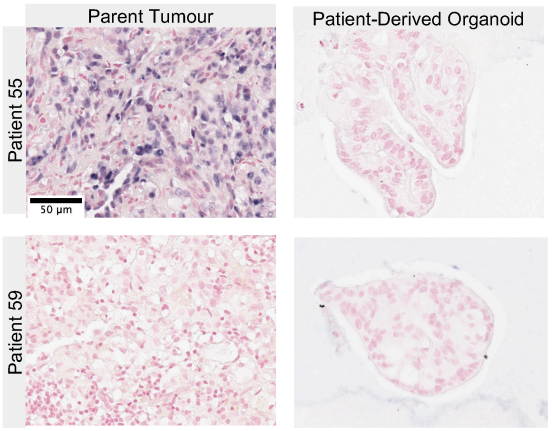
We previously established that our supervised molecular classifier produces subtype classes and scores that significantly reapproximates the intended phenotype. For example, EBV scores derived from our model provides an area under the curve value of 99.7% for predicting gold-standard EBV ISH status. In our own samples, we assessed the accuracy of

our Nanostring codeset in reproducing the accuracy of our model with reference to gold-standard tests for EBV and MSI tumours.

A

Epstein-Barr Virus Gastric Cancer		TRUE CLASS	
		Positive	Negative
PREDICTED CLASS	Positive	1	0
	Negative	0	19

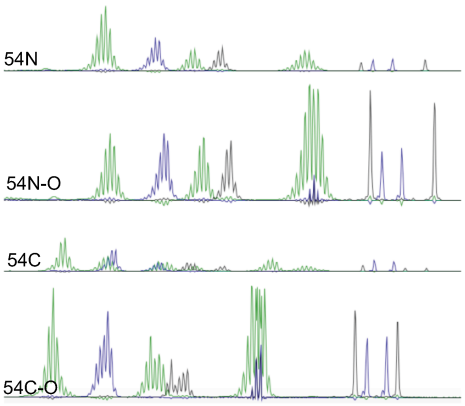
B



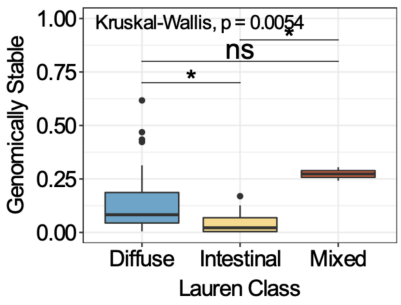
C

Microsatellite Instability Gastric Cancer		TRUE CLASS	
		Positive	Negative
PREDICTED CLASS	Positive	2	0
	Negative	0	18

D



E



F

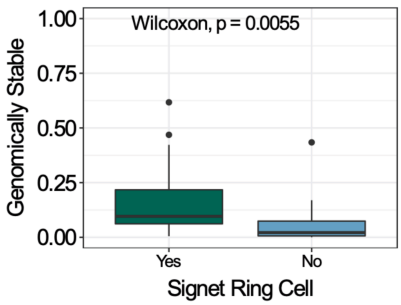


Figure 5-3 External validation of our Nanostring test in 12 parent and patient-derived organoid pairs. **A.** Contingency table comparing allocation of Epstein-Barr Virus (EBV) subtypes using our test versus the gold-standard EBER ISH. The tests were 100% concordant. **B.** Representative images of EBER ISH in two tumour organoid pairs. Parent tumours are shown in the left column and the patient-derived organoid is shown in the right. The parent tumour for patient 55 was the only EBV positive sample. **C.** Contingency table comparing allocation of Microsatellite Instability (MSI) subtypes using our test versus the gold-standard pentaplex pcr. The tests were 100% concordant. **D.** Electropherogram of pentaplex pcr for one patient's set of normal parent (54N) and normal organoid (54N-O) versus cancer parent (54C) and cancer organoid (54C-O). Allelic shift, which signifies MSI, is evident in both cancer tissue and cancer organoids relative to normal. **E.** Boxplot of genomically stable scores in 39 formalin-fixed paraffin-embedded samples versus Lauren classification. **F.** Boxplot of genomically stable scores in 39 formalin-fixed paraffin-embedded samples versus signet ring cell status. Statistical tests and significance is presented in the plots. For (E), a post-hoc Dunn's test with Bonferroni adjusted p values was performed (*P<0.05).

We performed EBV ISH and pentaplex PCR on 10 paired parent and tumour organoid samples. In these samples, our Nanostring test was 100% accurate in assigning EBV and MSI classes with reference to these tests (Figure 5-3A and C). In Figure 5-3B, we illustrate a positive and negative EBV ISH test in parent tumour samples (top left and bottom left, respectively) and their corresponding organoid (top and bottom right). We found that our test was able to discern the presence of EBV in the parent tumour and absence of EBV within the derived organoid. We identified MSI-High status in one parent tumour and organoid pair. In Figure 5-3D, we present the capillary electrophoresis measurements for patient 54 in cancer (54C) and cancer organoid (54C-O) tissue with reference to normal stomach controls. Three allelic shifts were identified in 54C and 54C-O. Once again, our test correctly identified the presence of MSI status in FFPE and fresh tissue.

Given that we did not perform whole genome sequencing in this preliminary study, we assessed the validity of CIN and GS assignments using pathology characteristics. Genomically stable tumours have previously been demonstrated to be enriched in diffuse type gastric cancer and signet ring cells.[26,82] Our test recapitulated this relationship with significantly greater GS score in diffuse and mixed-type cancers relative to intestinal-type (Figure 5-3E) (Dunn's Test, Bonferroni adjusted $p < 0.05$). In Figure 5-3F we also confirmed greater GS scores in signet ring cell tumours (Wilcoxon, $p<0.01$).

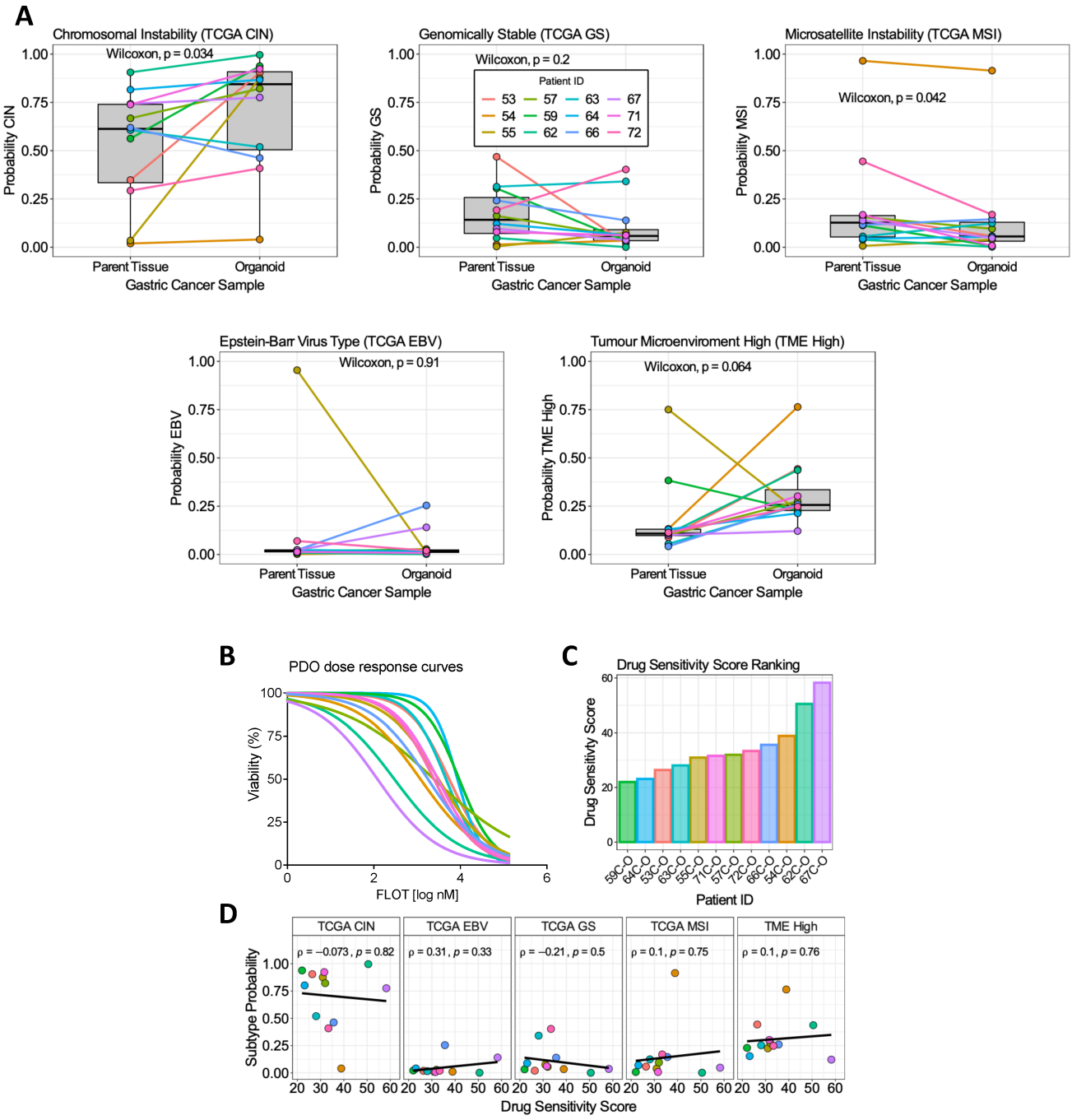


Figure 5-4 Assessment of concordance of molecular subtype scores between tumour organoid pairs and association of subtype scores with FLOT drug sensitivity. **A.** Paired boxplots comparing 12 parent tumours and patient-derived organoid pairs (x-axis). The subtype is designated by the plot title. The y-axis represents the corresponding subtype score. P values for a paired Wilcoxon test are presented in each plot. Each colour represents a tumour organoid pair with the colour corresponding to those identified in (C). **B.** Dose response curves for 10 patient derived organoids treated with FLOT chemotherapy *in-vitro*. The x-axis represents the log of FLOT concentration in nanomoles. Cell viability was measured using a

CCK-8 assay. **C.** Bar plot illustrating drug sensitivity scores (DSS) in ascending order (least favourable to most favourable response) for corresponding tumour organoids identified on the x-axis. **D.** Scatterplots illustrating the association of molecular subtype probability versus DSS for each molecular subtype as designated in the plot title. Pearson's correlation coefficient and p value are represented in each plot. The black line represents the line of best fit using simple linear regression.

5.3.5 Molecular subtype discordance exists between tumour and tumour organoid pairs

After establishing the validity of our test, we assessed the propensity of patient-derived organoids to recapitulate the molecular subtype of their parent tumour sample. Using categorical assignment of subtypes, we found 83.3% and 75.0% concordance between tumours and tumour organoids for TME and TCGA classes, respectively (Appendix .E.3). Paired analysis of continuous molecular subtype scores found that CIN and MSI scores were significantly different between tumour and tumour organoids (Paired Wilcoxon, unadjusted p=0.03 and p =0.04, respectively) (Figure 5-4A). The effect of organoid culture was observed to increase CIN and TME High scores, yet decrease GS scores.

5.3.6 Organoid drug assay

Dose-response assays using FLOT chemotherapy were completed for all 12 patient-derived organoids. First, we successfully validated the reproducibility of our method using the AGS cancer cell line across three independent dose-response assays (Appendix Figure E4-25). Least squares nonlinear regression were well fit to our dose-response data with a median adjusted goodness of fit value of 0.92 (IQR 0.84, 0.94). The dose response curves and corresponding drug sensitivity scores (DSS) are presented in Figure 5-4B and C. There were no significant associations found between DSS and molecular subtype scores (Figure 5-4D). We observed increasing efficacy of FLOT with increasing EBV score and decreasing efficacy with increasing GS scores, although all Spearman Rho effects were less than 0.31.

5.3.7 2-dimensional cell line remain a valuable tool to infer drug effects on certain molecular subtypes

Given the indiscriminate effects of cytotoxic FLOT on patient derived cancer organoids we assessed whether molecular subtype scores provide insight into targeted therapy effects. Furthermore, we assessed the potential utility of existing gastric cancer cell lines to approximate molecular classification heterogeneity. Using the CCLE we accessed multi-omics and drug sensitivity data for 37 gastric cancer cell lines.

The distribution of subtype probability scores in a cohort of 2,202 gastric cancer patients was generally recapitulated among gastric cancer cell lines (Figure 5-5A). The only significant difference between bulk tumour sequenced molecular subtype scores and those of cancer cell lines were found between TME High and TME Low scores (Wilcoxon, adjusted $p = 0.01$). Next, we assessed the distribution of subtype scores for the 20 cell lines with accompanying dose response assay data from the Sanger GDSC cohort. There was no significant difference identified in subtype scores between excluded and included cell lines and thus the molecular subtype scores for drug assay analysis approximated TCGA and ACRG subtypes.

We investigated the *in-vitro* effect of various anti-cancer drugs according to each molecular subtype score at a univariate level. Using Pearson correlation, 48 drugs were found to be statistically associated with improved or decreased efficacy according to subtype score. However, after multiple comparisons correction there were no statistically significant relationships. Amongst potential false positive comparisons, a variety of therapeutic targets were identified including Temozolomide, the survinin inhibitor YM-155 and Dactinomycin (Figure 5-5C). Amongst the largest effect sizes, TCGA Genomically Stable subtype was most commonly associated with decreased efficacy, whereas TCGA EBV was associated with improved efficacy. The results for the complete analysis are available in the online supplement.

Next, the efficacy of drugs against specific targeted pathways was evaluated according to molecular subtype scores. We utilized the classification system of targeted pathways proposed by GDSC. For each molecular subtype a Kruskal-Wallis test followed by Dunn's post-hoc analysis with multiple comparisons correction was conducted. In Figure 5-5D we illustrate the relationship of CIN scores with targeted drugs for various pathways. We identified statistically significant efficacy of EGFR signaling targets versus targets against DNA replication, Genome integrity, PI3K/MTOR signaling, Mitosis and Other targets including RNA helicase A, NAE, IKK-1 and IKK-2. This is consistent with hypotheses generated by the TCGA STAD paper that highlighted the presence of receptor tyrosine kinase mutations in CIN tumours.[57] Importantly, DNA replication targets represent conventional chemotherapy agents such as epirubicin, oxaliplatin, irinotecan and gemcitabine, among others. Given that CIN represents approximately 70% of gastric cancer cases this analysis suggests that improved survival in a large proportion of gastric cancers could be achieved by adopting targeted therapy techniques against EGFR signaling.

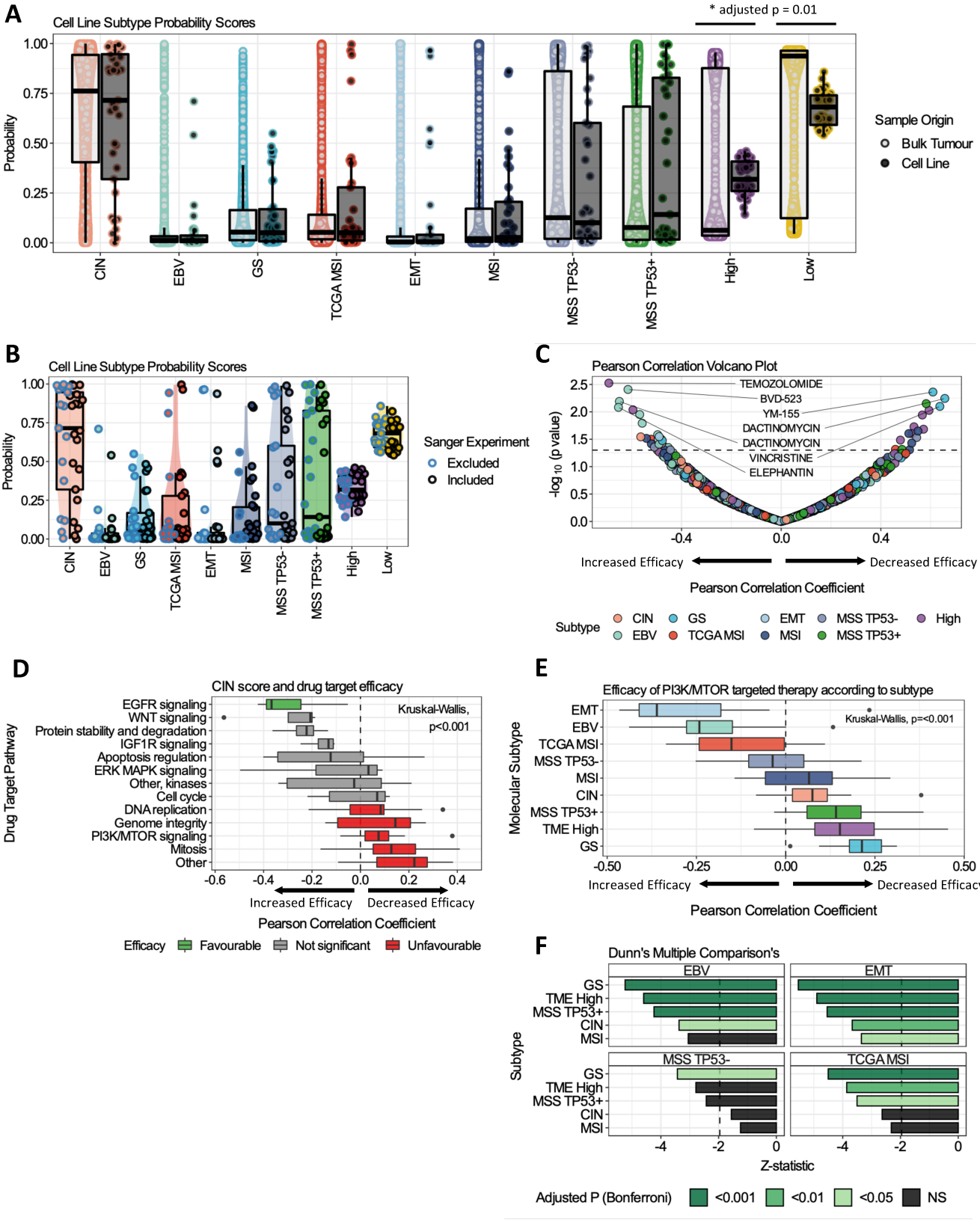

A
Cell Line Subtype Probability Scores
* adjusted p = 0.01
Probability
1.00
0.75
0.50
0.25
0.00
CIN
EBV
GS
TCGA MSI
EMT
MSI
MSS TP53-
MSS TP53+
High
Low
Sample Origin
Bulk Tumour
Cell Line
B
Cell Line Subtype Probability Scores
Probability
Sanger Experiment
Excluded
Included
C
Pearson Correlation Volcano Plot
TEMOZOLOMIDE
BVD-523
YM-155
DACTINOMYCIN
DACTINOMYCIN
VINCRISTINE
ELEPHANTIN
-log10 (p value)
Increased Efficacy
Decreased Efficacy
Pearson Correlation Coefficient
Subtype
CIN
EBV
GS
TCGA MSI
EMT
MSI
MSS TP53-
MSS TP53+
High
D
CIN score and drug target efficacy
Drug Target Pathway
EGFR signaling
WNT signaling
Protein stability and degradation
IGF1R signaling
Apoptosis regulation
ERK MAPK signaling
Other, kinases
Cell cycle
DNA replication
Genome integrity
PI3K/MTOR signaling
Mitosis
Other
Kruskal-Wallis, p<0.001
Increased Efficacy
Decreased Efficacy
Pearson Correlation Coefficient
Efficacy
Favourable
Not significant
Unfavourable
E
Efficacy of PI3K/MTOR targeted therapy according to subtype
Molecular Subtype
EMT
EBV
TCGA MSI
MSS TP53-
MSI
CIN
MSS TP53+
TME High
GS
Kruskal-Wallis, p=<0.001
Increased Efficacy
Decreased Efficacy
Pearson Correlation Coefficient
F
Dunn's Multiple Comparison's
EBV
EMT
MSS TP53-
TCGA MSI
Subtype
GS
TME High
MSS TP53+
CIN
MSI
Z-statistic
Adjusted P (Bonferroni)
<0.001
<0.01
<0.05
NS

Figure 5-5 Landscape of the role of 2D gastric cancer cell lines in pre-clinical models for personalized medicine. **A.** Boxplot comparing the distributions of subtype scores for each molecular subtype between bulk tumour tissue and 2D cell lines. Bulk tumour distributions were determined from 2,202 publicly available gastric cancer patients. Cell lines included 37 gastric cancer cell lines used in the CCLE. Bars and p values denote statistically significant differences using a Wilcox test. **B.** Boxplot comparing the distributions of subtype scores for cell lines included in GDSC drug screen to those excluded for each molecular subtype. **C.** Volcano plot for univariate Pearson's correlation of molecular subtype scores and area under the curve drug sensitivity for 111 cancer drugs. The dotted line identifies the line of alpha = 0.05 for a significant unadjusted p value. Annotated drugs are those with the largest effect size. **D.** Boxplot of drug target pathway versus Pearson's correlation coefficient for CIN subtype probability. A post-hoc Dunn's test was performed. The coloured boxes identify significant favourable (green) and unfavourable (red) efficacy of drug targets in the post-hoc test with Bonferroni corrected P values. **E.** Boxplot of molecular subtype versus Pearson's correlation coefficient for PI3K/MTOR signaling. Favourable efficacy is represented by a negative score. **F.** Bar plots illustrating results of Dunn's post-hoc test from plot (E). Statistical significance is denoted by the plot legend for subtypes that achieved any significant results.

In Figure 5-5E, we provide an alternative perspective by investigating the effect of a given target across all molecular subtypes. The complete analysis for all subtypes and targets is available in the online supplement. In this case, Pi3K/MTOR signaling targets are efficacious in EMT, EBV, TCGA MSI and MSS TP53- tumours relative to all others (Figure 5-5F). Once again, this analysis is consistent with prior studies that found TCGA EBV subtypes are enriched in PI3K mutations.[57]

5.4 Discussion

In this study we assessed the ability of a custom 107 gene Nanostring assay to learn TCGA and TME molecular subtypes in FFPE and patient-derived organoid gastric cancer tissue. This test was informed by our previously characterized supervised machine learning classifiers which were demonstrated to accurately represent intended molecular phenomena. Here we used gold-standard reference tests EBV ISH and pentaplex PCR for EBV and MSI type tumours, respectively, to demonstrate that our test is 100% accurate in capturing these subtypes in paired tumour and tumour organoid samples. Although we did not assess copy number variation or other characteristics associated with whole genome sequencing, we were able to approximate assay performance for GS scores. Among 39 FFPE gastric cancer specimens our results were concordant with prior research demonstrating

that diffuse-type gastric cancer is significantly enriched in GS scores. Furthermore, we identified nearly identical proportion of TCGA subtypes within our population compared to a cohort of 2,202 gastric cancer patients. Although we found fewer TME High tumours than expected, the increased presence of TME High subtypes with MSI and EBV tumours is also concordant with other populations. Together, these results demonstrate that our Nanostring test is accurate and feasible to perform additional study and validation.

The advantage of our test is that it provides accurate and concordant measurement of molecular subtypes in FFPE and fresh tumour tissue. Thus, our test can serve as a translational measurement tool for future gastric cancer research. The Nanostring platform does not require reverse transcription of mRNA and therefore is able to provide accurate gene counts using fragmented FFPE-derived RNA. Furthermore, we designed our initial machine learning models with the intent of developing a downstream translational test. Previous literature has demonstrated that concordance in gene counts between fresh and FFPE tissue is greatest when assessing high variance and high abundance genes.[342,343] Thus, we actively filtered troublesome genes prior to generating our models.

Appropriate normalization and translation of our Nanostring data was essential to achieving our results. Failure to adequately normalize data and account for heterogeneity in gene measurements across separate technological platforms is a significant problem in developing reproducible and reliable tests.[344,345] Here we implemented novel approaches to Nanostring normalization using the framework proposed by Molania et al.[334] This method does not rely on spike-in control probes or housekeeping genes. Instead, it assesses variation between genes of interest to define negative control genes that are then utilized in the RUV-III algorithm.

The next major challenge was accounting for cross-platform effects in our gene expression data. Our machine learning classifiers were constructed using distributions of data unique to their respective RNA-seq and microarray platforms. Thus, to effectively use these models they needed to be supplied with data that follows the same distribution from that which they were developed. To accomplish this task, we used Feature Specific Quantile Normalization. A desirable property of FSQN is that it only requires genes that are common between distributions to work. Thus, we were able to approximate our 50 and 57 gene measurements directly to the target distributions of each model.

Given the accuracy of our test we were able to apply it to personalized medicine models such as patient-derived organoids. Despite nearly a decade of promising research, the potential clinical utility of organoids was recently questioned by the SENSOR trial.[174] One important question in organoid research is the degree to which an organoid recapitulates the molecular characteristics of the parent tumour. Numerous studies have demonstrated similarity between gene expression profiles in tumour organoid pairs.[180,182] However, in this study we assessed whether this approximation of gene expression data actually translated to similar molecular subtype assignments. Using our test, we identified that significant molecular subtype discordance between paired tumour and tumour organoids can occur. This result is troublesome but not surprising given the presence of intratumour heterogeneity, issues with biopsy accuracy and selective pressures exerted by tissue processing, transport and constituents of organoid media itself. Additional research investigating the dynamic nature of molecular subtype classes between parent tumour-organoid pairs would enhance our understanding of potential confounding in pre-clinical experiments.

In this study we also assessed if more efficient and affordable 2-dimensional cell culture can still provide a valuable pre-clinical model. We tested the landscape of TCGA, ACRG and TME molecular subtypes, as informed by our models, in 37 gastric cancer cell lines from the CCLE.[338] With this analysis we demonstrated that representative distributions between a large publicly available clinical cohort and cell lines are present for all molecular subtypes except for those pertaining to the tumour immune microenvironment. Using a large drug screening assay performed by the GDSC, we also demonstrated that the strongest association between dose response and molecular subtype scores is typically observed with molecularly targeted therapy as opposed to traditional cytotoxic chemotherapy targeting DNA replication.[341] Thus, 2D cell culture may still provide an important pipeline to identify relevant therapeutic targets that can subsequently be validated using more expensive and cumbersome 3D cell culture or xenograft models.

Our study contains multiple limitations. We were not able to include a "gold-standard" reference for CIN and GS tumours. Additional validation of these subtype using whole genome sequencing is required in the future. Another limitation is that we do not have adequate follow-up time to perform a direct comparison of dose-response to progression free survival or overall survival. Once adequate follow up time is achieved correlation of *in-vitro* dose-response to tangible clinical outcomes should be performed.

5.5 Conclusion

Our 107 gene Nanostring assay informed by previously characterized supervised machine learning classifiers allows molecular classification of TCGA and TME subtypes in FFPE or fresh gastric cancer tissue. Our analyses suggest that a multimodal approach, which may consist of molecularly informed clinical samples, 2D and 3D cell culture models, is required to leverage molecular classification towards personalized medicine in gastric cancer. Promising results from this initial external validation should be expanded in additional and more robust studies.

Chapter 6: General Discussion and Future Study

In this thesis we explored several concepts related to personalized medicine in gastric cancer. Through these analyses we either improved our understanding of gastric cancer prognosis and treatment, or, provided a foundation for future study. In this section we will provide a summary of our findings and their relationship to the literature as well as considerations for future study.

In Chapter 2, we performed a pilot study to investigate the utility of immunohistochemistry-based biomarkers in predicting neoadjuvant chemotherapy response. In western nations, including Canada, neoadjuvant chemotherapy is the standard of care for locally advanced gastric cancer given its propensity to downsize bulky tumours and facilitate adequate surgical resection. However, treatment response is variable and only approximately 40% of patients achieve complete or partial treatment response. Thus, the majority of patients given neoadjuvant chemotherapy may experience adverse medical events without sufficient benefit. Given this scenario, there is a need to develop methods that allow appropriate allocation of patients to neoadjuvant chemotherapy in order to optimize survival outcomes.

First, we assessed the expression of E-cadherin, Galectin-3, CD4 and CD8 in forty-three gastric cancer patients from all stages. This survey of Canadian gastric cancer patients showed that CD4, CD4/CD8 ratio, Galectin-3 and E-cadherin expression is significantly different between cancer and normal tissue prior to chemotherapy. Eighteen patients underwent neoadjuvant chemotherapy in our cohort. We characterized the relative effect of neoadjuvant FLOT on biomarker expression in fifteen pre- and post-chemotherapy tumours. Notably, we observed increase infiltration of $CD4^+$ and $CD8^+$ lymphocytes following chemotherapy, which suggests that FLOT chemotherapy may facilitate anti-tumour immune response. This finding is consistent with more thorough analysis completed in breast cancer neoadjuvant chemotherapy. In Park et al., dynamic immune responses following neoadjuvant chemotherapy in breast cancer were observed and specifically, increased infiltration of TILs and $CD8^+$ lymphocytes were associated with greater pathologic response.[346]

In this work, we are the first to report that the CD4/CD8 ratio is significantly associated with complete or partial pathologic response to neoadjuvant chemotherapy in gastric cancer. In preliminary univariate and multivariate machine learning models, increased CD4/CD8 ratio provided significant predictive benefit. Central to our analysis was the use of digital

pathology using QuPath, which is a useful tool to enhance data acquisition for personalized medicine.[347–349] Beyond expanding our work based on our sample size calculations, future research should focus on establishing reproducible digital pathology protocols and assessing inter-rater reliability.

Organoids are a promising tool to advance personalized medicine. In chapter 3 we explored the feasibility of shipping fresh tissues to enhance the procurement of organoid samples. In this proof of concept study, we successfully established and propagated normal mouse gastric organoids after 24 or 48 hours of mock-shipment on ice. We assessed the effect of cold-ischemia time and various transport media on organoid viability and growth. Although increased cold-ischemia time was associated with decreased growth rate, organoid viability and dose-response to cytotoxic therapy was unaffected. We observed minimal differences in the long-term growth of organoids transported in HBSS, UW and HTK solutions. Given these findings we chose to use HBSS as transport solution for future experiments due to easier accessibility and affordability. Using our mock-shipment model we were also able to establish and grow two human gastric cancer organoids after 24 and 48 hours of cold-ischemia time.

In our mouse organoid mock-shipment model we assessed the longitudinal expression of important gastric stem cell markers TROY and LGR5. Beyond their essential role in organoid culture, these stem cell markers also have potential clinical implications related to chemoresistance.[257] To our knowledge, we are the first to establish that LGR5 and TROY expression initially decreases upon dissociation of tissue but subsequently increases over successive organoid passage to approximate endogenous levels by passage 6. Furthermore, we assessed the effect of passage number and stem cell gene expression on the *in-vitro* dose-response of mock-shipped organoids to FOLFOX therapy. Here, we found that organoid dose-response is unaffected by transport time and media. However, decreased cytotoxic effects were identified with increasing organoid passage and increasing TROY and LGR5 gene expression. The causal relationship between passage number, stem cell gene expression and dose-response remain unknown and could be a potential area of future study. Regardless, both passage number and stem cell gene expression should be considered as important confounders in evaluating organoid dose-response assays. These findings need to be expanded to cancer organoid models.

Molecular classification of gastric cancer has significantly shifted the landscape of personalized medicine for this deadly disease. Several molecular classification systems exist

but prominent classification systems include the TCGA, ACRG and TME score.[26,57,207] In this thesis, we recognized that the clinical implementation of molecular classification for gastric cancer is lacking due to several limitations. The cost and logistical issues behind data pipelines and processing suggests that widespread implementation of next-generation sequencing for multi-omics molecular classification is imminently distant, especially for public health care systems.[201] Thus, in chapter 4 and 5 we aimed to develop a translational gene expression-based test to assign molecular classification subtypes to clinical samples in an efficient and affordable manner.

Chapter 4 intended to address important challenges facing molecular classification in gastric cancer. Although several classification systems exist we do not know which classification systems are helpful and which are redundant. Thus, we prioritized integrated analysis of TCGA, ACRG and TME molecular subtypes. We also identified issues related to the common practice of using discrete molecular classes and models centred on estimating population-based effects.[268,270,308] In our study we actively interpreted molecular subtypes as continuous variables and demonstrated their superior statistical properties.

We designed and tested a series of supervised machine learning classifiers using whole-transcriptome data to assign TCGA, ACRG and TME subtypes. Central to our analysis was the use of unbiased cross-validation methods to ensure there was no data leakage to falsely inflate accuracy.[293,294] From our selected models we performed an in-depth analysis of calibration in the multiclass setting. We identified discrepancies between various calibration metrics and the need for additional calibration methods to optimize machine learning for personalized medicine.[299,301,302,330] The accuracy of our optimally calibrated models were 89.5 %, 89.4 % and 84.7% for TCGA, TME and ACRG classification, respectively.

Small sample size is a significant barrier to identifying clinically relevant molecular subtypes. For example, initial development datasets for TCGA, ACRG and TME subtypes only consist of 295-300 patients.[26,57,207] Once again this is related to massive costs of multi-omics or whole transcriptome technology. To address this problem, we used our machine learning classifier models to assign TCGA, ACRG and TME subtypes to 2,202 publicly available gastric cancer patients. The fundamental technique to properly assign subtypes across multiple gene expression technologies was a method developed by Franks et al. called Feature Specific Quantile Normalization (FSQN).[297] In the original study, FSQN was demonstrated to assign molecular subtypes with high concordance between RNA-seq and microarray data compared to controls and other normalization methods. Even more, FSQN is predicated on

normalizing data using only common features/genes. Thus, this method allows normalization of gene expression data with smaller number of genes compared to traditional normalization techniques such as quantile normalization. Ultimately this provided opportunity for downstream application of our machine learning classifiers to cross-platform technologies such as Nanostring.

To our knowledge, our 2,202 patient cohort is the largest integrated assessment of three independent molecular classification systems. Using these data, we demonstrated that the TME High score is the most significant prognostic and therapeutic molecular subtype biomarker. In a propensity score matched cohort consisting of Stage II-IV gastric cancer patients we identified that greater TME High score is associated with increasing efficacy of chemotherapy with respect to overall survival. Limitations of this analysis include heterogeneity in treatment regimens between studies and the potential of unmeasured confounding. Additional questions remain as to why certain molecular subtypes have a greater propensity to interact with the immune system.

We identified a distinct relationship between the proportion of TME High tumours and overall survival for each molecular subtype. For example, even among molecularly unique tumours such as MSI and EMT, stratification of subtypes by TME status demonstrated profound survival advantages for TME High tumours. We do not understand why a small proportion of MSI tumours remain TME Low. Additional research on the effect of TME status in MSI tumours could provide insight into debate surrounding the utility of neoadjuvant chemotherapy in MSI gastric cancer.[262,263]

Additional survival analysis was performed using individual survival distribution models. We compared the utility and interpretation of overall survival using novel MTLR models versus traditional Cox proportional hazards model. Using MTLR, we are able to illustrate survival curves for each individual patient that provide a probability of survival for any future time point. We conclude that individual survival distribution models, such as MTLR, are a powerful tool to communicate personalized medicine to patients and clinicians. We advocate for the use of individual survival distributions in future research.

Following our integrated molecular classification analysis in Chapter 5 we aimed to translate our machine learning models to our own population. We developed a custom Nanostring codeset for TCGA and TME subtypes using the genes selected by our models in Chapter 4. Before creating our models in Chapter 4 we recognized that eventual translation to a clinical

test could one day provide greater clinical uptake of molecular classification. Thus, we purposefully selected high abundance and high variance genes from each public database.[342,343] We chose FSQN because it provides normalization among public datasets with whole transcriptome data and with Nanostring gene expression data.

Nanostring is a desirable technology to use for development of clinical tests. The utility of Nanostring in evaluating FFPE specimens and fragmented RNA has been well documented.[334] Thus, a test validated with Nanostring is likely applicable to almost all future clinical specimens and can even be used for retrospective analysis of prior research. Furthermore, novel normalization methods for Nanostring assays, as proposed by Molania et al. do not require traditional housekeeping genes.[334] Instead variation among genes of interest is interpreted and a subset of these genes are used within the RUV-II algorithm to achieving superior normalization. This means that the 57 genes we used for TCGA classification are essentially isolated in the analysis process. Thus, future gene sets derived from other models could be added to or removed from a given Nanostring assay. This plug and play approach allow optimally designed genesets to be combined to maximize clinical utility.

We established 12 tumour and tumour organoid pairs to externally validate our Nanostring test. We performed gold-standard EBER ISH and multiplex PCR to test for EBV and MSI tumours. Next, we allocated molecular subtypes to our samples using our Nanostring test. In the public cohort of 2,202 and our cohort, we found >98.7% concordance of our EBV and MSI subtypes measured by our test in reference to gold-standard results. Next, we illustrated that some organoids do not recapitulate molecular subtypes from their parent tumour. This was a significant finding. It is likely that the appropriate recapitulation of the parent tumour's molecular subtype is an important confounder in studies aiming to translate *in-vitro* organoid drug response directly to clinical therapy. The true effect of molecular subtype discordance between tumour and tumour organoid pairs certainly requires additional research.

Next, we used this accurate test to assign molecular subtypes to 39 patients from our prospective clinical cohort in Chapter 1. Here, we found that the proportion of TCGA subtypes (CIN, GS, MSI and EBV) in our population was nearly identical to that in the public 2,202 patient cohort. Furthermore, we identified a decreased proportion of TME High tumours relative to the public cohort (12.8% vs. 33.2%). Application of our test to a larger sample is required to determine if this discrepancy is legitimate. In this limited sample we

could argue that we may have less TME High tumours in our population given that we appropriately observed a prominence of TME High tumours in MSI and EBV tumours.

As a proof of concept, we performed an *in-vitro* organoid dose-response assay using FLOT chemotherapy and compared drug sensitivity to molecular subtype scores. Given the non-specific cytotoxic effect of chemotherapy we were not surprised to find limited relationship between drug sensitivity and subtype scores. We hypothesized that significant association between molecular subtype scores and dose-response would be observed in the setting of targeted therapy.

Using The Cancer Cell Line Encyclopedia we were able to assign TCGA, ACRG and TME molecular subtypes to 37 gastric cancer cell lines and assess their relationship to dose responses in 111 anti-cancer drugs.[338,341] First, we established that the distribution of molecular subtype scores for TCGA and ACRG subtypes among these 37 gastric cancer cell lines approximates the distribution of scores observed in the 2,202 patient public cohort. This finding suggests that an appropriate selection of 2D cell lines provides an accurate representation of TCGA and ACRG intertumour heterogeneity in gastric cancer patients. Within the body of this thesis we provide an example of the utility of this approach. For example, we identify that EGFR therapy could be an effective treatment strategy in CIN tumours and that therapy targeting mitosis or PI3K/MTOR signaling should be avoided. Additional analyses for all drug and subtype combinations can be explored using the online supplement.

We also observed that TME subtype scores are not adequately represented. It is not surprising that 2D cell lines, which consist of homogenous epithelial cells, are not represented by a classification system designed to identify immune and tumour microenvironment factors. Unfortunately, the TME signature is arguably the most important molecular subtype for prognosis and treatment in gastric cancer. Thus, 2D cell lines may provide a prominent role in the discovery of certain therapies but more comprehensive analyses that consider the TME will require organoid, *in-vitro* immune cell co-culture or humanized xenograft models.[189,350,351]

Despite a strong theoretical basis, the excitement surrounding personalized medicine does not make it empiric reality. In this thesis I attempted to use a multimodal approach to enhance personalized medicine in gastric cancer and provide a foundation for future study. The development of personalized medicine in gastric cancer is a rapidly evolving field.

Ongoing clinical trials aim to establish the efficacy of immunotherapy in the neoadjuvant setting.[1,2] Particular attention to EBV and MSI high gastric cancer as subpopulations for treatment efficacy show promise.[206] However, current investigations into targeted therapy only benefits approximately 20-30% of gastric cancer patients (considering the prevalence of EBV, MSI and/or TME High tumours). Furthermore, EBV and MSI tumours already possess excellent outcomes. Thus, immunotherapy is serving to improve outcomes in patients who already experience superior outcomes. The efficacy of these treatments in the context of TME High scores in EBV and MSI gastric cancer is unknown and could further improve treatment assignment. In contrast, subtypes that carry the worst prognosis such as CIN, GS or EMT subtypes have no recommended or suggested targeted therapy regimens. For example, beyond HER2 therapy, there is limited perspective on how to advance care in CIN tumours. Advancement of care in these subtypes is paramount as they represent a majority of gastric cancer patients. Even if immunotherapy improves neoadjuvant treatment, improved prediction methods using biomarkers or molecular subtypes are required to optimize allocation of patients to existing cytotoxic based regimens such as FLOT. Given that only approximately 40% of patients achieve complete or partial treatment response, the majority of patients may not be receiving appropriate therapy.[213,223,232]

Within this current landscape of personalized medicine in gastric cancer our findings provide guidance for future research. Additional study of the CD4/CD8 Ratio in larger patient cohorts may provide a relatively simple biomarker approach to predict treatment response to neoadjuvant chemotherapy. Implementation of our tissue shipment methods could be used to increase procurement of organoid tissue for translational research. Further characterization of the relationship between organoid passage number and stem cell gene expression may enhance our understanding of *in-vitro* dose-response. Finally, our custom Nanostring test provides opportunity for additional integrated analysis of molecular classification systems in pre-clinical 2D cell culture, translational 3D organoid culture and/or our gastric cancer patient population.

References

1. Bang Y-J, Van Cutsem E, Fuchs CS, Ohtsu A, Tabernero J, Ilson DH, et al. KEYNOTE-585: Phase III study of perioperative chemotherapy with or without pembrolizumab for gastric cancer. Future Oncol. 2019 Mar;15(9):943–52.
2. Kennedy T, Shah MM, In H, Chuy JW, Moore DF, Kooby DA, et al. Preoperative pembrolizumab for MSI high, EBV positive or PD-L1 positive locally advanced gastric cancer followed by surgery and adjuvant chemoradiation with pembrolizumab: Interim results of a phase 2 multi-center trial (NCT03257163). J Clin Oncol [Internet]. 2021 May 20;39(15_suppl):e16111–e16111. Available from: https://doi.org/10.1200/JCO.2021.39.15_suppl.e16111
3. Leja M, Linē A. Early detection of gastric cancer beyond endoscopy - new methods. Best Pract Res Clin Gastroenterol. 2021;50–51:101731.
4. Desiderio J, Chao J, Melstrom L, Warner S, Tozzi F, Fong Y, et al. The 30-year experience-A meta-analysis of randomised and high-quality non-randomised studies of hyperthermic intraperitoneal chemotherapy in the treatment of gastric cancer. Eur J Cancer. 2017 Jul;79:1–14.
5. Götze TO, Piso P, Lorenzen S, Bankstahl US, Pauligk C, Elshafei M, et al. Preventive HIPEC in combination with perioperative FLOT versus FLOT alone for resectable diffuse type gastric and gastroesophageal junction type II/III adenocarcinoma – the phase III "PREVENT"- (FLOT9) trial of the AIO /CAOGI /ACO. BMC Cancer. 2021;21(1):1–9.
6. Waddingham W, Nieuwenburg SA V, Carlson S, Rodriguez-Justo M, Spaander M, Kuipers EJ, et al. Recent advances in the detection and management of early gastric cancer and its precursors. Frontline Gastroenterol [Internet]. 2021 Jul 1;12(4):322 LP – 331. Available from: http://fg.bmj.com/content/12/4/322.abstract
7. Pietrantonio F, Miceli R, Raimondi A, Kim YW, Kang WK, Langley RE, et al. Individual patient data meta-analysis of the value of microsatellite instability as a biomarker in gastric cancer. J Clin Oncol. 2019;37(35):3392–400.
8. Haag GM, Czink E, Ahadova A, Schmidt T, Sisic L, Blank S, et al. Prognostic significance of microsatellite-instability in gastric and gastroesophageal junction cancer patients undergoing neoadjuvant chemotherapy. Int J cancer. 2019 Apr;144(7):1697–703.
9. Schiappacasse Cocio G V, Schiappacasse ED. Is Adjuvant Chemotherapy Efficient in Colon Cancer with High Microsatellite Instability? A Look Towards the Future. Cancer Res [Internet]. 2019 Feb 1;79(3):441–4. Available from: https://doi.org/10.1158/0008-5472.CAN-18-2991
10. Cohen R, Taieb J, Fiskum J, Yothers G, Goldberg R, Yoshino T, et al. Microsatellite Instability in Patients With Stage III Colon Cancer Receiving Fluoropyrimidine With or Without Oxaliplatin: An ACCENT Pooled Analysis of 12 Adjuvant Trials. J Clin Oncol. 2021;39(6):642–51.
11. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. CA Cancer J Clin. 2015 Mar;65(2):87–108.
12. Todua F, Gagua R, Maglakelidze M, Maglakelidze D. Cancer incidence and mortality - Major patterns in GLOBOCAN 2012, worldwide and Georgia. Bull Georg Natl Acad Sci. 2015;9(1):168–73.
13. Stewart B, Wild International Agency for Research on Cancer, WHO CP (eds. . World Cancer Report 2014 [Online]. . 2014;
14. Bertuccio P, Chatenoud L, Levi F, Praud D, Ferlay J, Negri E, et al. Recent patterns in gastric cancer: a global overview. Int J cancer. 2009 Aug;125(3):666–73.
15. Parkin DM, Stjernsward J, Muir CS. Estimates of the worldwide frequency of twelve major cancers. Bull World Health Organ. 1984;62(2):163–82.

16. Teitelbaum EN, Hungness ES, Mahvi DM. Stomach [Internet]. Twentieth. Sabiston Textbook of Surgery. Elsevier Inc.; 2018. 1188–1236 p. Available from: http://dx.doi.org/10.1016/B978-0-323-29987-9.00048-5
17. National Cancer Institute. SEER Cancer Stat Facts: Stomach Cancer. [Internet]. Bethesda. MD; Available from: https://seer.cancer.gov/statfacts/html/stomach.html
18. Canadian Cancer Statistics Advisory Committee. Canadian Cancer Statistics. Can Cancer Soc [Internet]. 2018;52. Available from: cancer.ca/Canadian-Cancer-Statistics
19. Devesa SS, Blot WJ, Fraumeni JFJ. Changing patterns in the incidence of esophageal and gastric carcinoma in the United States. Cancer. 1998 Nov;83(10):2049–53.
20. Levi F, Randimbison R, La Vecchia C. Esophageal and gastric carcinoma in Vaud, Switzerland, 1976-1994. Vol. 75, International journal of cancer. United States; 1998. p. 160–1.
21. Jemal A, Center MM, DeSantis C, Ward EM. Global patterns of cancer incidence and mortality rates and trends. Cancer Epidemiol Biomarkers Prev. 2010;19(8):1893–907.
22. Powell J, McConkey CC. Increasing incidence of adenocarcinoma of the gastric cardia and adjacent sites. Br J Cancer. 1990 Sep;62(3):440–3.
23. Ma J, Shen H, Kapesa L, Zeng S. Lauren classification and individualized chemotherapyin gastric cancer (Review). Oncol Lett. 2016;11(5):2959–64.
24. Berlth F, Bollschweiler E, Drebber U, Hoelscher AH, Moenig S. Pathohistological classification systems in gastric cancer: Diagnostic relevance and prognostic value. World J Gastroenterol. 2014;20(19):5679–84.
25. LAUREN P. THE TWO HISTOLOGICAL MAIN TYPES OF GASTRIC CARCINOMA: DIFFUSE AND SO-CALLED INTESTINAL-TYPE CARCINOMA. AN ATTEMPT AT A HISTO-CLINICAL CLASSIFICATION. Acta Pathol Microbiol Scand. 1965;64:31–49.
26. Cristescu R, Lee J, Nebozhyn M, Kim K-M, Ting JC, Wong SS, et al. Molecular analysis of gastric cancer identifies subtypes associated with distinct clinical outcomes. Nat Med [Internet]. 2015;21(5):449–56. Available from: http://www.nature.com/doifinder/10.1038/nm.3850
27. Qiu MZ, Shi SM, Chen M, Wang J, Wu QN, Sheng H, et al. Comparison of HER2 and lauren classification between biopsy and surgical resection samples, primary and metastatic samples of Gastric Cancer. J Cancer. 2017;8(17).
28. Kim H, An JY, Noh SH, Shin SK, Lee YC, Kim H. High microsatellite instability predicts good prognosis in intestinal-type gastric cancers. J Gastroenterol Hepatol. 2011;26(3):585–92.
29. Qiu MZ, Cai MY, Zhang DS, Wang ZQ, Wang DS, Li YH, et al. Clinicopathological characteristics and prognostic analysis of Lauren classification in gastric adenocarcinoma in China. J Transl Med [Internet]. 2013;11(1):1. Available from: Journal of Translational Medicine
30. Chen YC, Fang WL, Wang RF, Liu CA, Yang MH, Lo SS, et al. Clinicopathological Variation of Lauren Classification in Gastric Cancer. Pathol Oncol Res. 2016;22(1):197–202.
31. Wang H, Xing X, Ma L, Liu L, Hao J, Feng L, et al. Metastatic lymph node ratio and Lauren classification are independent prognostic markers for survival rates of patients with gastric cancer. Oncol Lett [Internet]. 2018;8853–62. Available from: http://www.spandidos-publications.com/10.3892/ol.2018.8497
32. Correa P, Piazuelo MB. The gastric precancerous cascade. J Dig Dis. 2012;13(1):2–9.
33. Correa P. A human model of gastric carcinogenesis. Cancer Res. 1988 Jul;48(13):3554–60.
34. Pittman ME, Voltaggio L, Bhaijee F, Robertson SA, Montgomery EA. Autoimmune Metaplastic Atrophic Gastritis: Recognizing Precursor Lesions for Appropriate Patient Evaluation. Am J Surg Pathol [Internet]. 2015;39(12):1611–20. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=

26291507
35. Hunt RH, Yaghoobi M. The Esophageal and Gastric Microbiome in Health and Disease. Gastroenterol Clin North Am [Internet]. 2017;46(1):121–41. Available from: http://dx.doi.org/10.1016/j.gtc.2016.09.009
36. Noto JM, Peek RM. The gastric microbiome, its interaction with Helicobacter pylori, and its potential role in the progression to stomach cancer. PLoS Pathog. 2017;13(10):3–9.
37. Klinkenberg-Knol EC, Nelis F, Dent J, Snel P, Mitchell B, Prichard P, et al. Long-term omeprazole treatment in resistant gastroesophageal reflux disease: efficacy, safety, and influence on gastric mucosa. Gastroenterology. 2000 Apr;118(4):661–9.
38. Lundell L, Miettinen P, Myrvold HE, Pedersen SA, Thor K, Andersson A, et al. Lack of effect of acid suppression therapy on gastric atrophy. Nordic Gerd Study Group. Gastroenterology. 1999 Aug;117(2):319–26.
39. Busuttil RA, Boussioutas A. Intestinal metaplasia: A premalignant lesion involved in gastric carcinogenesis. J Gastroenterol Hepatol. 2009;24(2):193–201.
40. Boussioutas A, Li H, Liu J, Waring P, Lade S, Holloway AJ, et al. Distinctive patterns of gene expression in premalignant gastric mucosa and gastric cancer. Cancer Res. 2003 May;63(10):2569–77.
41. Conchillo JM, Houben G, de Bruine A, Stockbrugger R. Is type III intestinal metaplasia an obligatory precancerous lesion in intestinal-type gastric carcinoma? Eur J Cancer Prev. 2001 Aug;10(4):307–12.
42. Rokkas T, Pistiolas D, Sechopoulos P, Robotis I, Margantinis G. The long-term impact of Helicobacter pylori eradication on gastric histology: A systematic review and meta-analysis. Helicobacter. 2007;12(SUPPL. 2):32–8.
43. Rugge M, Farinati F, Baffa R, Sonego F, Di Mario F, Leandro G, et al. Gastric epithelial dysplasia in the natural history of gastric cancer: a multicenter prospective follow-up study. Interdisciplinary Group on Gastric Epithelial Dysplasia. Gastroenterology. 1994 Nov;107(5):1288–96.
44. Sala N, Muñoz X, Travier N, Agudo A, Duell EJ, Moreno V, et al. Prostate stem-cell antigen gene is associated with diffuse and intestinal gastric cancer in Caucasians: Results from the EPIC-EURGAST study. Int J Cancer [Internet]. 2011 Jun 16;130(10):2417–27. Available from: https://doi.org/10.1002/ijc.26243
45. Saeki N, Sakamoto H, Yoshida T. Mucin 1 gene (MUC1) and gastric-cancer susceptibility. Int J Mol Sci. 2014 May;15(5):7958–73.
46. Saeki N, Ono H, Sakamoto H, Yoshida T. Genetic factors related to gastric cancer susceptibility identified using a genome-wide association study. Cancer Sci. 2013 Jan;104(1):1–8.
47. Worthley DL, Phillips KD, Wayte N, Schrader KA, Healey S, Kaurah P, et al. Gastric adenocarcinoma and proximal polyposis of the stomach (GAPPS): a new autosomal dominant syndrome. Gut. 2012 May;61(5):774–9.
48. Oliveira C, Pinheiro H, Figueiredo J, Seruca R, Carneiro F. Familial gastric cancer: genetic susceptibility, pathology, and implications for management. Lancet Oncol. 2015 Feb;16(2):e60-70.
49. Syngal S, Brand RE, Church JM, Giardiello FM, Hampel HL, Burt RW. ACG clinical guideline: Genetic testing and management of hereditary gastrointestinal cancer syndromes. Am J Gastroenterol. 2015 Feb;110(2):223–62; quiz 263.
50. van der Post RS, Vogelaar IP, Carneiro F, Guilford P, Huntsman D, Hoogerbrugge N, et al. Hereditary diffuse gastric cancer: updated clinical guidelines with an emphasis on germline CDH1 mutation carriers. J Med Genet. 2015 Jun;52(6):361–74.
51. Fitzgerald RC, Hardwick R, Huntsman D, Carneiro F, Guilford P, Blair V, et al. Hereditary diffuse gastric cancer: updated consensus guidelines for clinical management and directions for future research. J Med Genet. 2010 Jul;47(7):436–44.

52. Graziano F, Humar B, Guilford P. The role of the E-cadherin gene (CDH1) in diffuse gastric cancer susceptibility: from the laboratory to clinical practice. Ann Oncol Off J Eur Soc Med Oncol. 2003 Dec;14(12):1705–13.
53. Oliveira C, Sousa S, Pinheiro H, Karam R, Bordeira-Carrico R, Senz J, et al. Quantification of epigenetic and genetic 2nd hits in CDH1 during hereditary diffuse gastric cancer syndrome progression. Gastroenterology. 2009 Jun;136(7):2137–48.
54. Barber M, Murrell A, Ito Y, Maia A-T, Hyland S, Oliveira C, et al. Mechanisms and sequelae of E-cadherin silencing in hereditary diffuse gastric cancer. J Pathol. 2008 Nov;216(3):295–306.
55. Carneiro F, Huntsman DG, Smyrk TC, Owen DA, Seruca R, Pharoah P, et al. Model of the early development of diffuse gastric cancer in E-cadherin mutation carriers and its implications for patient screening. J Pathol. 2004 Jun;203(2):681–7.
56. Franco AT, Israel DA, Washington MK, Krishna U, Fox JG, Rogers AB, et al. Activation of beta-catenin by carcinogenic Helicobacter pylori. Proc Natl Acad Sci U S A. 2005 Jul;102(30):10646–51.
57. Cancer Genome Atlas Research Network, Bass AJ, Thorsson V, Shmulevich I, Reynolds SM, Miller M, et al. Comprehensive molecular characterization of gastric adenocarcinoma. Nature. 2014;513(7517):202–9.
58. Piessen G, Messager M, Leteurtre E, Jean-Pierre T, Mariette C. Signet ring cell histology is an independent predictor of poor prognosis in gastric adenocarcinoma regardless of tumoral clinical presentation. Ann Surg. 2009 Dec;250(6):878–87.
59. Research WCRF/ AI for C. Food, Nutrition, Physical Activity, and the Prevention of Cancer: a Global Perspective. Washington DC: AICR, 2007;
60. Yang P, Zhou Y, Chen B, Wan H-W, Jia G-Q, Bai H-L, et al. Overweight, obesity and gastric cancer risk: results from a meta-analysis of cohort studies. Eur J Cancer. 2009 Nov;45(16):2867–73.
61. Lauby-Secretan B, Scoccianti C, Loomis D, Grosse Y, Bianchini F, Straif K, et al. Body Fatness and Cancer--Viewpoint of the IARC Working Group. N Engl J Med. 2016 Aug;375(8):794–8.
62. Biological agents. Volume 100 B. A review of human carcinogens. IARC Monogr Eval Carcinog risks to humans. 2012;100(Pt B):1–441.
63. Parsonnet J, Friedman GD, Vandersteen DP, Chang Y, Vogelman JH, Orentreich N, et al. Helicobacter pylori infection and the risk of gastric carcinoma. N Engl J Med. 1991 Oct;325(16):1127–31.
64. Schistosomes, liver flukes and Helicobacter pylori. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Lyon, 7-14 June 1994. Vol. 61, IARC monographs on the evaluation of carcinogenic risks to humans. France; 1994. p. 1–241.
65. Cave DR. Transmission and Epidemiology of Helicobacter pylori. Am J Med [Internet]. 1996;100(96):12S-18S. Available from: https://www.sciencedirect.com/science/article/pii/S0002934396802245
66. Hu LT, Foxall PA, Russell R, Mobley HL. Purification of recombinant Helicobacter pylori urease apoenzyme encoded by ureA and ureB. Infect Immun. 1992 Jul;60(7):2657–66.
67. Watari J, Chen N, Amenta PS, Fukui H, Oshima T, Tomita T, et al. Helicobacter pylori associated chronic gastritis , clinical syndromes , precancerous lesions , and pathogenesis of gastric cancer development. 2014;20(18):5461–73.
68. Blanca Piazuelo PC. The Gastric Precancerous Cascade. J Clin Exp Pathol [Internet]. 2013;03(03):2–9. Available from: https://www.omicsonline.org/the-gastric-precancerous-cascade-2161-0681-3-147.php?aid=20275
69. Blaser MJ. Role of vacA and the cagA locus of Helicobacter pylori in human disease. Aliment Pharmacol Ther. 1996 Apr;10 Suppl 1:73–7.
70. Uemura N, Okamoto S, Yamamoto S, Matsumura N, Yamaguchi S, Yamakido M, et al.

Helicobacter pylori infection and the development of gastric cancer. N Engl J Med. 2001 Sep;345(11):784–9.
71. Hansson LE, Nyren O, Hsing AW, Bergstrom R, Josefsson S, Chow WH, et al. The risk of stomach cancer in patients with gastric or duodenal ulcer disease. N Engl J Med. 1996 Jul;335(4):242–9.
72. Fuccio L, Zagari RM, Eusebi LH, Laterza L, Cennamo V, Ceroni L, et al. Meta-analysis: can Helicobacter pylori eradication treatment reduce the risk for gastric cancer? Ann Intern Med. 2009 Jul;151(2):121–8.
73. Maehata Y, Nakamura S, Fujisawa K, Esaki M, Moriyama T, Asano K, et al. Long-term effect of Helicobacter pylori eradication on the development of metachronous gastric cancer after endoscopic resection of early gastric cancer. Gastrointest Endosc. 2012 Jan;75(1):39–46.
74. Kato M, Nishida T, Yamamoto K, Hayashi S, Kitamura S, Yabuta T, et al. Scheduled endoscopic surveillance controls secondary cancer after curative endoscopic resection for early gastric cancer: a multicentre retrospective cohort study by Osaka University ESD study group. Gut. 2013 Oct;62(10):1425–32.
75. Lee Y-C, Chiang T-H, Chou C-K, Tu Y-K, Liao W-C, Wu M-S, et al. Association Between Helicobacter pylori Eradication and Gastric Cancer Incidence: A Systematic Review and Meta-analysis. Gastroenterology. 2016 May;150(5):1113-1124.e5.
76. Leja M, Axon A, Brenner H. Epidemiology of Helicobacter pylori infection. Helicobacter. 2016;21:3–7.
77. Zamani M, Ebrahimtabar F, Zamani V, Miller WH, Alizadeh-Navaei R, Shokri-Shirvani J, et al. Systematic review with meta-analysis: the worldwide prevalence of Helicobacter pylori infection. Aliment Pharmacol Ther. 2018 Apr;47(7):868–76.
78. Kutok JL, Wang F. Spectrum of Epstein-Barr virus-associated diseases. Annu Rev Pathol. 2006;1:375–404.
79. Camargo MC, Kim W-H, Chiaravalli AM, Kim K-M, Corvalan AH, Matsuo K, et al. Improved survival of gastric cancer with tumour Epstein-Barr virus positivity: an international pooled analysis. Gut. 2014 Feb;63(2):236–43.
80. Zhao J, Jin H, Cheung KF, Tong JHM, Zhang S, Go MYY, et al. Zinc finger E-box binding factor 1 plays a central role in regulating Epstein-Barr virus (EBV) latent-lytic switch and acts as a therapeutic target in EBV-associated gastric cancer. Cancer. 2012 Feb;118(4):924–36.
81. Jing J-J, Wang Z-Y, Li H, Sun L-P, Yuan Y. Key elements involved in Epstein-Barr virus-associated gastric cancer and their network regulation. Cancer Cell Int. 2018;18:146.
82. Comprehensive molecular characterization of gastric adenocarcinoma. Nature. 2014 Sep;513(7517):202–9.
83. Sohn BH, Hwang JE, Jang HJ, Lee HS, Oh SC, Shim JJ, et al. Clinical significance of four molecular subtypes of gastric cancer identified by The Cancer Genome Atlas project. Clin Cancer Res. 2017;23(15):4441–9.
84. Dong M, Wang HY, Zhao XX, Chen JN, Zhang YW, Huang Y, et al. Expression and prognostic roles of PIK3CA, JAK2, PD-L1, and PD-L2 in Epstein-Barr virus-associated gastric carcinoma. Hum Pathol [Internet]. 2016;53:25–34. Available from: http://dx.doi.org/10.1016/j.humpath.2016.02.007
85. Ma C, Patel K, Singhi AD, Ren B, Zhu B, Shaikh F, et al. Programmed death-ligand 1 expression is common in gastric cancer associated with Epstein-Barr virus or microsatellite instability. Am J Surg Pathol. 2016;40(11):1496–506.
86. Wanebo HJ, Kennedy BJ, Chmiel J, Steele GJ, Winchester D, Osteen R. Cancer of the stomach. A patient care study by the American College of Surgeons. Ann Surg. 1993 Nov;218(5):583–92.
87. Mcfadden DW, Brenner BM. The Management of Gastric Adenocarcinoma [Internet]. Twelfth Ed. Current Surgical Therapy. Elsevier Inc.; 2018. 94–98 p. Available from:

http://dx.doi.org/10.1016/B978-0-323-37691-4.00018-5
88. Deans DAC, Tan BH, Wigmore SJ, Ross JA, de Beaux AC, Paterson-Brown S, et al. The influence of systemic inflammation, dietary intake and stage of disease on rate of weight loss in patients with gastro-oesophageal cancer. Br J Cancer. 2009 Jan;100(1):63–9.
89. Lowy AM, Mansfield PF, Leach SD, Ajani J. Laparoscopic staging for gastric cancer. Surgery. 1996 Jun;119(6):611–4.
90. Burke EC, Karpeh MS, Conlon KC, Brennan MF. Laparoscopy in the management of gastric adenocarcinoma. Ann Surg. 1997 Mar;225(3):262–7.
91. Spolverato G, Ejaz A, Kim Y, Squires MH, Poultsides GA, Fields RC, et al. Use of endoscopic ultrasound in the preoperative staging of gastric cancer: a multi-institutional study of the US gastric cancer collaborative. J Am Coll Surg. 2015 Jan;220(1):48–56.
92. Yeh JM, Hur C, Ward Z, Schrag D, Goldie SJ. Gastric adenocarcinoma screening and prevention in the era of new biomarker and endoscopic technologies: a cost-effectiveness analysis. Gut. 2016 Apr;65(4):563–74.
93. Dan YY, So JBY, Yeoh KG. Endoscopic screening for gastric cancer. Clin Gastroenterol Hepatol. 2006 Jun;4(6):709–16.
94. Kim GH, Bang SJ, Ende AR, Hwang JH. Is screening and surveillance for early detection of gastric cancer needed in Korean Americans? Korean J Intern Med. 2015 Nov;30(6):747–58.
95. In H, Solsky I, Palis B, Langdon-Embry M, Ajani J, Sano T. Validation of the 8th Edition of the AJCC TNM Staging System for Gastric Cancer using the National Cancer Database. Ann Surg Oncol. 2017;24(12):3683–91.
96. He X, Wu W, Lin Z, Ding Y, Si J, Sun L. Validation of the American Joint Committee on Cancer (AJCC) 8th edition stage system for gastric cancer patients: a population-based analysis. Gastric Cancer [Internet]. 2017;21(3):391–400. Available from: http://link.springer.com/10.1007/s10120-017-0770-1
97. Internation Agency for Research on Cancer. WHO Classification of Tumours of the Digestive System. 4th ed. 2010.
98. Okada K, Fujisaki J, Kasuga A, Omae M, Yoshimoto K, Hirasawa T, et al. Endoscopic ultrasonography is valuable for identifying early gastric cancers meeting expanded-indication criteria for endoscopic submucosal dissection. Surg Endosc. 2011 Mar;25(3):841–8.
99. Yamao T, Shirao K, Ono H, Kondo H, Saito D, Yamaguchi H, et al. Risk factors for lymph node metastasis from intramucosal gastric carcinoma. Cancer. 1996 Feb;77(4):602–6.
100. Kim JJ, Lee JH, Jung H-Y, Lee GH, Cho JY, Ryu CB, et al. EMR for early gastric cancer in Korea: a multicenter retrospective study. Gastrointest Endosc. 2007 Oct;66(4):693–700.
101. Gotoda T, Yanagisawa A, Sasako M, Ono H, Nakanishi Y, Shimoda T, et al. Incidence of lymph node metastasis from early gastric cancer: estimation with a large number of cases at two large centers. Gastric Cancer. 2000 Dec;3(4):219–25.
102. Ahn JY, Jung H-Y, Choi KD, Choi JY, Kim M-Y, Lee JH, et al. Endoscopic and oncologic outcomes after endoscopic resection for early gastric cancer: 1370 cases of absolute and extended indications. Gastrointest Endosc. 2011 Sep;74(3):485–93.
103. National Comprehensive Cancer Network. Gastric Cancer (Version 2.2018) [Internet]. NCCN Clinical Practice Guidelines in Oncology. 2018 [cited 2018 May 22]. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18267927
104. Cuschieri A, Weeden S, Fielding J, Bancewicz J, Craven J, Joypaul V, et al. Patient survival after D1 and D2 resections for gastric cancer: long-term results of the MRC randomized surgical trial. Surgical Co-operative Group. Br J Cancer. 1999 Mar;79(9–10):1522–30.

105. Songun I, Putter H, Kranenbarg EMK, Sasako M, van de Velde CJH. Surgical treatment of gastric cancer: 15-year follow-up results of the randomised nationwide Dutch D1D2 trial. Lancet Oncol. 2010 May;11(5):439–49.
106. Jiang L, Yang K-H, Guan Q-L, Zhao P, Chen Y, Tian J-H. Survival and recurrence free benefits with different lymphadenectomy for resectable gastric cancer: a meta-analysis. J Surg Oncol. 2013 Jun;107(8):807–14.
107. Japanese gastric cancer treatment guidelines 2014 (ver. 4). Gastric Cancer. 2017 Jan;20(1):1–19.
108. Macdonald JS, Smalley SR, Benedetti J, Hundahl SA, Estes NC, Stemmermann GN, et al. Chemoradiotherapy after surgery compared with surgery alone for adenocarcinoma of the stomach or gastroesophageal junction. N Engl J Med [Internet]. 2001;345(10):725–30. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11547741
109. Sasako M, Sakuramoto S, Katai H, Kinoshita T, Furukawa H, Yamaguchi T, et al. Five-year outcomes of a randomized phase III trial comparing adjuvant chemotherapy with S-1 versus surgery alone in stage II or III gastric cancer. J Clin Oncol. 2011 Nov;29(33):4387–93.
110. Cunningham D, Allum WH, Stenning SP, Thompson JN, Van de Velde CJH, Nicolson M, et al. Perioperative chemotherapy versus surgery alone for resectable gastroesophageal cancer. N Engl J Med. 2006 Jul;355(1):11–20.
111. Al-Batran SE, Hofheinz RD, Pauligk C, Kopp HG, Haag GM, Luley KB, et al. Histopathological regression after neoadjuvant docetaxel, oxaliplatin, fluorouracil, and leucovorin versus epirubicin, cisplatin, and fluorouracil or capecitabine in patients with resectable gastric or gastro-oesophageal junction adenocarcinoma (FLOT4-AIO. Lancet Oncol. 2016;17(12):1697–708.
112. Al-Batran S-E, Homann N, Pauligk C, Goetze TO, Meiler J, Kasper S, et al. Perioperative chemotherapy with fluorouracil plus leucovorin, oxaliplatin, and docetaxel versus fluorouracil or capecitabine plus cisplatin and epirubicin for locally advanced, resectable gastric or gastro-oesophageal junction adenocarcinoma (FLOT4): a r. Lancet (London, England). 2019 May;393(10184):1948–57.
113. Al-Batran S-E, Homann N, Schmalenberg H, Kopp H-G, Haag GM, Luley KB, et al. Perioperative chemotherapy with docetaxel, oxaliplatin, and fluorouracil/leucovorin (FLOT) versus epirubicin, cisplatin, and fluorouracil or capecitabine (ECF/ECX) for resectable gastric or gastroesophageal junction (GEJ) adenocarcinoma (FLOT4-AIO): A mul [Internet]. Vol. 35, Journal of Clinical Oncology. 2017 [cited 2018 Dec 16]. p. (suppl; abstr 4004). Available from: http://abstracts.asco.org/199/AbstView_199_191595.html
114. Yee N. Update in Systemic and Targeted Therapies in Gastrointestinal Oncology. Biomedicines [Internet]. 2018;6(1):34. Available from: http://www.mdpi.com/2227-9059/6/1/34
115. Kiyozumi Y, Iwatsuki M, Yamashita K, Koga Y, Yoshida N, Baba H. Update on targeted therapy and immune therapy for gastric cancer, 2018. J Cancer Metastasis Treat [Internet]. 2018;4(6):31. Available from: http://jcmtjournal.com/article/view/2638
116. Kim C, Mulder K, Spratlin J. How prognostic and predictive biomarkers are transforming our understanding and management of advanced gastric cancer. Oncologist. 2014;19(10):1046–55.
117. Song H, Zhu J, Lu D. Molecular-targeted first-line therapy for advanced gastric cancer (Review). Cochrane Database Syst Rev. 2016;(7).
118. Ohtsu A, Shah MA, Van Cutsem E, Rha SY, Sawaki A, Park SR, et al. Bevacizumab in combination with chemotherapy as first-line therapy in advanced gastric cancer: A randomized, double-blind, placebo-controlled phase III study. J Clin Oncol. 2011;29(30):3968–76.
119. Lordick F, Kang Y-K, Chung H-C, Salman P, Oh SC, Bodoky G, et al. Capecitabine and

cisplatin with or without cetuximab for patients with previously untreated advanced gastric cancer (EXPAND): a randomised, open-label phase 3 trial. Lancet Oncol. 2013 May;14(6):490–9.
120. Ohtsu A, Ajani JA, Bai Y-X, Bang Y-J, Chung H-C, Pan H-M, et al. Everolimus for previously treated advanced gastric cancer: results of the randomized, double-blind, phase III GRANITE-1 study. J Clin Oncol. 2013 Nov;31(31):3935–43.
121. Bang YJ, Van Cutsem E, Feyereislova A, Chung HC, Shen L, Sawaki A, et al. Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive advanced gastric or gastro-oesophageal junction cancer (ToGA): A phase 3, open-label, randomised controlled trial. Lancet [Internet]. 2010;376(9742):687–97. Available from: http://dx.doi.org/10.1016/S0140-6736(10)61121-X
122. Hecht JR, Bang YJ, Qin SK, Chung HC, Xu JM, Park JO, et al. Lapatinib in combination with capecitabine plus oxaliplatin in human epidermal growth factor receptor 2-positive advanced or metastatic gastric, esophageal, or gastroesophageal adenocarcinoma: TRIO-013/LOGiC - A randomized phase III trial. J Clin Oncol. 2016;34(5):443–51.
123. Wilke H, Muro K, Van Cutsem E, Oh S-C, Bodoky G, Shimada Y, et al. Ramucirumab plus paclitaxel versus placebo plus paclitaxel in patients with previously treated advanced gastric or gastro-oesophageal junction adenocarcinoma (RAINBOW): a double-blind, randomised phase 3 trial. Lancet Oncol. 2014 Oct;15(11):1224–35.
124. Fuchs CS, Tomasek J, Yong CJ, Dumitru F, Passalacqua R, Goswami C, et al. Ramucirumab monotherapy for previously treated advanced gastric or gastro-oesophageal junction adenocarcinoma (REGARD): An international, randomised, multicentre, placebo-controlled, phase 3 trial. Lancet. 2014;383(9911):31–9.
125. Matsueda S, Graham DY. Immunotherapy in gastric cancer. World J Gastroenterol. 2014;20(7):1657–66.
126. Muro K, Chung HC, Shankaran V, Geva R, Catenacci D, Gupta S, et al. Pembrolizumab for patients with PD-L1-positive advanced gastric cancer (KEYNOTE-012): a multicentre, open-label, phase 1b trial. Lancet Oncol. 2016 Jun;17(6):717–26.
127. Fuchs CS, Doi T, Jang RW, Muro K, Satoh T, Machado M, et al. Safety and efficacy of pembrolizumab monotherapy in patients with previously treated advanced gastric and gastroesophageal junction cancer: Phase 2 clinical KEYNOTE-059 trial. JAMA Oncol. 2018;4(5):1–8.
128. Janjigian YY, Bendell J, Calvo E, Kim JW, Ascierto PA, Sharma P, et al. CheckMate-032 Study: Efficacy and Safety of Nivolumab and Nivolumab Plus Ipilimumab in Patients With Metastatic Esophagogastric Cancer. J Clin Oncol. 2018 Oct;36(28):2836–44.
129. Shitara K, Van Cutsem E, Bang YJ, Fuchs C, Wyrwicz L, Lee KW, et al. Efficacy and Safety of Pembrolizumab or Pembrolizumab plus Chemotherapy vs Chemotherapy Alone for Patients with First-line, Advanced Gastric Cancer: The KEYNOTE-062 Phase 3 Randomized Clinical Trial. JAMA Oncol. 2020;6(10):1571–80.
130. Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, et al. PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. N Engl J Med [Internet]. 2015;372(26):2509–20. Available from: http://www.nejm.org/doi/10.1056/NEJMoa1500596
131. Le DT, Durham JN, Smith KN, Wang H, Bartlett BR, Aulakh LK, et al. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. Science (80-) [Internet]. 2017;357(6349):409–13. Available from: http://science.sciencemag.org/content/sci/357/6349/409.full.pdf%0Ahttp://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed18&NEWS=N&AN=617529777
132. Atkinson AJ, Colburn WA, DeGruttola VG, DeMets DL, Downing GJ, Hoth DF, et al. Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. Clin Pharmacol Ther. 2001;69(3):89–95.

133. Strimbu K, Tavel JA. What are biomarkers? Curr Opin HIV AIDS. 2010;5(6):463–6.
134. Carlomagno N, Incollingo P, Tammaro V, Peluso G, Rupealta N, Chiacchio G, et al. Diagnostic, Predictive, Prognostic, and Therapeutic Molecular Biomarkers in Third Millennium: A Breakthrough in Gastric Cancer. Biomed Res Int. 2017;2017.
135. Durães C, Almeida GM, Seruca R, Oliveira C, Carneiro F. Biomarkers for gastric cancer: Prognostic, predictive or targets of therapy? Virchows Arch. 2014;464(3):367–78.
136. Bastien RR, Rodríguez-Lescure Á, Ebbert MT, Prat A, Munárriz B, Rowe L, et al. PAM50 breast cancer subtyping by RT-qPCR and concordance with standard clinical molecular markers. BMC Med Genomics. 2012;5.
137. Abrahao-Machado LF, Scapulatempo-Neto C. HER2 testing in gastric cancer: An update. World J Gastroenterol [Internet]. 2016 May 21;22(19):4619–25. Available from: https://pubmed.ncbi.nlm.nih.gov/27217694
138. Ratti M, Lampis A, Hahne JC, Passalacqua R, Valeri N. Microsatellite instability in gastric cancer: molecular bases, clinical perspectives, and new treatment approaches. Cell Mol Life Sci [Internet]. 2018/09/01. 2018 Nov;75(22):4151–62. Available from: https://pubmed.ncbi.nlm.nih.gov/30173350
139. Jiang Y, Zhang Q, Hu Y, Li T, Yu J, Zhao L, et al. ImmunoScore Signature: A Prognostic and Predictive Tool in Gastric Cancer. Ann Surg. 2018 Mar;267(3):504–13.
140. Mlecnik B, Bindea G, Angell HK, Maby P, Angelova M, Tougeron D, et al. Integrative Analyses of Colorectal Cancer Show Immunoscore Is a Stronger Predictor of Patient Survival Than Microsatellite Instability. Immunity. 2016 Mar;44(3):698–711.
141. Angell HK, Bruni D, Barrett JC, Herbst R, Galon J. The Immunoscore: Colon Cancer and Beyond. Clin Cancer Res. 2020 Jan;26(2):332–9.
142. Hendry S, Salgado R, Gevaert T, Russell PA, John T, Thapa B, et al. Assessing Tumor-Infiltrating Lymphocytes in Solid Tumors: A Practical Review for Pathologists and Proposal for a Standardized Method from the International Immuno-Oncology Biomarkers Working Group: Part 2: TILs in Melanoma, Gastrointestinal Tract Carcinom. Adv Anat Pathol. 2017;24(6):311–35.
143. Galon J. Approaches to treat immune hot , altered and cold tumours with combination immunotherapies. Nat Rev Drug Discov [Internet]. 2019;18(March):10–3. Available from: http://dx.doi.org/10.1038/s41573-018-0007-y
144. Khalil DN, Smith EL, Brentjens RJ, Wolchok JD. The future of cancer treatment: Immunomodulation, CARs and combination immunotherapy. Nat Rev Clin Oncol. 2016;13(5):273–90.
145. Humar B, Graziano F, Cascinu S, Catalano V, Ruzzo AM, Magnani M, et al. Association of CDH1 haplotypes with susceptibility to sporadic diffuse gastric cancer. Oncogene. 2002 Nov;21(53):8192–5.
146. Tamura G, Yin J, Wang S, Fleisher AS, Zou T, Abraham JM, et al. E-Cadherin gene promoter hypermethylation in primary human gastric carcinomas. J Natl Cancer Inst. 2000 Apr;92(7):569–73.
147. Corso G, Carvalho J, Marrelli D, Vindigni C, Carvalho B, Seruca R, et al. Somatic mutations and deletions of the E-cadherin gene predict poor survival of patients with gastric cancer. J Clin Oncol. 2013 Mar;31(7):868–75.
148. Agnes A, Estrella JS, Badgwell B. The significance of a nineteenth century definition in the era of genomics: linitis plastica. World J Surg Oncol. 2017 Jul;15(1):123.
149. Kadrofske MM, Openo KP, Wang JL. The human LGALS3 (galectin-3) gene: determination of the gene structure and functional characterization of the promoter. Arch Biochem Biophys. 1998 Jan;349(1):7–20.
150. Cherayil BJ, Weiner SJ, Pillai S. The Mac-2 antigen is a galactose-specific lectin that binds IgE. J Exp Med. 1989 Dec;170(6):1959–72.
151. Sciacchitano S, Lavra L, Morgante A, Ulivieri A, Magi F, De Francesco GP, et al. Galectin-3: One molecule for an alphabet of diseases, from A to Z. Int J Mol Sci.

2018;19(2).
152. Fortuna-Costa A, Gomes AM, Kozlowski EO, Stelling MP, Pavão MSG. Extracellular galectin-3 in tumor progression and metastasis. Front Oncol. 2014;4:138.
153. Thijssen VL, Heusschen R, Caers J, Griffioen AW. Galectin expression in cancer diagnosis and prognosis: A systematic review. Biochim Biophys Acta. 2015 Apr;1855(2):235–47.
154. Haudek KC, Spronk KJ, Voss PG, Patterson RJ, Wang JL, Arnoys EJ. Dynamics of galectin-3 in the nucleus and cytoplasm. Biochim Biophys Acta. 2010 Feb;1800(2):181–9.
155. Yang R-Y, Rabinovich GA, Liu F-T. Galectins: structure, function and therapeutic potential. Expert Rev Mol Med. 2008 Jun;10:e17.
156. Tas F, Bilgin E, Tastekin D, Erturk K, Duranyildiz D. Clinical Significance of Serum Galectin-3 Levels in Gastric Cancer Patients. J Gastrointest Cancer. 2016 Jun;47(2):182–6.
157. Okada K, Shimura T, Suehiro T, Mochiki E, Kuwano H. Reduced galectin-3 expression is an indicator of unfavorable prognosis in gastric cancer. Anticancer Res. 2006;26(2B):1369–76.
158. Stasenko M, Smith E, Yeku O, Park KJ, Laster I, Lee K, et al. Targeting galectin-3 with a high-affinity antibody for inhibition of high-grade serous ovarian cancer and other MUC16/CA-125-expressing malignancies. Sci Rep [Internet]. 2021;11(1):3718. Available from: https://doi.org/10.1038/s41598-021-82686-3
159. Harrison SA, Marri SR, Chalasani N, Kohli R, Aronstein W, Thompson GA, et al. Randomised clinical study: GR-MD-02, a galectin-3 inhibitor, vs. placebo in patients having non-alcoholic steatohepatitis with advanced fibrosis. Aliment Pharmacol Ther. 2016 Dec;44(11–12):1183–98.
160. Frigeri LG, Liu FT. Surface expression of functional IgE binding protein, an endogenous lectin, on mast cells and macrophages. J Immunol [Internet]. 1992 Feb 1;148(3):861 LP – 867. Available from: http://www.jimmunol.org/content/148/3/861.abstract
161. Jia W, Kidoya H, Yamakawa D, Naito H, Takakura N. Galectin-3 Accelerates M2 Macrophage Infiltration and Angiogenesis in Tumors. Am J Pathol [Internet]. 2013;182(5):1821–31. Available from: https://www.sciencedirect.com/science/article/pii/S0002944013001016
162. MacKinnon AC, Farnworth SL, Hodkinson PS, Henderson NC, Atkinson KM, Leffler H, et al. Regulation of Alternative Macrophage Activation by Galectin-3. J Immunol [Internet]. 2008 Feb 15;180(4):2650 LP – 2658. Available from: http://www.jimmunol.org/content/180/4/2650.abstract
163. Liu J, Geng X, Hou J, Wu G. New insights into M1/M2 macrophages: key modulators in cancer progression. Cancer Cell Int [Internet]. 2021;21(1):1–7. Available from: https://doi.org/10.1186/s12935-021-02089-2
164. Turley SJ, Cremasco V, Astarita JL. Immunological hallmarks of stromal cells in the tumour microenvironment. Nat Rev Immunol [Internet]. 2015;15(11):669–82. Available from: https://doi.org/10.1038/nri3902
165. Long B, Yu Z, Zhou H, Ma Z, Ren Y, Zhan H, et al. Clinical characteristics and prognostic significance of galectins for patients with gastric cancer: A meta-analysis. Int J Surg [Internet]. 2018 Aug;56(82):242–9. Available from: https://doi.org/10.1016/j.ijsu.2018.06.033
166. Li F, Li C, Cai X, Xie Z, Zhou L, Cheng B, et al. The association between CD8+ tumor-infiltrating lymphocytes and the clinical outcome of cancer immunotherapy: A systematic review and meta-analysis. eClinicalMedicine. 2021;41.
167. Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. Cell [Internet]. 2011 Mar 4;144(5):646–74. Available from: https://doi.org/10.1016/j.cell.2011.02.013

168. Schreiber RD, Old LJ, Smyth MJ. Cancer immunoediting: Integrating immunity's roles in cancer suppression and promotion. Science (80-). 2011;331(6024):1565–70.
169. Jiang Z, Liu Z, Li M, Chen C, Wang X. Immunogenomics Analysis Reveals that TP53 Mutations Inhibit Tumor Immunity in Gastric Cancer. Transl Oncol [Internet]. 2018;11(5):1171–87. Available from: https://doi.org/10.1016/j.tranon.2018.07.012
170. Wormann SM, Song L, Ai J, Diakopoulos KN, Kurkowski MU, Gorgulu K, et al. Loss of P53 Function Activates JAK2-STAT3 Signaling to Promote Pancreatic Tumor Growth, Stroma Modification, and Gemcitabine Resistance in Mice and Is Associated With Patient Survival. Gastroenterology. 2016 Jul;151(1):180-193.e12.
171. Galluzzi L, Chan TA, Kroemer G, Wolchok JD, López-Soto A. The hallmarks of successful anticancer immunotherapy. Sci Transl Med. 2018;10(459):1–15.
172. Jensen C, Teng Y. Is It Time to Start Transitioning From 2D to 3D Cell Culture? [Internet]. Vol. 7, Frontiers in Molecular Biosciences . 2020. Available from: https://www.frontiersin.org/article/10.3389/fmolb.2020.00033
173. Costa EC, Moreira AF, de Melo-Diogo D, Gaspar VM, Carvalho MP, Correia IJ. 3D tumor spheroids: an overview on the tools and techniques used for their analysis. Biotechnol Adv. 2016 Dec;34(8):1427–41.
174. Ooft SN, Weeber F, Schipper L, Dijkstra KK, McLean CM, Kaing S, et al. Prospective experimental treatment of colorectal cancer patients based on organoid drug responses. ESMO Open [Internet]. 2021;6(3):100103. Available from: https://doi.org/10.1016/j.esmoop.2021.100103
175. Zhang F, Wang W, Long Y, Liu H, Cheng J, Guo L, et al. Characterization of drug responses of mini patient-derived xenografts in mice for predicting cancer patient clinical therapeutic response. Cancer Commun (London, England) [Internet]. 2018;38(1):60. Available from: https://doi.org/10.1186/s40880-018-0329-5
176. Kim J, Koo BK, Knoblich JA. Human organoids: model systems for human biology and medicine. Nat Rev Mol Cell Biol [Internet]. 2020;21(10):571–84. Available from: http://dx.doi.org/10.1038/s41580-020-0259-3
177. Method of the Year 2017: Organoids. Nat Methods. 2018;15(1):1–1.
178. Gao M, Lin M, Rao M, Thompson H, Hirai K, Choi M, et al. Development of Patient-Derived Gastric Cancer Organoids from Endoscopic Biopsies and Surgical Tissues. Ann Surg Oncol [Internet]. 2018;25(9):2767–75. Available from: https://doi.org/10.1245/s10434-018-6662-8
179. Bartfeld S, Bayram T, Van De Wetering M, Huch M, Begthel H, Kujala P, et al. In vitro expansion of human gastric epithelial stem cells and their responses to bacterial infection. Gastroenterology. 2015;148(1):126-136.e6.
180. Steele NG, Chakrabarti J, Wang J, Biesiada J, Holokai L, Chang J, et al. An Organoid-Based Preclinical Model of Human Gastric Cancer. Cmgh [Internet]. 2019;7(1):161–84. Available from: https://doi.org/10.1016/j.jcmgh.2018.09.008
181. Weeber F, Ooft SN, Dijkstra KK, Voest EE. Tumor Organoids as a Pre-clinical Cancer Model for Drug Discovery. Cell Chem Biol [Internet]. 2017;24(9):1092–100. Available from: http://dx.doi.org/10.1016/j.chembiol.2017.06.012
182. Yan HHN, Siu HC, Law S, Ho SL, Yue SSK, Tsui WY, et al. A Comprehensive Human Gastric Cancer Organoid Biobank Captures Tumor Subtype Heterogeneity and Enables Therapeutic Screening. Cell Stem Cell [Internet]. 2018;23(6):882-897.e11. Available from: https://doi.org/10.1016/j.stem.2018.09.016
183. Van De Wetering M, Francies HE, Francis JM, Bounova G, Iorio F, Pronk A, et al. Prospective derivation of a living organoid biobank of colorectal cancer patients. Cell [Internet]. 2015;161(4):933–45. Available from: http://dx.doi.org/10.1016/j.cell.2015.03.053
184. Seidlitz T, Merker SR, Rothe A, Zakrzewski F, Von Neubeck C, Grützmann K, et al. Human gastric cancer modelling using organoids. Gut. 2019;68(2):207–17.
185. Sato T, Vries RG, Snippert HJ, Van De Wetering M, Barker N, Stange DE, et al. Single

Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. Nature. 2009;459(7244):262–5.
186. Sato T, Stange DE, Ferrante M, Vries RGJ, Van Es JH, Van den Brink S, et al. Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. Gastroenterology. 2011 Nov;141(5):1762–72.
187. Zhao Y. Stem cells in gastric cancer. World J Gastroenterol [Internet]. 2015;21(1):112. Available from: http://www.wjgnet.com/1007-9327/full/v21/i1/112.htm
188. Stange DE, Koo BK, Huch M, Sibbel G, Basak O, Lyubimova A, et al. Differentiated Troy+ chief cells act as reserve stem cells to generate all lineages of the stomach epithelium. Cell [Internet]. 2013;155(2):357. Available from: http://dx.doi.org/10.1016/j.cell.2013.09.008
189. Neal JT, Li X, Zhu J, Giangarra V, Grzeskowiak CL, Ju J, et al. Organoid Modeling of the Tumor Immune Microenvironment. Cell [Internet]. 2018;175(7):1972-1988.e16. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30550791
190. Pietras K, Ostman A. Hallmarks of cancer: interactions with the tumor stroma. Exp Cell Res. 2010 May;316(8):1324–31.
191. Shimoda M, Mellody KT, Orimo A. Carcinoma-associated fibroblasts are a rate-limiting determinant for tumour progression. Semin Cell Dev Biol. 2010 Feb;21(1):19–25.
192. Lin M, Kim J. ASO Author Reflections: We Can Stomach It: Personalizing Medicine with EGD-Derived Gastric Cancer Organoids. Ann Surg Oncol [Internet]. 2018;25(S3):972–3. Available from: https://doi.org/10.1245/s10434-018-6962-z
193. Aberle MR, Burkhart RA, Tiriac H, Olde Damink SWM, Dejong CHC, Tuveson DA, et al. Patient-derived organoid models help define personalized management of gastrointestinal cancer. Br J Surg. 2018;105(2):e48–60.
194. Wallaschek N, Niklas C, Pompaiah M, Wiegering A, Germer CT, Kircher S, et al. Establishing Pure Cancer Organoid Cultures: Identification, Selection and Verification of Cancer Phenotypes and Genotypes. J Mol Biol [Internet]. 2019;431(15):2884–93. Available from: https://doi.org/10.1016/j.jmb.2019.05.031
195. Tiriac H, Belleau P, Engle DD, Plenker D, Deschênes A, Somerville TDD, et al. Organoid profiling identifies common responders to chemotherapy in pancreatic cancer. Cancer Discov. 2018;8(9):1112–29.
196. Vlachogiannis G, Hedayat S, Vatsiou A, Jamin Y, Fernández-Mateos J, Khan K, et al. Patient-derived organoids model treatment response of metastatic gastrointestinal cancers. Science (80-). 2018;359(6378):920–6.
197. Koedijk MS, Heijmen BJM, Groot Koerkamp B, Eskens FALM, Sprengers D, Poley JW, et al. Protocol for the STRONG trial: Stereotactic body radiation therapy following chemotherapy for unresectable perihilar cholangiocarcinoma, a phase i feasibility study. BMJ Open. 2018;8(10):1–8.
198. Ahmed Z. Practicing precision medicine with intelligently integrative clinical and multi-omics data analysis. Hum Genomics. 2020;14(1):1–5.
199. Malta TM, Sokolov A, Gentles AJ, Burzykowski T, Poisson L, Weinstein JN, et al. Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation. Cell [Internet]. 2018 Apr 5;173(2):338-354.e15. Available from: https://doi.org/10.1016/j.cell.2018.03.034
200. Bailey MH, Tokheim C, Porta-Pardo E, Sengupta S, Bertrand D, Weerasinghe A, et al. Comprehensive Characterization of Cancer Driver Genes and Mutations. Cell [Internet]. 2018 Apr 5;173(2):371-385.e18. Available from: https://doi.org/10.1016/j.cell.2018.02.060
201. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, et al. The Immune Landscape of Cancer. Immunity. 2018;48(4):812-830.e14.
202. Sanchez-Vega F, Mina M, Armenia J, Chatila WK, Luna A, La KC, et al. Oncogenic

Signaling Pathways in The Cancer Genome Atlas. Cell [Internet]. 2018 Apr 5;173(2):321-337.e10. Available from: https://doi.org/10.1016/j.cell.2018.03.035
203. Liu Y, Sethi NS, Hinoue T, Schneider BG, Cherniack AD, Sanchez-Vega F, et al. Comparative Molecular Analysis of Gastrointestinal Adenocarcinomas. Cancer Cell. 2018;33(4):721-735.e8.
204. Lafarge MW, Koelzer VH. Towards computationally efficient prediction of molecular signatures from routine histology images. Lancet Digit Heal [Internet]. 2021;3(12):e752–3. Available from: http://dx.doi.org/10.1016/S2589-7500(21)00232-6
205. Heo YJ, Park C, Yu D, Lee J, Kim K-M. Reproduction of molecular subtypes of gastric adenocarcinoma by transcriptome sequencing of archival tissue. Sci Rep [Internet]. 2019;(June):1–8. Available from: http://feeds.nature.com/~r/srep/rss/current/~3/zK1oFzoS8yM/s41598-019-46216-6?utm_source=researcher_app&utm_medium=referral&utm_campaign=MKEF_USG_Researcher_inbound
206. Kim ST, Cristescu R, Bass AJ, Kim KM, Odegaard JI, Kim K, et al. Comprehensive molecular characterization of clinical responses to PD-1 inhibition in metastatic gastric cancer. Nat Med [Internet]. 2018;24(9):1449–58. Available from: http://dx.doi.org/10.1038/s41591-018-0101-z
207. Zeng D, Li M, Zhou R, Zhang J, Sun H, Shi M, et al. Tumor microenvironment characterization in gastric cancer identifies prognostic and immunotherapeutically relevant gene signatures. Cancer Immunol Res. 2019;7(5):737–50.
208. Oh SC, Sohn BH, Cheong JH, Kim SB, Lee JE, Park KC, et al. Clinical and genomic landscape of gastric cancer with a mesenchymal phenotype. Nat Commun [Internet]. 2018;9(1). Available from: http://dx.doi.org/10.1038/s41467-018-04179-8
209. Li L, Wang X. Identification of gastric cancer subtypes based on pathway clustering. npj Precis Oncol [Internet]. 2021;5(1):1–17. Available from: http://dx.doi.org/10.1038/s41698-021-00186-z
210. Kumar V, Ramnarayanan K, Sundar R, Padmanabhan N, Srivastava S, Koiwa M, et al. Single-Cell Atlas of Lineage States, Tumor Microenvironment, and Subtype-Specific Expression Programs in Gastric Cancer. Cancer Discov. 2022 Mar;12(3):670–91.
211. Gabbert HE, Mueller W, Schneiders A, Meier S, Moll R, Birchmeier W, et al. Prognostic value of E-cadherin expression in 413 gastric carcinomas. Int J cancer. 1996 Jun;69(3):184–9.
212. Misra S, Pedroso FE, DiPasco PJ, Solomon NL, Gennis E, Franceschi D, et al. Does neoadjuvant chemotherapy improve outcomes for patients with gastric cancer? J Surg Res [Internet]. 2012;178(2):623–31. Available from: https://www.sciencedirect.com/science/article/pii/S0022480412004222
213. Xu A-M, Huang L, Liu W, Gao S, Han W-X, Wei Z-J. Neoadjuvant Chemotherapy Followed by Surgery versus Surgery Alone for Gastric Carcinoma: Systematic Review and Meta-Analysis of Randomized Controlled Trials. PLoS One [Internet]. 2014 Jan 30;9(1):e86941. Available from: https://doi.org/10.1371/journal.pone.0086941
214. Ruvolo PP. Galectin 3 as a guardian of the tumor microenvironment. Biochim Biophys Acta - Mol Cell Res [Internet]. 2016;1863(3):427–37. Available from: https://www.sciencedirect.com/science/article/pii/S0167488915002700
215. Fukumori T, Kanayama H-O, Raz A. The role of galectin-3 in cancer drug resistance. Drug Resist Updat [Internet]. 2007/06/04. 2007 Jun;10(3):101–8. Available from: https://pubmed.ncbi.nlm.nih.gov/17544840
216. Ilmer M, Mazurek N, Byrd JC, Ramirez K, Hafley M, Alt E, et al. Cell surface galectin-3 defines a subset of chemoresistant gastrointestinal tumor-initiating cancer cells with heightened stem cell characteristics. Cell Death Dis [Internet]. 2016;7(8):e2337–e2337. Available from: https://doi.org/10.1038/cddis.2016.239
217. Wang L, Li Z, Wang C, Yang Y, Sun L, Yao W, et al. E-cadherin decreased human

breast cancer cells sensitivity to staurosporine by up-regulating Bcl-2 expression. Arch Biochem Biophys. 2009 Jan;481(1):116–22.
218. Berezhnaya NM, Belova OB, Vinnichuk YD, Tarutinov VI. Expression of E-cadherin in drug resistant human breast cancer cells and their sensitivity to lymphokine-activated lymphocytes action. Exp Oncol. 2009 Dec;31(4):242–5.
219. Galon J, Costes A, Sanchez-Cabo F, Kirilovsky A, Mlecnik B, Lagorce-Pagès C, et al. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. Science. 2006 Sep;313(5795):1960–4.
220. Zou W, Zhou M-L, Zhang L-Y, Yang J-N, Yang W, Wang Y-Q, et al. Immune Score Predicts Outcomes of Gastric Cancer Patients Treated with Adjuvant Chemoradiotherapy. J Oncol [Internet]. 2021 Dec 27;2021:9344124. Available from: https://pubmed.ncbi.nlm.nih.gov/34987582
221. Lee HE, Chae SW, Lee YJ, Kim MA, Lee HS, Lee BL, et al. Prognostic implications of type and density of tumour-infiltrating lymphocytes in gastric cancer. Br J Cancer. 2008 Nov;99(10):1704–11.
222. Liu K, Yang K, Wu B, Chen H, Chen X, Chen X, et al. Tumor-Infiltrating Immune Cells Are Associated With Prognosis of Gastric Cancer. Medicine (Baltimore). 2015 Sep;94(39):e1631.
223. Derieux S, Svrcek M, Manela S, Lagorce-Pages C, Berger A, André T, et al. Evaluation of the prognostic impact of pathologic response to preoperative chemotherapy using Mandard's Tumor Regression Grade (TRG) in gastric adenocarcinoma. Dig Liver Dis [Internet]. 2019; Available from: https://linkinghub.elsevier.com/retrieve/pii/S1590865819307121
224. Bankhead P, Loughrey MB, Fernández JA, Dombrowski Y, McArt DG, Dunne PD, et al. QuPath: Open source software for digital pathology image analysis. Sci Rep. 2017;7(1):1–7.
225. McClelland RA, Finlay P, Walker KJ, Nicholson D, Robertson JF, Blamey RW, et al. Automated quantitation of immunocytochemically localized estrogen receptors in human breast cancer. Cancer Res. 1990 Jun;50(12):3545–50.
226. R Core Team. R: A language and environment for statistical computing. R Found Stat Comput Vienna, Austria [Internet]. 2020; Available from: https://www.r-project.org/
227. Kuhn M. Building predictive models in R using the caret package. J Stat Softw. 2008;28(5):1–26.
228. Friedman JH, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. J Stat Softw [Internet]. 2010 Feb 2;33(1 SE-Articles):1–22. Available from: https://www.jstatsoft.org/index.php/jss/article/view/v033i01
229. Ychou M, Boige V, Pignon J-P, Conroy T, Bouché O, Lebreton G, et al. Perioperative chemotherapy compared with surgery alone for resectable gastroesophageal adenocarcinoma: an FNCLCC and FFCD multicenter phase III trial. J Clin Oncol Off J Am Soc Clin Oncol. 2011 May;29(13):1715–21.
230. Schuhmacher C, Gretschel S, Lordick F, Reichardt P, Hohenberger W, Eisenberger CF, et al. Neoadjuvant chemotherapy compared with surgery alone for locally advanced cancer of the stomach and cardia: European Organisation for Research and Treatment of Cancer randomized trial 40954. J Clin Oncol Off J Am Soc Clin Oncol. 2010 Dec;28(35):5210–8.
231. Kaltenmeier C, Althans A, Mascara M, Nassour I, Khan S, Hoehn R, et al. Pathologic Complete Response Following Neoadjuvant Therapy for Gastric Adenocarcinoma: A National Cancer Database Analysis on Incidence, Predictors, and Outcomes. Am Surg. 2021 Jul;87(7):1145–54.
232. Li Z, Shan F, Wang Y, Zhang Y, Zhang L, Li S, et al. Correlation of pathological complete response with survival after neoadjuvant chemotherapy in gastric or gastroesophageal junction cancer treated with radical surgery: A meta-analysis. PLoS

One. 2018;13(1):1–13.
233. Jiang L, Ma Z, Ye X, Kang W, Yu J. Clinicopathological factors affecting the effect of neoadjuvant chemotherapy in patients with gastric cancer. World J Surg Oncol [Internet]. 2021;19(1):44. Available from: https://doi.org/10.1186/s12957-021-02157-x
234. Wang LB, Teng RY, Jiang ZN, Hu WX, Dong MJ, Yuan XM, et al. Clinicopathologic variables predicting tumor response to neoadjuvant chemotherapy in patients with locally advanced gastric cancer. J Surg Oncol. 2012 Mar;105(3):293–6.
235. Chen Y-H, Xiao J, Chen X-J, Wang H-S, Liu D, Xiang J, et al. Nomogram for predicting pathological complete response to neoadjuvant chemotherapy in patients with advanced gastric cancer. World J Gastroenterol [Internet]. 2020 May 21;26(19):2427–39. Available from: https://pubmed.ncbi.nlm.nih.gov/32476803
236. Oshima T, Yoshikawa T, Miyagi Y, Morita S, Yamamoto M, Tanabe K, et al. Biomarker analysis to predict the pathological response to neoadjuvant chemotherapy in locally advanced gastric cancer: An exploratory biomarker study of COMPASS, a randomized phase II trial. Oncotarget; Vol 11, No 30 [Internet]. 2020; Available from: https://www.oncotarget.com/article/27658/text/
237. Sun J, Wang X, Zhang Z, Zeng Z, Ouyang S, Kang W. The Sensitivity Prediction of Neoadjuvant Chemotherapy for Gastric Cancer. Front Oncol. 2021;11:641304.
238. Wang K, Shen T, Siegal GP, Wei S. The CD4/CD8 ratio of tumor-infiltrating lymphocytes at the tumor-host interface has prognostic value in triple-negative breast cancer. Hum Pathol. 2017 Nov;69:110–7.
239. Ganesh K, Wu C, O'Rourke KP, Szeglin BC, Zheng Y, Sauvé CEG, et al. A rectal cancer organoid platform to study individual responses to chemoradiation. Nat Med [Internet]. 2019;25(10):1607–14. Available from: http://dx.doi.org/10.1038/s41591-019-0584-2
240. Foo MA, You M, Chan SL, Sethi G, Bonney GK, Yong WP, et al. Clinical translation of patient-derived tumour organoids- bottlenecks and strategies. Biomark Res. 2022;10(1).
241. Fujii M, Shimokawa M, Date S, Takano A, Matano M, Nanki K, et al. A Colorectal Tumor Organoid Library Demonstrates Progressive Loss of Niche Factor Requirements during Tumorigenesis. Cell Stem Cell. 2016 Jun;18(6):827–38.
242. Puca L, Bareja R, Prandi D, Shaw R, Benelli M, Karthaus WR, et al. Patient derived organoids to model rare prostate cancer phenotypes. Nat Commun [Internet]. 2018;9(1):2404. Available from: https://doi.org/10.1038/s41467-018-04495-z
243. Mowat C, Mosley SR, Namdar A, Schiller D, Baker K. Anti-tumor immunity in mismatch repair-deficient colorectal cancers requires type I IFN–driven CCL5 and CXCL10. J Exp Med. 2021;218(9).
244. Miyoshi H, Stappenbeck TS. In vitro expansion and genetic modification of gastrointestinal stem cells in spheroid culture. Nat Protoc. 2013 Dec;8(12):2471–82.
245. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods [Internet]. 2012;9(7):671–5. Available from: https://doi.org/10.1038/nmeth.2089
246. Applied Biosystems. Guide to Performing Relative Quantitation of Gene Expression Using Real-Time Quantitative PCR. Appl Biosyst [Internet]. 2008;2009:1–60. Available from: http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Guide+to+Performing+Relative+Quantitation+of+Gene+Expression+Using+Real-Time+Quantitative+PCR#0
247. Wood SN. Stable and Efficient Multiple Smoothing Parameter Estimation for Generalized Additive Models. J Am Stat Assoc [Internet]. 2004 Sep 1;99(467):673–86. Available from: https://doi.org/10.1198/016214504000000980
248. CORFIELD AP, MYERSCOUGH N, LONGMAN R, SYLVESTER P, ARUL S, PIGNATELLI M.

Mucins and mucosal protection in the gastrointestinal tract: new prospects for mucins in the pathology of gastrointestinal disease. Gut [Internet]. 2000 Oct 1;47(4):589 LP – 594. Available from: http://gut.bmj.com/content/47/4/589.abstract
249. Yoon J-Y, Brezden-Masley C, Streutker CJ. Lgr5 and stem/progenitor gene expression in gastric/gastroesophageal junction carcinoma – significance of potentially retained stemness. BMC Cancer [Internet]. 2020;20(1):860. Available from: https://doi.org/10.1186/s12885-020-07362-7
250. Evelina S, Wen Z, Linn F, Max K, Nicholas M, Csaba A, et al. An atlas of the protein-coding genes in the human, pig, and mouse brain. Science (80-) [Internet]. 2020 Mar 6;367(6482):eaay5947. Available from: https://doi.org/10.1126/science.aay5947
251. Pauli C, Hopkins BD, Prandi D, Shaw R, Fedrizzi T, Sboner A, et al. Personalized in vitro and in vivo cancer models to guide precision medicine. Cancer Discov. 2017;7(5):462–77.
252. Walsh AJ, Cook RS, Sanders ME, Arteaga CL, Skala MC. Drug response in organoids generated from frozen primary tumor tissues. Sci Rep [Internet]. 2016;6(November 2015):1–12. Available from: http://dx.doi.org/10.1038/srep18889
253. Georgiou M, Chichagova V, Hilgen G, Dorgau B, Sernagor E, Armstrong L, et al. Room temperature shipment does not affect the biological activity of pluripotent stem cell-derived retinal organoids. PLoS One [Internet]. 2020 Jun 1;15(6):e0233860. Available from: https://doi.org/10.1371/journal.pone.0233860
254. Singh RK, Winkler P, Binette F, Glickman RD, Seiler M, Petersen-Jones SM, et al. Development of a protocol for maintaining viability while shipping organoid-derived retinal tissue. J Tissue Eng Regen Med. 2020 Feb;14(2):388–94.
255. Xu L, Lin W, Wen L, Li G. Lgr5 in cancer biology: functional identification of Lgr5 in cancer progression and potential opportunities for novel therapy. Stem Cell Res Ther [Internet]. 2019;10(1):219. Available from: https://doi.org/10.1186/s13287-019-1288-8
256. Hata M, Hayakawa Y, Koike K. Gastric stem cell and cellular origin of cancer. Biomedicines. 2018;6(4):1–9.
257. XI H-Q, CUI J-X, SHEN W-S, WU X-S, BIAN S-B, LI J-Y, et al. Increased expression of Lgr5 is associated with chemotherapy resistance in human gastric cancer. Oncol Rep [Internet]. 2014;32(1):181–8. Available from: https://www.spandidos-publications.com/
258. Pellino A, Riello E, Nappo F, Brignola S, Murgioni S, Djaballah SA, et al. Targeted therapies in metastatic gastric cancer: Current knowledge and future perspectives. World J Gastroenterol [Internet]. 2019 Oct 14;25(38):5773–88. Available from: https://pubmed.ncbi.nlm.nih.gov/31636471
259. Apicella M, Corso S, Giordano S. Targeted therapies for gastric cancer: failures and hopes from clinical trials. Oncotarget [Internet]. 2017 Jan 26;8(34):57654–69. Available from: https://pubmed.ncbi.nlm.nih.gov/28915702
260. Smyth EC, Wotherspoon A, Peckitt C, Gonzalez D, Hulkki-Wilson S, Eltahir Z, et al. Mismatch repair deficiency, microsatellite instability, and survival: An exploratory analysis of the Medical Research Council Adjuvant Gastric Infusional Chemotherapy (MAGIC) trial. JAMA Oncol. 2017;3(9):1197–203.
261. Choi YY, Kim H, Shin S-J, Kim HY, Lee J, Yang H-K, et al. Microsatellite Instability and Programmed Cell Death-Ligand 1 Expression in Stage II/III Gastric Cancer: Post Hoc Analysis of the CLASSIC Randomized Controlled study. Ann Surg. 2019 Aug;270(2):309–16.
262. Lordick F. Chemotherapy for resectable microsatellite instability-high gastric cancer? Lancet Oncol. 2020 Feb;21(2):203.
263. Smyth EC. Chemotherapy for resectable microsatellite instability-high gastric cancer? Lancet Oncol. 2020 Feb;21(2):204.

264. Guinney J, Dienstmann R, Wang X, De Reyniès A, Schlicker A, Soneson C, et al. The consensus molecular subtypes of colorectal cancer. Nat Med. 2015;21(11):1350–6.
265. Kahles A, Lehmann K Van, Toussaint NC, Hüser M, Stark SG, Sachsenberg T, et al. Comprehensive Analysis of Alternative Splicing Across Tumors from 8,705 Patients. Cancer Cell. 2018;34(2):211-224.e6.
266. von Ahlfen S, Missel A, Bendrat K, Schlumpberger M. Determinants of RNA Quality from FFPE Samples. PLoS One [Internet]. 2007 Dec 5;2(12):e1261. Available from: https://doi.org/10.1371/journal.pone.0001261
267. Abdueva D, Wing M, Schaub B, Triche T, Davicioni E. Quantitative expression profiling in formalin-fixed paraffin-embedded samples by Affymetrix microarrays. J Mol Diagnostics [Internet]. 2010;12(4):409–17. Available from: http://dx.doi.org/10.2353/jmoldx.2010.090155
268. Senn SJ. Dichotomania: an obsessive compulsive disorder that is badly affecting the quality of analysis of pharmaceutical trials. Proc Int Stat Institute, 55th Sess. 2005;1–13.
269. Giannoni A, Baruah R, Leong T, Rehman MB, Pastormerlo LE, Harrell FE, et al. Do optimal prognostic thresholds in continuous physiological variables really exist? Analysis of origin of apparent thresholds, with systematic review for peak oxygen consumption, ejection fraction and BNP. PLoS One. 2014;9(1):e81699.
270. Altman DG, Royston P. The cost of dichotomising continuous variables. Br Med J. 2006;332(7549):1080.
271. Lee J, Sohn I, Do I-G, Kim K-M, Park SH, Park JO, et al. Nanostring-based multigene assay to predict recurrence for gastric cancer patients after surgery. PLoS One. 2014;9(3):e90133.
272. Ooi CH, Ivanova T, Wu J, Lee M, Tan IB, Tao J, et al. Oncogenic pathway combinations predict clinical prognosis in gastric cancer. PLoS Genet. 2009 Oct;5(10):e1000676.
273. Cho JY, Lim JY, Cheong JH, Park Y-Y, Yoon S-L, Kim SM, et al. Gene expression signature-based prognostic risk score in gastric cancer. Clin cancer Res an Off J Am Assoc Cancer Res. 2011 Apr;17(7):1850–7.
274. Lei Z, Tan IB, Das K, Deng N, Zouridis H, Pattison S, et al. Identification of molecular subtypes of gastric cancer with different responses to PI3-kinase inhibitors and 5-fluorouracil. Gastroenterology. 2013 Sep;145(3):554–65.
275. Yoon S-J, Park J, Shin Y, Choi Y, Park SW, Kang S-G, et al. Deconvolution of diffuse gastric cancer and the suppression of CD34 on the BALB/c nude mice model. BMC Cancer. 2020 Apr;20(1):314.
276. Qian Z, Zhu G, Tang L, Wang M, Zhang L, Fu J, et al. Whole genome gene copy number profiling of gastric cancer identifies PAK1 and KRAS gene amplification as therapy targets. Genes Chromosomes Cancer. 2014 Nov;53(11):883–94.
277. Sean D, Meltzer PS. GEOquery: A bridge between the Gene Expression Omnibus (GEO) and BioConductor. Bioinformatics. 2007;23(14):1846–7.
278. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal. 2013 Apr;6(269):pl1.
279. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012 May;2(5):401–4.
280. Mounir M, Lucchetta M, Silva TC, Olsen C, Bontempi G, Chen X, et al. New functionalities in the TCGAbiolinks package for the study and integration of cancer data from GDC and GTEx. PLOS Comput Biol [Internet]. 2019 Mar 5;15(3):e1006701. Available from: https://doi.org/10.1371/journal.pcbi.1006701
281. Silva TC, Colaprico A, Olsen C, D'Angelo F, Bontempi G, Ceccarelli M, et al. TCGA Workflow: Analyze cancer genomics and epigenomics data using Bioconductor

packages. F1000Research. 2016;5:1542.
282. Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, et al. TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. Nucleic Acids Res. 2016 May;44(8):e71.
283. Parrish N, Hormozdiari F, Eskin E. Assembly of non-unique insertion content using next-generation sequencing. Bioinforma Impact Accurate Quantif Proteomic Genet Anal Res. 2014;21–40.
284. Tetzlaff MT, Calderone TL, Davies MA, Lazar AJ, Jeffrey E. Biological Validation of RNA Sequencing Data From Formalin- Fixed Paraffin-Embedded Primary Melanomas. 2020;1–19.
285. Richard AC, Lyons PA, Peters JE, Biasci D, Flint SM, Lee JC, et al. Comparison of gene expression microarray data with count-based RNA measurements informs microarray interpretation. 2014;1–10.
286. Smyth GK, Ritchie M, Thorne N. Linear Models for Microarray Data User's Guide. Bioinformatics [Internet]. 2011;20(May):3705–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15297296
287. Liaw A, Wiener M. Classification and Regression by randomForest. R News. 2002;2(3):18–22.
288. Karatzoglou A, Smola A, Hornik K, Zeileis A. kernlab - An S4 Package for Kernel Methods in R. J Stat Softw [Internet]. 2004 Nov 2;11(9 SE-Articles):1–20. Available from: https://www.jstatsoft.org/index.php/jss/article/view/v011i09
289. Wickham H. The Split-Apply-Combine Strategy for Data Analysis. J Stat Softw [Internet]. 2011 Apr 7;40(1 SE-Articles):1–29. Available from: https://www.jstatsoft.org/index.php/jss/article/view/v040i01
290. Greenwell B, Boehmke B, Cunningham J, GBM Developers. gbm: Generalized Boosted Regression Models [Internet]. 2020. Available from: https://cran.r-project.org/package=gbm
291. Venables WN, Ripley BD. Modern Applied Statistics with S. In: 4th ed. Springer New York, NY; 2002.
292. Weihs C, Ligges U, Luebke K, Raabe N. klaR Analyzing German Business Cycles BT - Data Analysis and Decision Support. In: Baier D, Decker R, Schmidt-Thieme L, editors. Berlin, Heidelberg: Springer Berlin Heidelberg; 2005. p. 335–43. Available from: https://doi.org/10.1007/3-540-28397-8_36
293. Id AV, Gowen E, Poliakoff E, Casson AJ. Machine learning algorithm validation with a limited sample size. 2019;1–20. Available from: http://dx.doi.org/10.1371/journal.pone.0224365
294. Diamantidis NA, Karlis D, Giakoumakis EA. Unsupervised stratification of cross-validation for accuracy estimation. Artif Intell. 2000;116(1–2):1–16.
295. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977 Mar;33(1):159–74.
296. BRIER GW. VERIFICATION OF FORECASTS EXPRESSED IN TERMS OF PROBABILITY. Mon Weather Rev [Internet]. 1950;78(1):1–3. Available from: https://journals.ametsoc.org/view/journals/mwre/78/1/1520-0493_1950_078_0001_vofeit_2_0_co_2.xml
297. Franks JM, Cai G, Whitfield ML. Gene expression Feature specific quantile normalization enables cross-platform classification of molecular subtypes using gene expression data. 2018;34(January):1868–74.
298. Austin PC, Steyerberg EW. Graphical assessment of internal and external calibration of logistic regression models by using loess smoothers. Stat Med [Internet]. 2013/08/23. 2014 Feb 10;33(3):517–35. Available from: https://pubmed.ncbi.nlm.nih.gov/24002997
299. Kull M, Perello-Nieto M, Kängsepp M, Filho TS, Song H, Flach P. Beyond temperature scaling: Obtaining well-calibrated multiclass probabilities with dirichlet calibration.

Adv Neural Inf Process Syst. 2019;32(NeurIPS 2019).
300. Platt J. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. Adv large margin Classif. 1999;10(3):61–74.
301. Guo C, Pleiss G, Sun Y, Weinberger KQ. On Calibration of Modern Neural Networks. In: Precup D, Teh YW, editors. Proceedings of the 34th International Conference on Machine Learning [Internet]. PMLR; 2017. p. 1321–30. (Proceedings of Machine Learning Research; vol. 70). Available from: https://proceedings.mlr.press/v70/guo17a.html
302. Widmann D, Lindsten F, Zachariah D. Calibration tests in multi-class classification: A unifying framework. Adv Neural Inf Process Syst. 2019;32(NeurIPS).
303. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics [Internet]. 2011;12(1):77. Available from: https://doi.org/10.1186/1471-2105-12-77
304. Ho D, Imai K, King G, Stuart EA. MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. J Stat Softw [Internet]. 2011 Jun 14;42(8 SE-Articles):1–28. Available from: https://www.jstatsoft.org/index.php/jss/article/view/v042i08
305. VanderWeele TJ, Ding P. Sensitivity Analysis in Observational Research: Introducing the E-Value. Ann Intern Med. 2017 Aug;167(4):268–74.
306. Therneau TM, Grambsch PM. Modeling Survival Data: Extending the Cox Model. Springer, New York; 2000.
307. Gandrud C. simPH: An R Package for Illustrating Estimates from Cox Proportional Hazard Models Including. J Stat Softw. 2015;65(3):1–20.
308. Haider H, Hoehn B, Davis S, Greiner R. Effective ways to build and evaluate individual survival distributions. J Mach Learn Res. 2020;21:1–63.
309. Gulley ML, Tang W. Laboratory assays for Epstein-Barr virus-related disease. J Mol Diagn. 2008 Jul;10(4):279–92.
310. Zerhouni E, Umar A, Boland CR, Terdiman JP, Syngal S, de la Chapelle A, et al. Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. J Natl Cancer Inst [Internet]. 2004 Feb 18;96(4):261–8. Available from: https://www.ncbi.nlm.nih.gov/pubmed/14970275
311. Suraweera N, Duval A, Reperant M, Vaury C, Furlan D, Leroy K, et al. Evaluation of tumor microsatellite instability using five quasimonomorphic mononucleotide repeats and pentaplex PCR. Gastroenterology. 2002;123(6):1804–11.
312. Marisa L, Blum Y, Taieb J, Ayadi M, Pilati C, Le Malicot K, et al. Intratumor CMS heterogeneity impacts patient prognosis in localized colon cancer. Clin Cancer Res. 2021;27(17):4768–80.
313. Carter SL, Cibulskis K, Helman E, McKenna A, Shen H, Zack T, et al. Absolute quantification of somatic DNA alterations in human cancer. Nat Biotechnol [Internet]. 2012 May;30(5):413–21. Available from: https://pubmed.ncbi.nlm.nih.gov/22544022
314. Jiang Y, Li T, Liang X, Hu Y, Huang L, Liao Z, et al. Association of adjuvant chemotherapy with survival in patients with stage II or III gastric cancer. JAMA Surg. 2017;152(7):e171087.
315. Paoletti X, Oba K, Burzykowski T, Michiels S, Ohashi Y, Pignon J-P, et al. Benefit of adjuvant chemotherapy for resectable gastric cancer: a meta-analysis. JAMA. 2010 May;303(17):1729–37.
316. Sun K, Jia K, Lv H, Wang S-Q, Wu Y, Lei H, et al. EBV-Positive Gastric Cancer: Current Knowledge and Future Perspectives [Internet]. Vol. 10, Frontiers in Oncology . 2020. Available from: https://www.frontiersin.org/article/10.3389/fonc.2020.583463
317. van Beek J, zur Hausen A, Snel SN, Berkhof J, Kranenbarg EK, van de Velde CJH, et al. Morphological evidence of an activated cytotoxic T-cell infiltrate in EBV-positive

gastric carcinoma preventing lymph node metastases. Am J Surg Pathol. 2006 Jan;30(1):59–65.
318. Le DT, Durham JN, Smith KN, Wang H, Bjarne R, Aulakh LK, et al. Mismatch-repair deficiency predicts response of solid tumors to PD-1 blockade. 2017;357(6349):409–13.
319. Ostroumov D, Fekete-Drimusz N, Saborowski M, Kühnel F, Woller N. CD4 and CD8 T lymphocyte interplay in controlling tumor growth. Cell Mol Life Sci [Internet]. 2017/10/14. 2018 Feb;75(4):689–713. Available from: https://pubmed.ncbi.nlm.nih.gov/29032503
320. Smith CC, Selitsky SR, Chai S, Armistead PM, Vincent BG, Serody JS. Alternative tumour-specific antigens. Nat Rev Cancer [Internet]. 2019;19(8):465–78. Available from: http://dx.doi.org/10.1038/s41568-019-0162-4
321. Marusyk A, Janiszewska M, Polyak K. Intratumor Heterogeneity: The Rosetta Stone of Therapy Resistance. Cancer Cell. 2020 Apr;37(4):471–84.
322. Wang R, Dang M, Harada K, Han G, Wang F, Pool Pizzi M, et al. Single-cell dissection of intratumoral heterogeneity and lineage diversity in metastatic gastric adenocarcinoma. Nat Med [Internet]. 2021;27(1):141–51. Available from: http://dx.doi.org/10.1038/s41591-020-1125-8
323. Röcken C, Amallraja A, Halske C, Opasic L, Traulsen A, Behrens HM, et al. Multiscale heterogeneity in gastric adenocarcinoma evolution is an obstacle to precision medicine. Genome Med. 2021;13(1):1–19.
324. Hiley C, de Bruin EC, McGranahan N, Swanton C. Deciphering intratumor heterogeneity and temporal acquisition of driver events to refine precision medicine. Genome Biol. 2014;15(8):1–10.
325. Blum Y, Meiller C, Quetel L, Elarouci N, Ayadi M, Tashtanbaeva D, et al. Dissecting heterogeneity in malignant pleural mesothelioma through histo-molecular gradients for clinical applications. Nat Commun [Internet]. 2019;10(1):1333. Available from: https://doi.org/10.1038/s41467-019-09307-6
326. Wu Z, Aryee MJ. Subset Quantile Normalization Using Negative Control Features. 2010;17(10):1385–95.
327. Zhao Y, Wong L, Wen W, Goh B. How to do quantile normalization correctly for gene expression data analyses. Sci Rep [Internet]. 2020;1–11. Available from: https://doi.org/10.1038/s41598-020-72664-6
328. Castillo D, Gálvez JM, Herrera LJ, Román BS, Rojas F, Rojas I. Integration of RNA-Seq data with heterogeneous microarray data for breast cancer profiling. 2017;1–15.
329. Olchanski N, Cohen JT, Neumann PJ, Wong JB, Kent DM. Understanding the Value of Individualized Information: The Impact of Poor Calibration or Discrimination in Outcome Prediction Models. Med Decis Mak [Internet]. 2017 Apr 11;37(7):790–801. Available from: https://doi.org/10.1177/0272989X17704855
330. Van Calster B, McLernon DJ, Van Smeden M, Wynants L, Steyerberg EW, Bossuyt P, et al. Calibration: The Achilles heel of predictive analytics. BMC Med. 2019;17(1):1–7.
331. Minderer M, Djolonga J, Romijnders R, Hubis F, Zhai X, Houlsby N, et al. Revisiting the calibration of modern neural networks. Adv Neural Inf Process Syst [Internet]. 2021;34(NeurIPS). Available from: http://arxiv.org/abs/2106.07998
332. Pectasides E, Stachler MD, Derks S, Liu Y, Maron S, Islam M, et al. Genomic heterogeneity as a barrier to precision medicine in gastroesophageal adenocarcinoma. Cancer Discov. 2018;8(1):37–48.
333. Conway JR, Warner JL, Rubinstein WS, Miller RS. Next-Generation Sequencing and the Clinical Oncology Workflow: Data Challenges, Proposed Solutions, and a Call to Action. JCO Precis Oncol [Internet]. 2019 Oct 1;(3):1–10. Available from: https://doi.org/10.1200/PO.19.00232
334. Molania R, Gagnon-Bartsch JA, Dobrovic A, Speed TP. A new normalization for Nanostring nCounter gene expression data. Nucleic Acids Res [Internet]. 2019 Jul

9;47(12):6073–83. Available from: https://doi.org/10.1093/nar/gkz433
335. Jacob L, Gagnon-Bartsch JA, Speed TP. Correcting gene expression data when neither the unwanted variation nor the factor of interest are observed. Biostatistics. 2016 Jan;17(1):16–28.
336. Gagnon-Bartsch JA, Speed TP. Using control genes to correct for unwanted variation in microarray data. Biostatistics. 2012 Jul;13(3):539–52.
337. Yadav B, Pemovska T, Szwajda A, Kulesskiy E, Kontro M, Karjalainen R, et al. Quantitative scoring of differential drug sensitivity for individually optimized anticancer therapies. Sci Rep [Internet]. 2014;4(1):5193. Available from: https://doi.org/10.1038/srep05193
338. Ghandi M, Huang FW, Jané-Valbuena J, Kryukov G V, Lo CC, McDonald ER, et al. Next-generation characterization of the Cancer Cell Line Encyclopedia. Nature [Internet]. 2019;569(7757):503–8. Available from: https://doi.org/10.1038/s41586-019-1186-3
339. Chan EM, Shibue T, McFarland JM, Gaeta B, Ghandi M, Dumont N, et al. WRN helicase is a synthetic lethal target in microsatellite unstable cancers. Nature [Internet]. 2019;568(7753):551–6. Available from: https://doi.org/10.1038/s41586-019-1102-x
340. Cohen-Sharir Y, McFarland JM, Abdusamad M, Marquis C, Bernhard S V, Kazachkova M, et al. Aneuploidy renders cancer cells vulnerable to mitotic checkpoint inhibition. Nature [Internet]. 2021;590(7846):486–91. Available from: https://doi.org/10.1038/s41586-020-03114-6
341. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, et al. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Res [Internet]. 2013 Jan 1;41(D1):D955–61. Available from: https://doi.org/10.1093/nar/gks1111
342. Veldman-Jones MH, Brant R, Rooney C, Geh C, Emery H, Harbron CG, et al. Evaluating robustness and sensitivity of the nanostring technologies ncounter platform to enable multiplexed gene expression analysis of clinical samples. Cancer Res. 2015;75(13):2587–93.
343. Veldman-Jones MH, Lai Z, Wappett M, Harbron CG, Barrett JC, Harrington EA, et al. Reproducible, quantitative, and flexible molecular subtyping of clinical DLBCL samples using the NanoString nCounter system. Clin Cancer Res. 2015;21(10):2367–78.
344. Junet V, Farrés J, Mas JM, Daura X. CuBlock: a cross-platform normalization method for gene-expression microarrays. Bioinformatics [Internet]. 2021 Aug 15;37(16):2365–73. Available from: https://doi.org/10.1093/bioinformatics/btab105
345. Zhang S, Shao J, Yu D, Qiu X, Zhang J. MatchMixeR: a cross-platform normalization method for gene expression data integration. Bioinformatics [Internet]. 2020 Apr 15;36(8):2486–91. Available from: https://doi.org/10.1093/bioinformatics/btz974
346. Park YH, Lal S, Lee JE, Choi Y La, Wen J, Ram S, et al. Chemotherapy induces dynamic immune responses in breast cancers that impact treatment outcome. Nat Commun [Internet]. 2020;11(1):1–14. Available from: http://dx.doi.org/10.1038/s41467-020-19933-0
347. Retamero JA, Aneiros-Fernandez J, Del Moral RG. Complete Digital Pathology for Routine Histopathology Diagnosis in a Multicenter Hospital Network. Arch Pathol Lab Med. 2020 Feb;144(2):221–8.
348. Volynskaya Z, Chow H, Evans A, Wolff A, Lagmay-Traya; C, Asa SL. Integrated Pathology Informatics Enables High-Quality Personalized and Precision Medicine: Digital Pathology and Beyond. Arch Pathol Lab Med [Internet]. 2017 Aug 29;142(3):369–82. Available from: https://doi.org/10.5858/arpa.2017-0139-OA
349. Baxi V, Edwards R, Montalto M, Saha S. Digital pathology and artificial intelligence in translational medicine and clinical practice. Mod Pathol [Internet]. 2022;35(1):23–32. Available from: https://doi.org/10.1038/s41379-021-00919-2
350. Lin S, Huang G, Cheng L, Li Z, Xiao Y, Deng Q, et al. Establishment of peripheral

blood mononuclear cell-derived humanized lung cancer mouse models for studying efficacy of PD-L1/PD-1 targeted immunotherapy. MAbs [Internet]. 2018;10(8):1301–11. Available from: https://doi.org/10.1080/19420862.2018.1518948
351. Choi Y, Lee S, Kim K, Kim SH, Chung YJ, Lee C. Studying cancer immunotherapy using patient-derived xenografts (PDXs) in humanized mice. Exp Mol Med [Internet]. 2018;50(8). Available from: http://dx.doi.org/10.1038/s12276-018-0115-0

Appendices

A. Appendix Material

Supplemental material for this thesis is presented here by chapter or in the online supplement:

Chapter 2 Appendix

Chapter 3 Appendix

Chapter 4 Appendix

Chapter 5 Appendix

Online: https://github.com/skubleny/Thesis-Supplemental-Files

B. Chapter 2 Appendix

1. Comparison of IHC biomarkers with clinicopathologic factors

Stage (n = 43)

Variable	Stage I, N = 11[1]	Stage II, N = 10[1]	Stage III, N = 8[1]	Stage IV, N = 14[1]	p-value[2]
CD4/CD8 Ratio	1.8 (1.3, 3.0)	2.4 (1.6, 2.8)	1.5 (1.0, 2.4)	1.4 (1.0, 2.8)	0.8
Unknown	0	0	0	1	
CD4	14 (11, 31)	16 (12, 21)	6 (4, 11)	15 (8, 23)	0.11
CD8	8 (4, 19)	7 (4, 11)	7 (3, 9)	8 (5, 12)	0.7
Unknown	0	0	0	1	
Galectin-3	43 (30, 53)	47 (35, 53)	45 (25, 55)	46 (33, 68)	0.9
E-cadherin	19 (10, 28)	16 (8, 30)	24 (21, 27)	7 (1, 21)	0.2
E-cadherin H-score	22 (13, 38)	21 (9, 46)	28 (26, 38)	8 (1, 32)	0.15

[1]Median (IQR)
[2]Kruskal-Wallis rank sum test

Lymphovascular Invasion (n = 35)

Variable	Negative, N = 19[1]	Positive, N = 16[1]	p-value[2]
CD4/CD8 Ratio	2.24 (1.46, 2.93)	1.34 (0.87, 2.56)	0.12
CD4	13 (10, 24)	14 (6, 23)	0.5
CD8	8 (3, 11)	8 (6, 11)	0.9
Galectin-3	43 (29, 53)	45 (37, 55)	0.7
E-cadherin	20 (13, 28)	8 (5, 25)	0.2
E-cadherin H-score	27 (14, 41)	11 (5, 30)	0.2

[1]Median (IQR)
[2]Wilcoxon rank sum exact test

Carcinomatosis (n = 38)

Variable	Negative, N = 29[1]	Positive, N = 9[1]	p-value[2]
CD4/CD8 Ratio	1.7 (1.2, 2.8)	1.7 (1.1, 4.9)	>0.9
Unknown	0	1	
CD4	12 (6, 24)	16 (10, 18)	0.9
CD8	8 (4, 11)	8 (4, 14)	0.8
Unknown	0	1	
Galectin-3	43 (30, 55)	47 (44, 70)	0.5
E-cadherin	20 (7, 27)	9 (2, 25)	0.4
E-cadherin H-score	25 (7, 38)	15 (3, 37)	0.4

[1]Median (IQR)
[2]Wilcoxon rank sum exact test

Perineural Invasion (n = 33)

Variable	Negative, N = 20[1]	Positive, N = 11[1]	p-value[2]
CD4/CD8 Ratio	2.49 (1.61, 3.09)	1.07 (0.76, 2.37)	0.054
CD4	15 (11, 24)	7 (5, 24)	0.2
CD8	8 (3, 10)	8 (5, 11)	0.9
Galectin-3	43 (30, 49)	53 (44, 63)	0.066
E-cadherin	20 (6, 30)	18 (6, 24)	0.6
E-cadherin H-score	25 (7, 44)	23 (7, 28)	0.7

[1]Median (IQR)
[2]Wilcoxon rank sum exact test

Tumour Grade (n = 43)

Variable	G1, N = 1[1]	G2, N = 10[1]	G3, N = 31[1]	Gx, N = 1[1]	p-value[2]
CD4/CD8 Ratio	1.4 (1.4, 1.4)	2.1 (1.6, 2.8)	1.9 (1.1, 3.0)	1.6 (1.6, 1.6)	0.9
Unknown	0	0	1	0	
CD4	6 (6, 6)	13 (8, 22)	14 (7, 24)	14 (14, 14)	0.7
CD8	4 (4, 4)	7 (3, 10)	8 (5, 12)	8 (8, 8)	0.7
Unknown	0	0	1	0	
Galectin-3	17 (17, 17)	39 (29, 47)	47 (34, 63)	69 (69, 69)	0.2
E-cadherin	26 (26, 26)	19 (7, 30)	16 (7, 28)	3 (3, 3)	0.5
E-cadherin H-score	32 (32, 32)	23 (8, 42)	22 (7, 40)	3 (3, 3)	0.5

[1]Median (IQR)
[2]Kruskal-Wallis rank sum test
G1 = Well differentiated, G2 = Moderately differentiated, G3 = Poorly differentiated, Gx = Undetermined

Lauren Classification (n = 41)

Variable	Diffuse, N = 26[1]	Intestinal, N = 13[1]	Mixed, N = 2[1]	p-value[2]
CD4/CD8 Ratio	2.1 (1.2, 2.8)	1.6 (1.2, 2.5)	2.6 (2.2, 3.0)	0.7
Unknown	1	0	0	
CD4	13 (8, 21)	15 (7, 24)	16 (11, 21)	>0.9
CD8	8 (5, 11)	8 (4, 14)	5 (4, 7)	0.7
Unknown	1	0	0	
Galectin-3	47 (32, 67)	37 (29, 48)	37 (34, 41)	0.4
E-cadherin	11 (6, 25)	25 (13, 31)	3 (2, 4)	0.043
E-cadherin H-score	17 (7, 35)	29 (15, 44)	3 (2, 5)	0.057

[1]Median (IQR)
[2]Kruskal-Wallis rank sum test

Tumour Location (n = 43)

Variable	Distal, N = 14[1]	Proximal, N = 26[1]	Whole stomach, N = 3[1]	p-value[2]
CD4/CD8 Ratio	2.0 (1.3, 3.1)	1.4 (1.0, 2.8)	2.8 (2.5, 10.3)	0.2
Unknown	0	1	0	
CD4	21 (11, 24)	12 (6, 21)	16 (12, 17)	0.4
CD8	8 (4, 14)	8 (5, 11)	6 (3, 7)	0.6
Unknown	0	1	0	
Galectin-3	45 (38, 53)	46 (30, 61)	20 (20, 45)	0.7
E-cadherin	18 (7, 29)	16 (6, 27)	26 (17, 34)	0.6
E-cadherin H-score	22 (8, 41)	20 (7, 37)	46 (28, 54)	0.5

[1]Median (IQR)
[2]Kruskal-Wallis rank sum test

C. Chapter 3 Appendix

1. Mouse organoid reagents

Reagent	Source	Catalog Number
Advanced DMEM/F12	Gibco	12634010
L-Glut	Sigma	G8540
HEPES	Fisher Scientific	BP310
Penicillin Streptomycin	Gibco	15140122
Amphotericin B	ThermoFisher	15290026
N2 Supplement	Gibco	17502011
B27 Supplement	Gibco	17504044
N-acetylcysteine	Sigma	A9165
Gastrin	Sigma	G9145
Nicotinamide	Sigma	N0636
SB202190/p 38 inhibitor	Sigma	S7067
A83-01 (ALK4/5/7 inhibitor)	Sigma	SML0788
Mouse EGF	Life Technologies	PMG8041
Human EGF	Gibco	PHG0313
Dispase II	Sigma	D4693
Collagenase IX	Sigma	C9407
Y-27632	Sigma	Y0503

2. Immunofluorescence protocols

Protocol	Primary Antibody	Secondary antibody	Antigen Retrieval
1	Rabbit anti-pan cytokeratin, 1:50, overnight 4 degrees Celsius, Abcam, ab9377	Goat anti-rabbit IgG Alexa Fluor 488, 1:100, 30 minutes room temperature, Abcam, ab150077	Sodium Citrate
	Mouse anti-MUC5AC, 1:100, overnight 4 degrees Celsius, Invitrogen,	Goat anti-mouse IgG Alexa Fluor 568, 1:50, 30 minutes room temperature, Invitrogen, A-11004	
2	Rabbit anti-pan cytokeratin, 1:50, overnight 4 degrees Celsius, Abcam, ab9377	Goat anti-rabbit IgG Alexa Fluor 488, 1:100, 30 minutes room temperature, Abcam, ab150077	Sodium Citrate
	Mouse anti-TROY, 1:50, overnight 4 degrees Celsius, Santa Cruz, sc-398526	Goat anti-mouse IgG Alexa Fluor 568, 1:50, 30 minutes room temperature, Invitrogen, A-11004	
3	Rabbit anti-pan cytokeratin, 1:50, overnight 4 degrees Celsius, Abcam, ab9377	Goat anti-rabbit IgG Alexa Fluor 488, 1:100, 30 minutes room temperature, Abcam, ab150077	Sodium Citrate
	Mouse anti-LGR5, 1:50, overnight 4 degrees Celsius, Invitrogen, MA5-25644	Goat anti-mouse IgG Alexa Fluor 568, 1:50, 30 minutes room temperature, Invitrogen, A-11004	

3. qRT-PCR validation and optimization

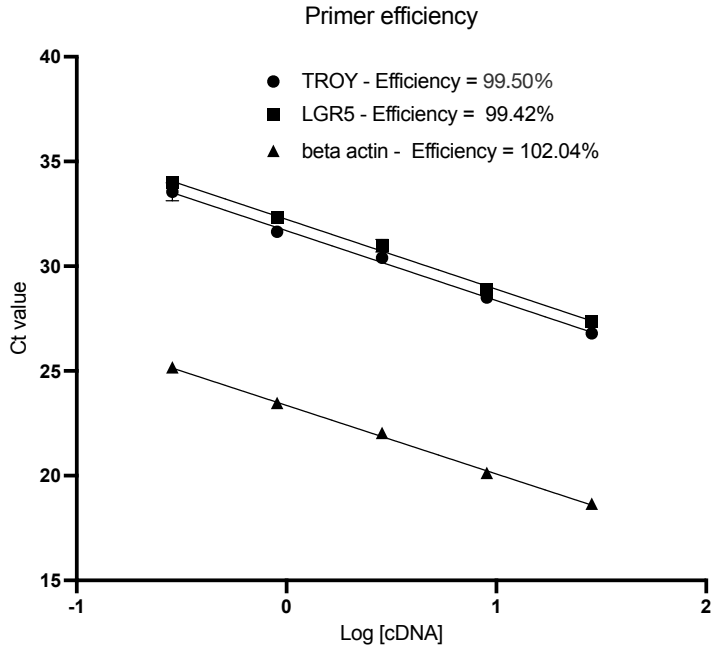


Figure C3-1 qRT-PCR assay primer efficiency.

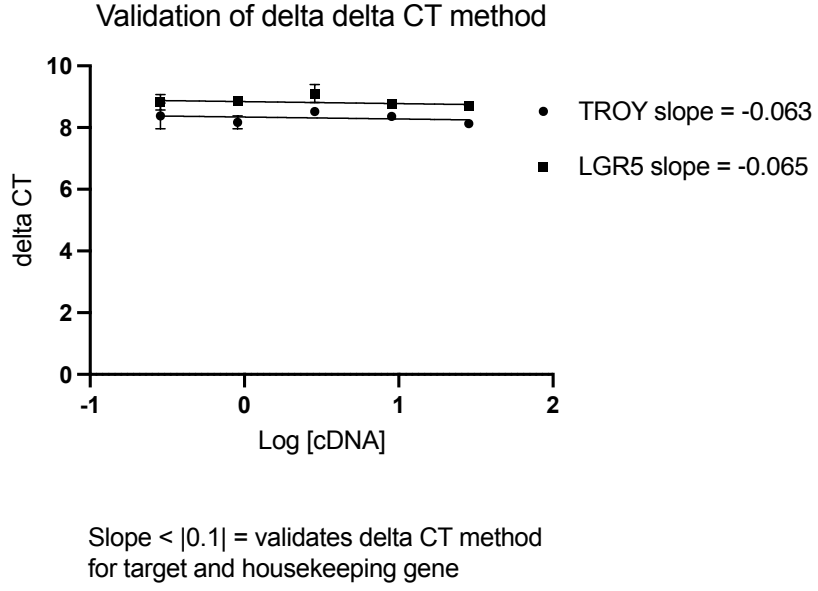


Figure C3-2 Delta-Delta CT validation experiment.

4. Drug assay validation

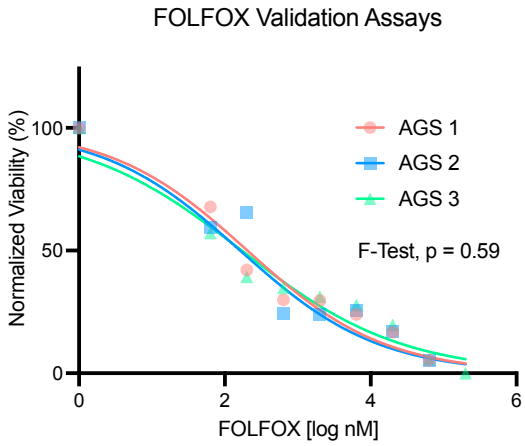


Figure C4-3 FOLFOX *in-vitro* drug assay validation with AGS cell line. Three independent dose-response assays were performed. Cell viability was assessed using a CCK-8 assay. Variable slope least squares linear regression was used to calculate dose-response curves. A sum-of-squares F-Test demonstrated that the was best characterized by a single dose-response curve.

5. Viability organoid assessment

5.1 Unadjusted viability

We analyzed the viability of mouse gastric organoids to determine the effect of shipment media and dissociation time. In Figure S-7-4, we observe the trend of all viability measurements over 10 passages in each cohort. Visually, the cohort variable can be identified as a source of heterogeneity in viability and no obvious pattern relates passage to viability. Thus, despite using consistent protocols and inbred mouse stomachs, the trend of viability in gastric organoids is not consistent or reproducible. In Figure S-7-5, we demonstrate that viability is relatively unaffected by treatment media or dissociation time.

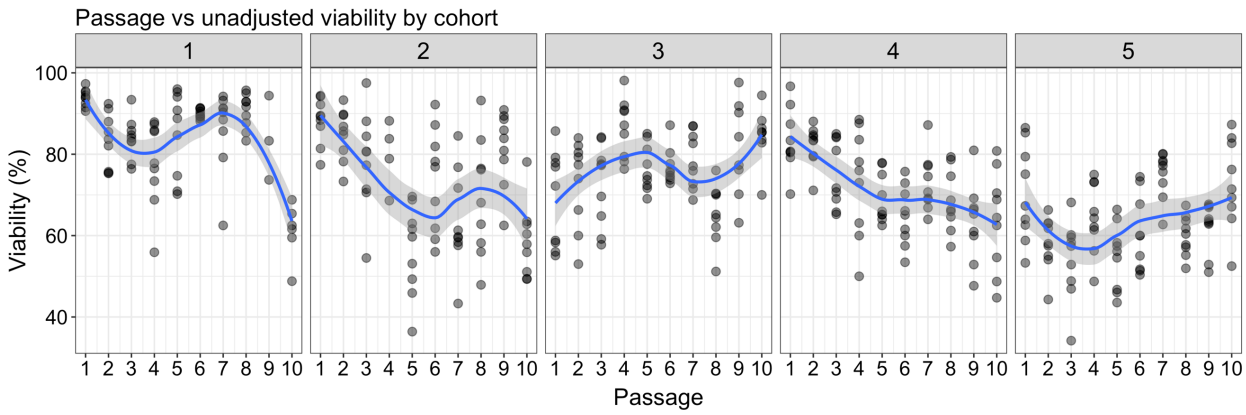


Figure C5-4 Unadjusted organoid viability vs passage number stratified by cohort. The y-axis is the viability measured by hemocytometer and trypan blue. The blue loess smooth is bound by the standard error in gray.

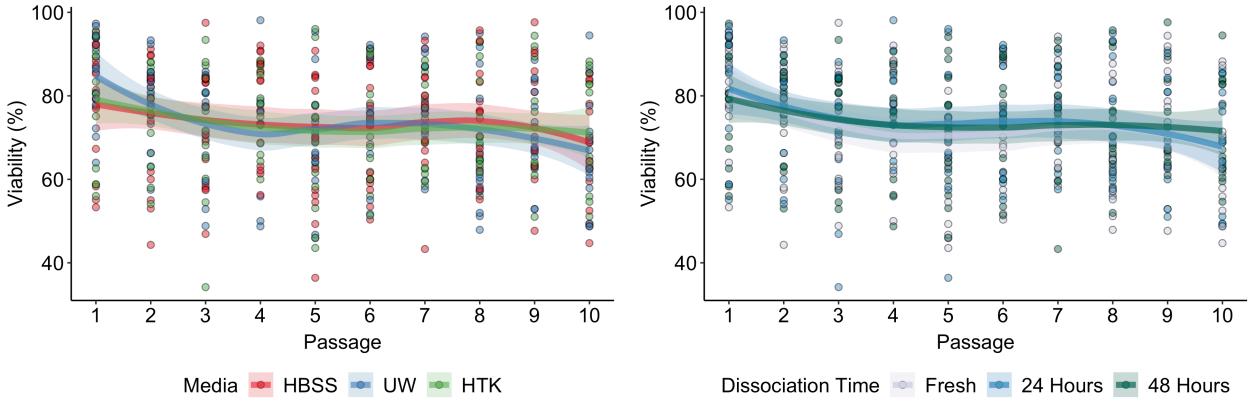


Figure C5-5 Unadjusted viability versus passage number by treatment media (left) and dissociation time (right). The loess smooth is bound by the standard error for each variable factor level and coloured according to the plot legend. Note, the y-axis only spans from 40-100% viability.

5.2 Multivariable models of viability over time

To determine the effect of various confounders to estimating viability in organoid culture we constructed a series of multivariable models. First, we fit multiple linear regression models with treatment media, dissociation time, number of days in culture, passage number and cohort number as variables. After examining the data visually, we observed a nonlinear relationship between viability our continuous variables. Next, we fit linear models with

natural cubic splines for passage number, days in culture and growth rate. Optimal spline degrees of freedom were determined using a likelihood ratio test of nested models and a full model with the optimal spline terms was fit. The model with spline terms was superior to the first model above with an improvement in adjusted R^2 from 0.42 to 0.5. Finally, we fit generalized additive models (GAM) with penalized cubic regression splines. The GAM model was statistically superior to the spline model (Adjusted R^2 = 0.5 to 0.56). We also assessed collinearity of the cubic spline variables with the variable inflation factor (VIF) (Note: GAM is unable to be assessed with variable inflation factor). All VIF values were less than 5, thus signifying the absence of collinearity.

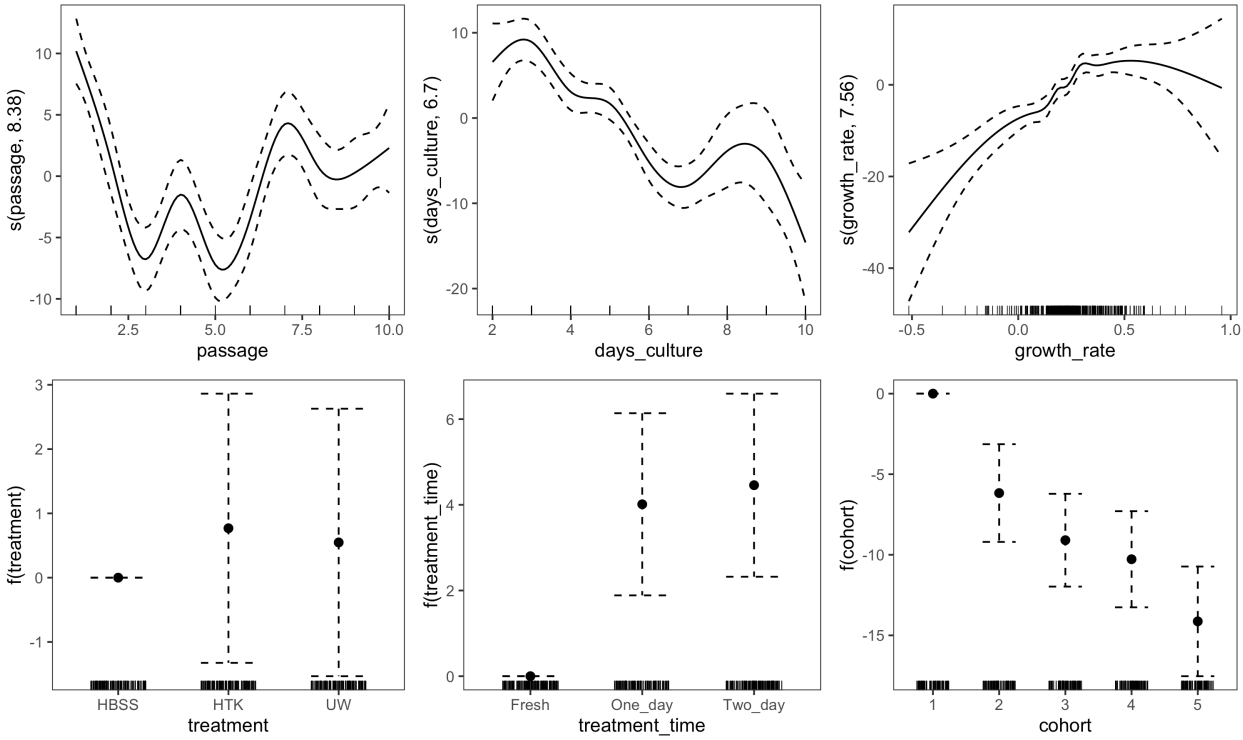


Figure C5-6 Covariable plots from the viability generalized additive model. The top three plots show the functional effect of passage, days in culture and growth rate on viability. The bottom three images show the coefficient estimates for media, dissociation time and cohort in the context of predicting viability.

6. Organoid growth rate assessment

6.1 Unadjusted growth rate

We assessed growth rate in a similar manner to our assessment of viability above.

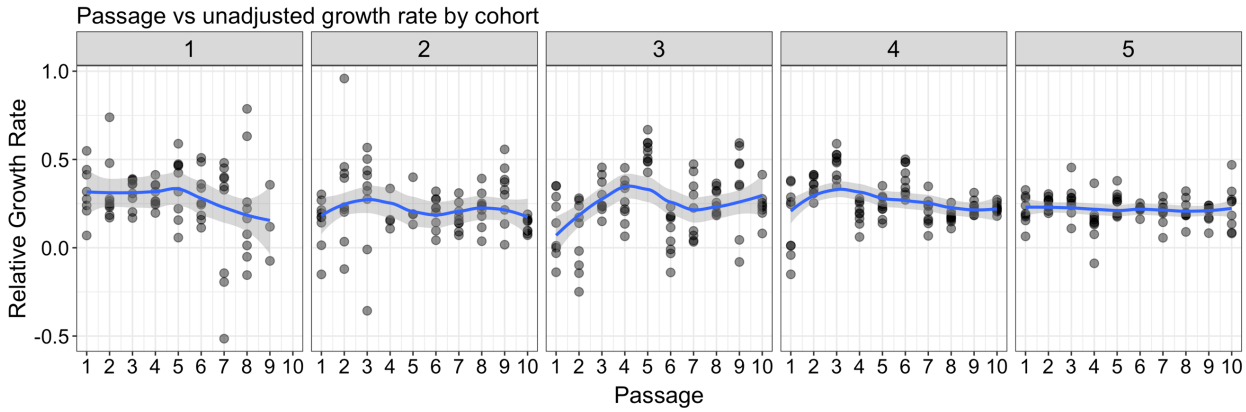


Figure C6-7 Unadjusted organoid growth rate vs passage number stratified by cohort. The y-axis is the growth rate measured by hemocytometer and trypan blue. The blue loess smooth is bound by the standard error in gray.

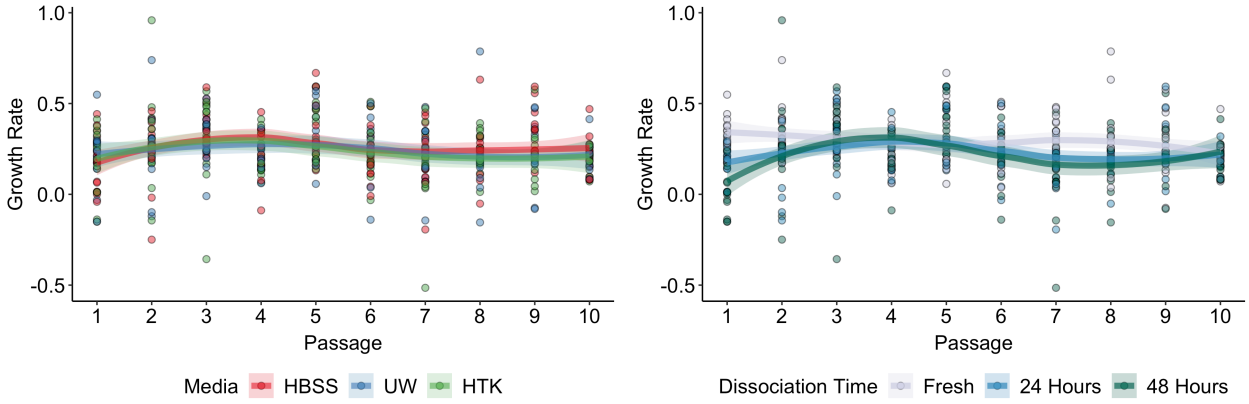


Figure C6-8 Unadjusted growth rate versus passage number by treatment media (left) and dissociation time (right). The loess smooth is bound by the standard error for each variable factor level and coloured according to the plot legend.

6.2 Multivariable models of relative growth rate over time

Similar to our analysis for viability above we performed a multivariable GAM regression to model the relationship of our experimental and confounding variables with growth rate.

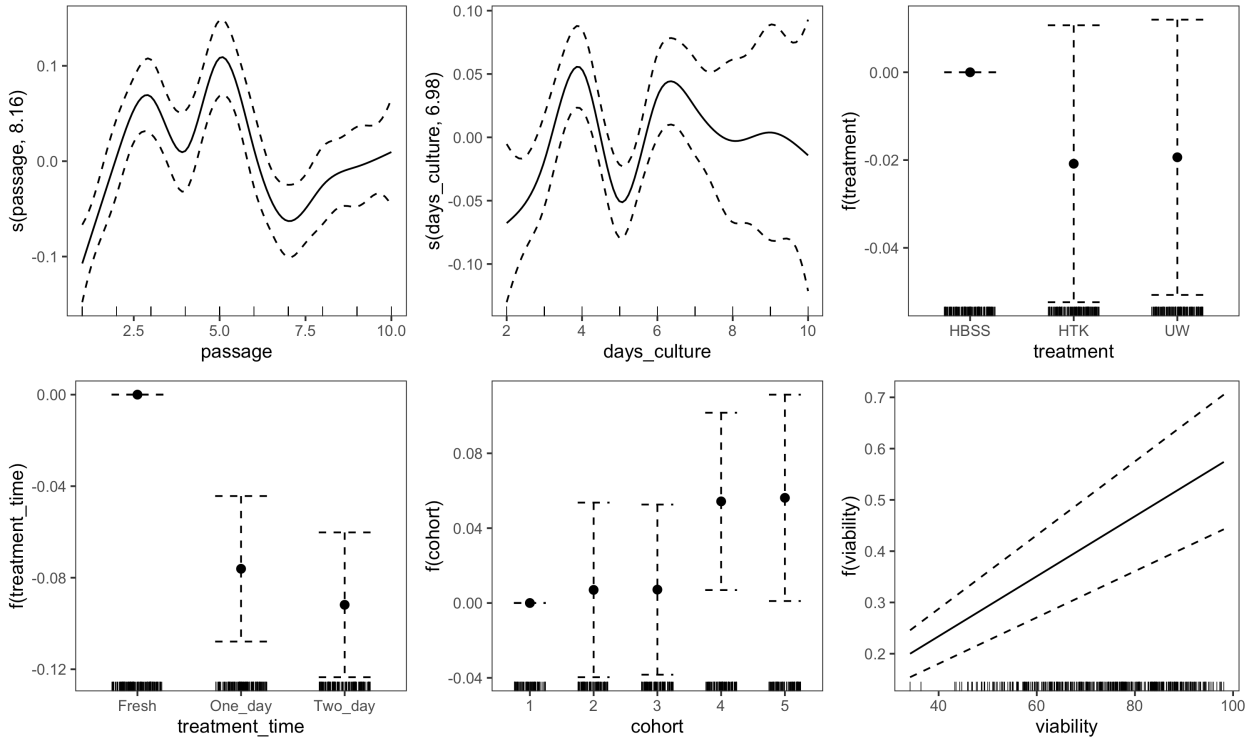


Figure C6-9 Covariable plots from the growth rate generalized additive model. The top three plots show the functional effect of passage, days in culture and shipment media on growth rate. The bottom three images show the coefficient estimates for dissociation time, cohort and viability in the context of predicting growth rate.

D. Chapter 4 Appendix

1. Example of model probability score from machine learning models

	TCGA Subtype				TME Subtype		ACRG Subtype			
ID	CIN	EBV	GS	TCGA MSI	High	Low	MSS TP53-	MSS TP53+	MSI	EMT
1	0.054	0.098	0.126	0.722	0.915	0.085	0.013	0.860	0.126	0.001
2	0.021	0.034	0.896	0.050	0.036	0.964	0.001	0.002	0.000	0.997
3	0.988	0.001	0.008	0.002	0.057	0.943	0.115	0.882	0.002	0.001
4	0.001	0.986	0.005	0.008	0.079	0.921	0.010	0.706	0.282	0.002
5	0.972	0.010	0.006	0.013	0.048	0.952	0.444	0.383	0.170	0.003

2. Propensity Score Matched Analysis

This section contains all components related to our propensity score matched analysis including model development and sensitivity analyses. Following initial matching, we evaluated the nonlinear relationship of continuous variables with overall survival to optimize our match. We found that a natural cubic spline with 2 degrees of freedom was the best representation for age (Table D2-1; Figure D2-10).

Table D2-1 Likelihood ratio tests for nonlinear relationship of age with overall survival

Analysis of Deviance Table

Cox model: response is Overall Survival

Model 1 ~ age

Model 2 ~ ns[1] (age, df=2)

Model 3 ~ ns[1] (age, df=3)

	loglik	Chisq	Df	P(>|Chi|)
Model 1	-2858.2			
Model 2	-2856.2	3.9986	1	0.04554*
Model 3	-2855.2	2.1125	1	0.1461

[1]Natural Cubic Spline fit with specified degrees of freedom (df)

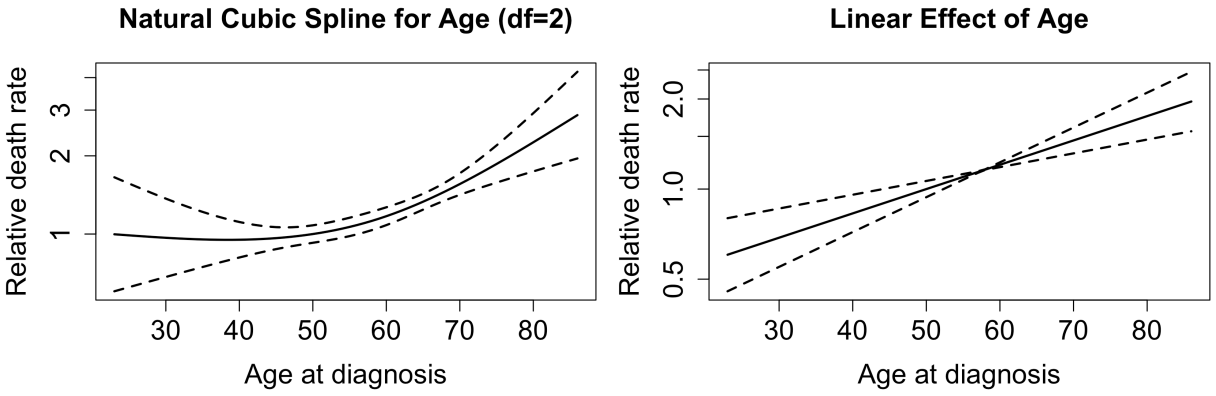


Figure D2-10 Partial residual plots of relative death rate versus age for the propensity score matched analysis. A natural cubic splie with 2 degrees of freedom (Left) provided a statistically optimal representation compared to a standard linear approach (Right). Dotted lines denote 95% confidence interval.

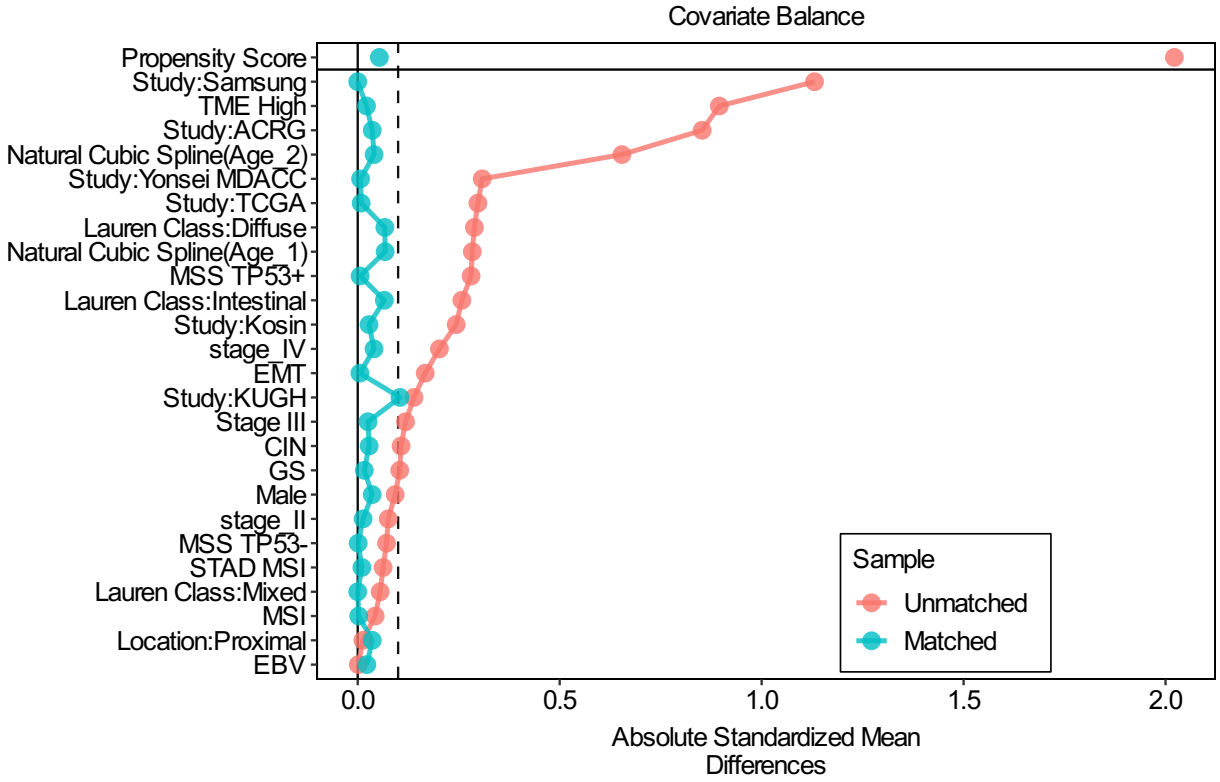


Figure D2-11 Love plot for propensity score matched cohort. Covariates are labeled on the y-axis and the x-axis represents the absolute standardized mean difference. The dotted line represents the optimal match value of 0.1 absolute standardized mean difference. The red points show the difference in unmatched data and the green points show the differences after matching.

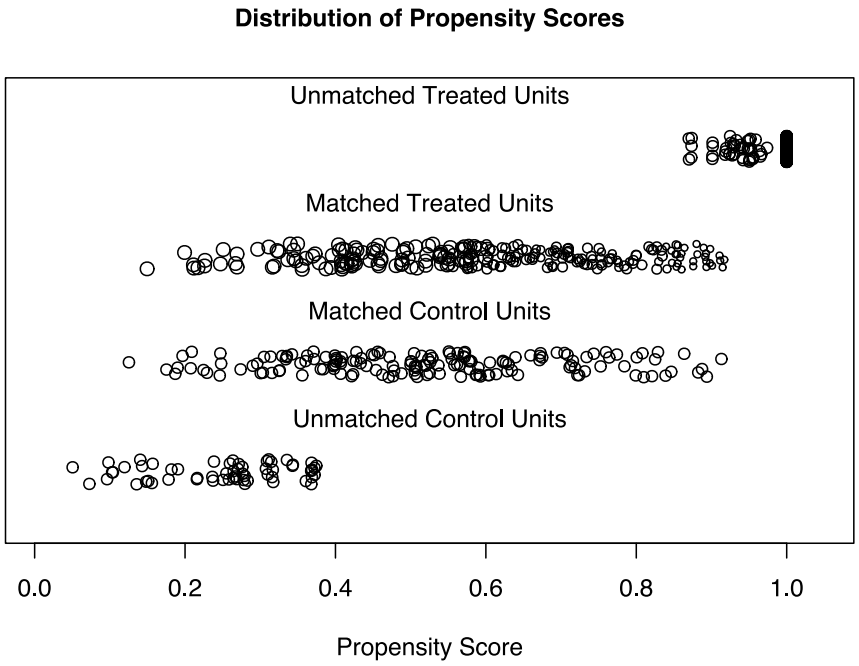


Figure D2-12 Distribution of propensity scores for receiving chemotherapy. The propensity score derived from a logistic regression model is presented on the x-axis.

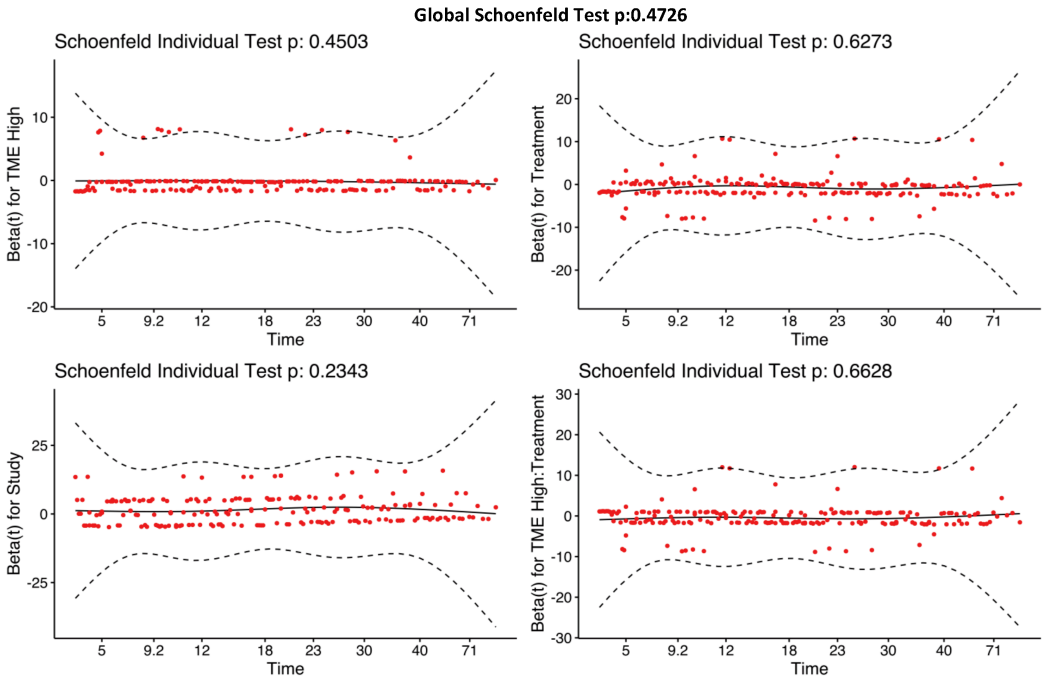


Figure D2-13 Tests for Cox Proportional Hazards Assumptions for variables included in TME and chemotherapy subgroup analysis. A p value > 0.05 means the proportional hazards assumption was satisfied.

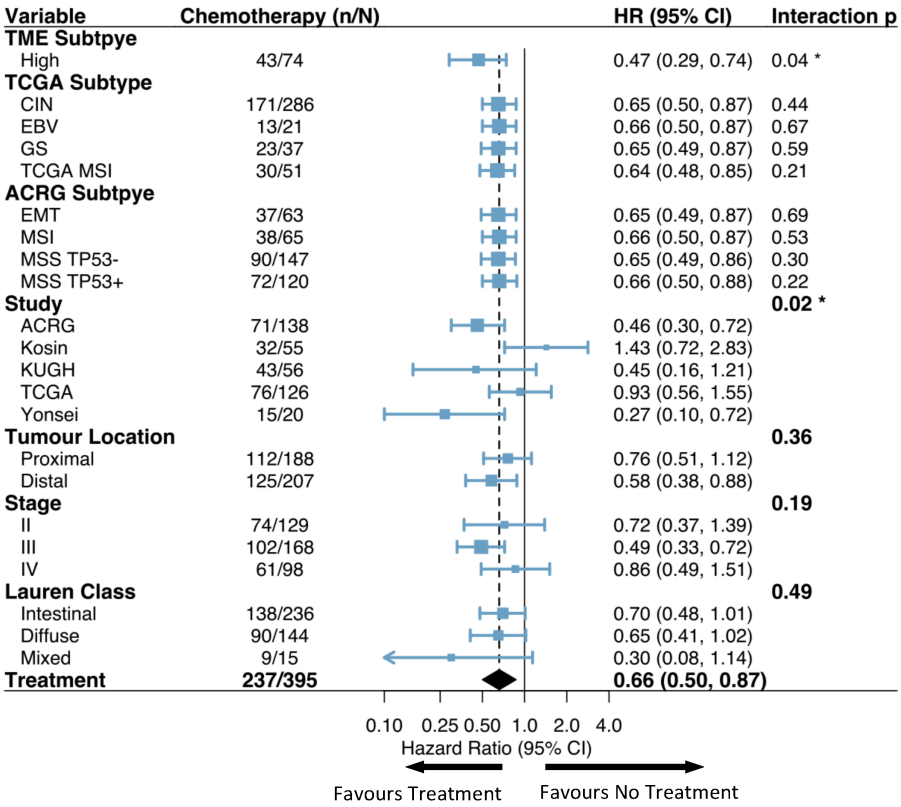


Figure D2-14 Forest plot of subgroup analyses for variables included in propensity score matched cohort. The blue box represents the hazards ratio estimate and the error bars represent the 95% confidence interval for the effect of chemotherapy. The p value for the interaction term between chemotherapy and variable is identified in the last column.

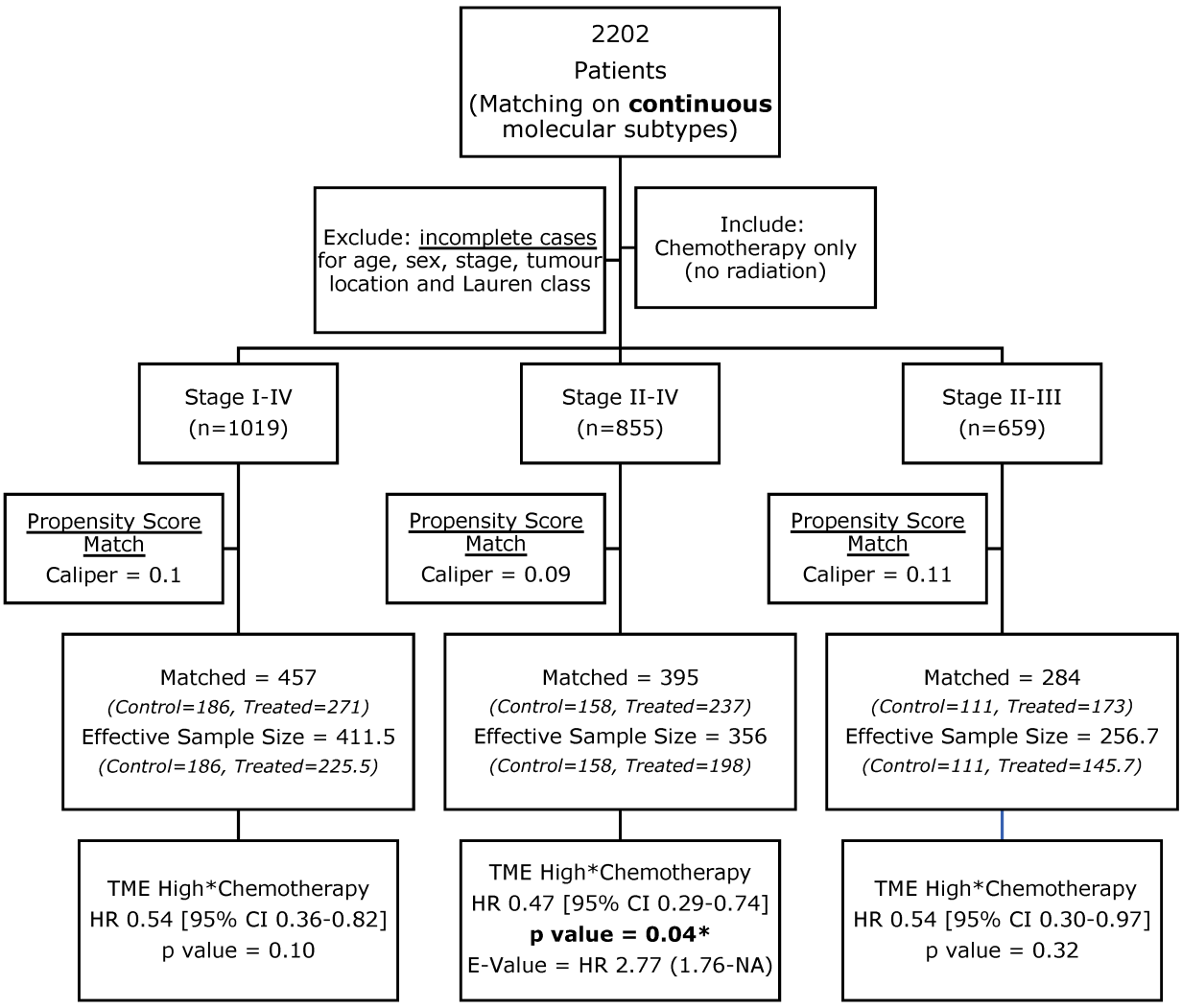


Figure D2-15 Overview of sensitivity analysis for propensity score matched cohort survival models using continuous interpretations of classification.

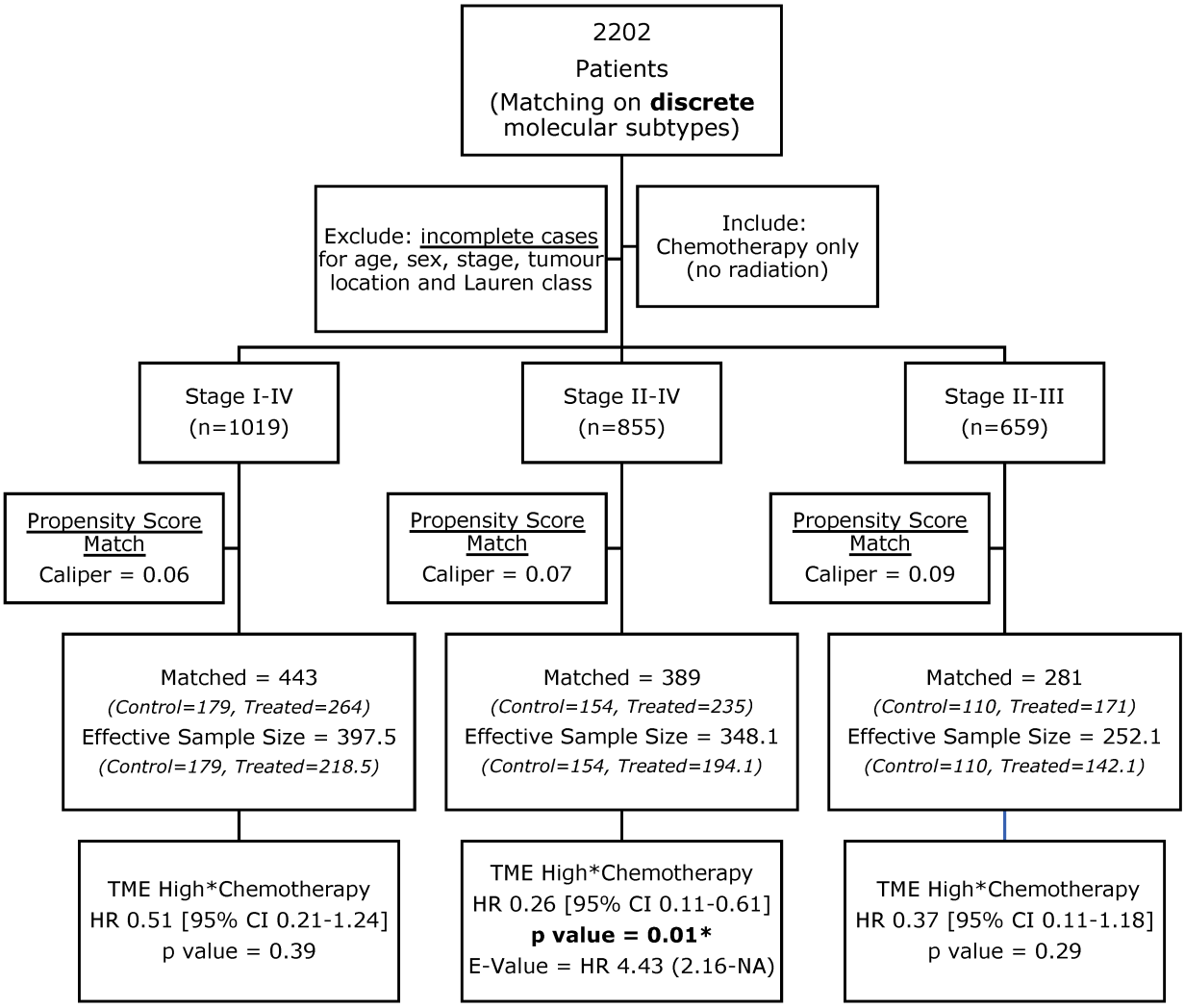


Figure D2-16 Overview of sensitivity analysis for propensity score matched cohort survival models using discrete interpretations of classification.

3. Critical Differences Analysis

We performed a Friedman test following by Nemenyi post-hoc test to calculate the critical differences. For TCGA predictions, GBM, NB and KNN models with SVM features and NB and KNN models with GBM features performed significantly worse relative to all other models. For ACRG predictions, NB, KNN and NNET models (GBM features), KNN and NNET models (LASSO features) and KNN and NB models (SVM features) were critically worse than all other combinations. Feature selection using GBM and LASSO models in combination with RF, GBM, KNN and NNET comprised the majority of best performing models in TME classification.

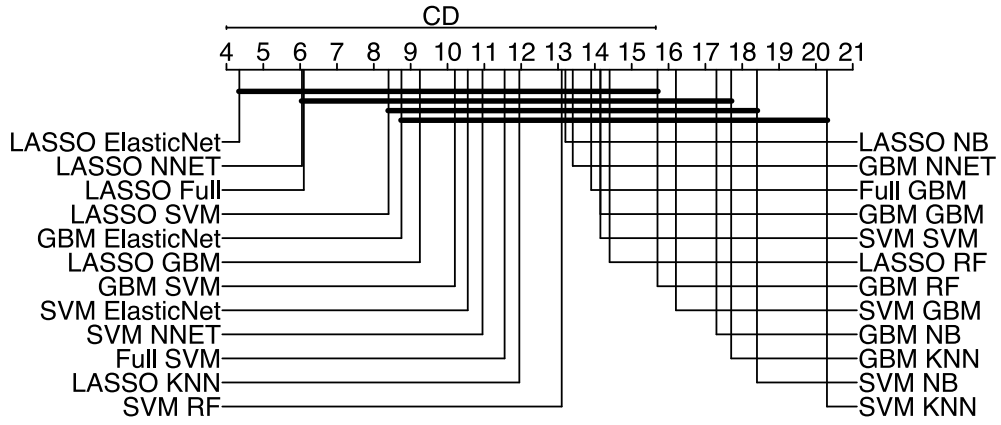


Figure D3-17 TCGA Models Critical Difference Plot. Models are ranked from most accurate (top, left column) to least accurate (bottom, right column). Statistically similar models are connected by a black line. The Critical Difference of top ranked models is identified by the line below CD.

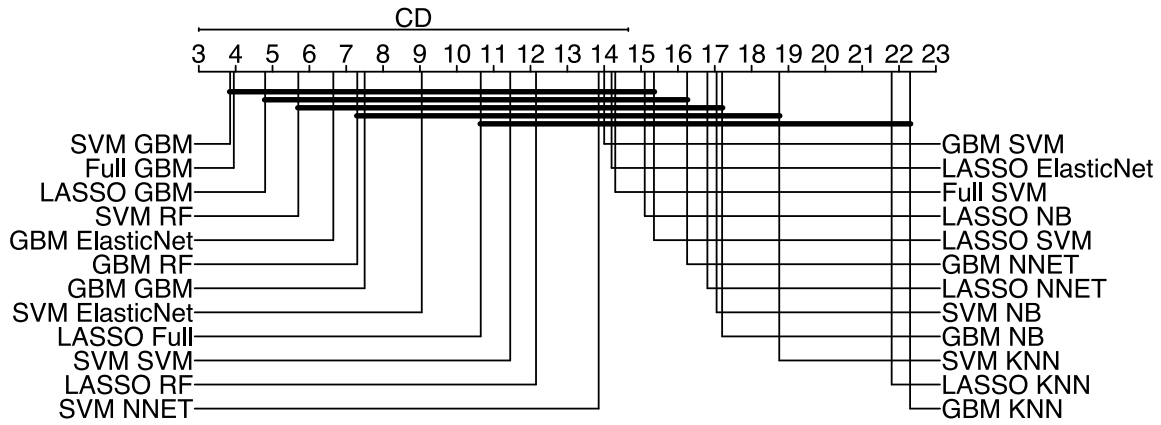


Figure D3-18 ACRG Models Critical Difference Plot. Models are ranked from most accurate (top, left column) to least accurate (bottom, right column). Statistically similar models are connected by a black line. The Critical Difference of top ranked models is identified by the line below CD

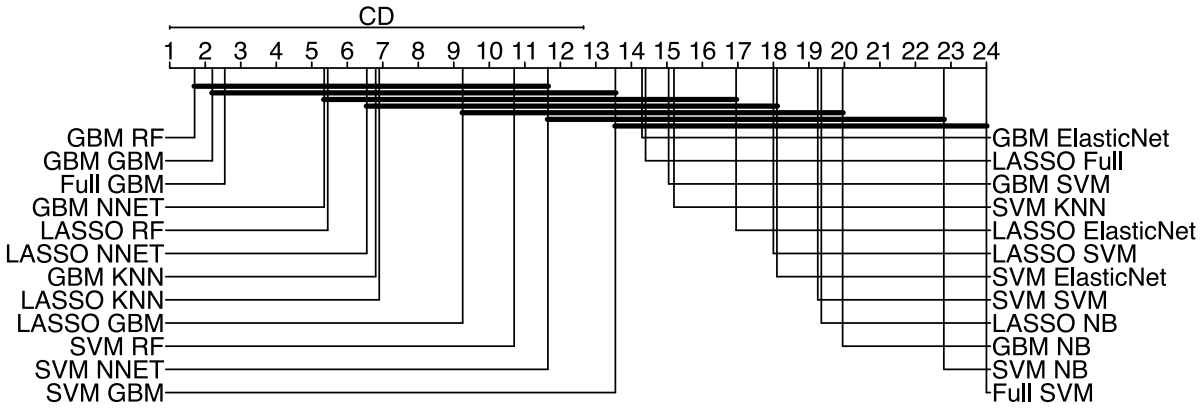


Figure D3-19 TME Models Critical Difference Plot. Models are ranked from most accurate (top, left column) to least accurate (bottom, right column). Statistically similar models are connected by a black line. The Critical Difference of top ranked models is identified by the line below CD

4. Calibration plots

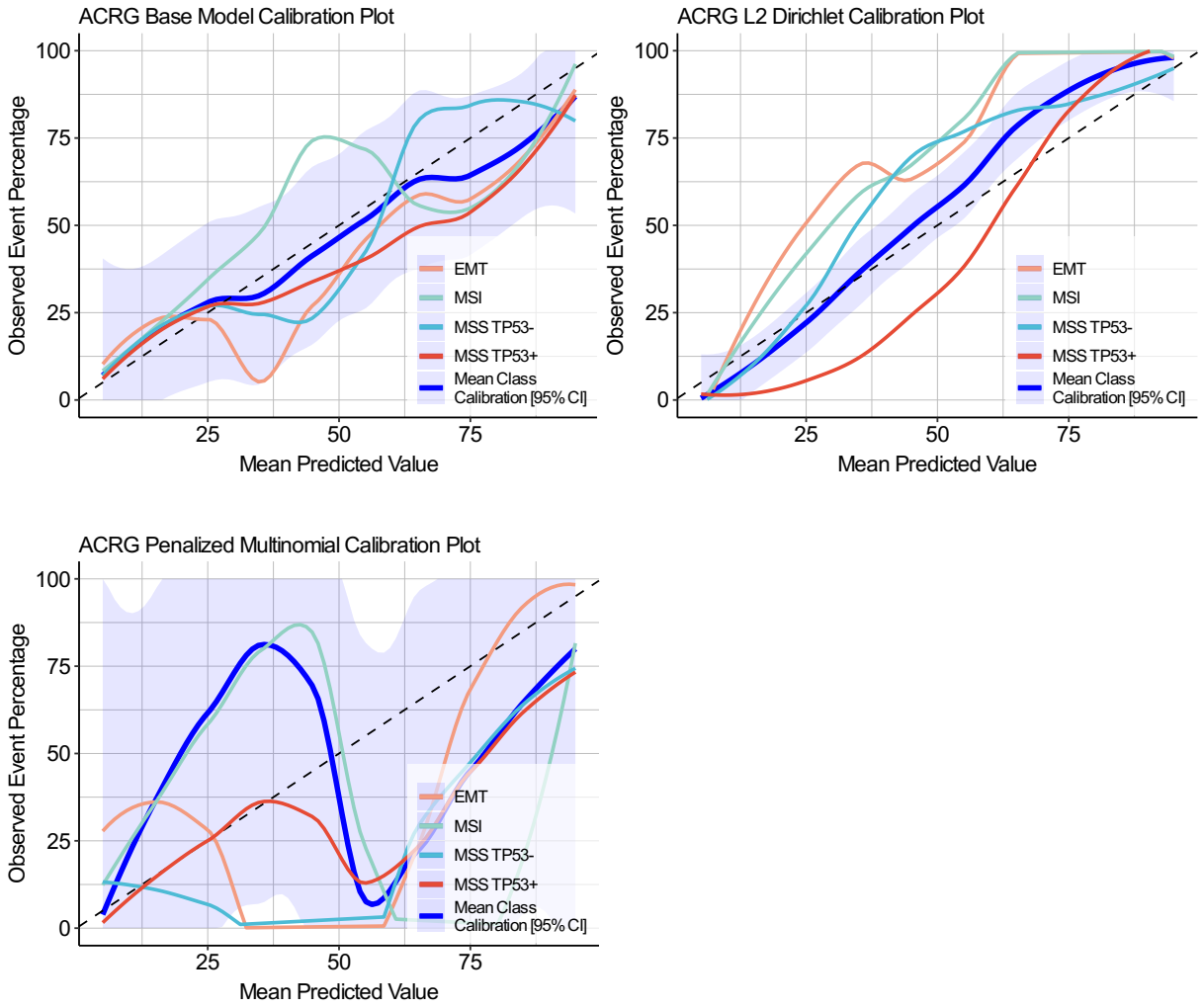


Figure D4-20 Calibration plots for ACRG uncalibrated, L2 Dirichlet and Penalized Multinomial models. Calibration was assessed using deciles and visualized using a loess smooth with span=0.75. The mean calibration for all subtypes is represented in blue with 95% confidence intervals in light blue.

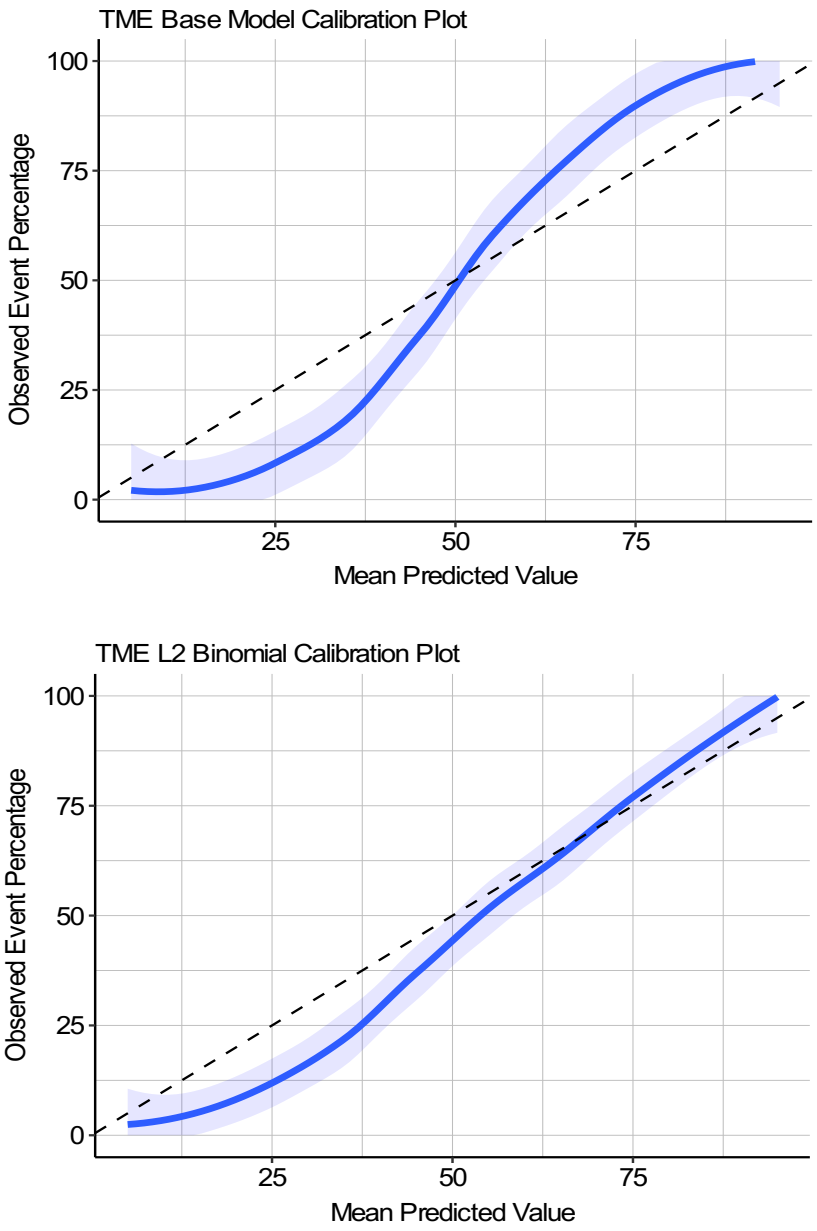


Figure D4-21 Calibration plot for TME uncalibrated and L2 Binomial (modified Platt) models. Calibration was assessed using deciles and visualized using a loess smooth with span=0.75. The mean calibration for all subtypes is represented in blue with 95% confidence intervals in light blue.

5. Detailed demographic data for each molecular subtype

Table D5-2: Demographic data for TCGA subtypes

Variable	**CIN**, N = 1,631[1]	**EBV**, N = 124[1]	**GS**, N = 189[1]	**TCGA MSI**, N = 258[1]	p-value[2]
Age	61 (51, 68)	61 (52, 67)	59 (50, 68)	68 (60, 74)	<0.001
Stage					<0.001
I	188 (15%)	11 (12%)	15 (9.9%)	44 (23%)	
II	361 (29%)	21 (24%)	40 (26%)	62 (33%)	
III	424 (34%)	43 (49%)	62 (41%)	61 (32%)	
IV	276 (22%)	13 (15%)	34 (23%)	23 (12%)	
Sex					<0.001
Female	495 (31%)	21 (17%)	81 (44%)	117 (47%)	
Male	1,084 (69%)	100 (83%)	103 (56%)	131 (53%)	
Study					<0.001
ACRG	190 (12%)	18 (15%)	43 (23%)	49 (19%)	
Kosin	97 (5.9%)	5 (4.0%)	2 (1.1%)	5 (1.9%)	
KUGH	70 (4.3%)	8 (6.5%)	5 (2.6%)	10 (3.9%)	
MDACC	32 (2.0%)	3 (2.4%)	3 (1.6%)	2 (0.8%)	
Samsung	395 (24%)	6 (4.8%)	6 (3.2%)	25 (9.7%)	
Shanghai	52 (3.2%)	3 (2.4%)	5 (2.6%)	10 (3.9%)	
Singapore	176 (11%)	12 (9.7%)	39 (21%)	21 (8.1%)	
TCGA	254 (16%)	31 (25%)	49 (26%)	78 (30%)	
Yonsei	319 (20%)	33 (27%)	31 (16%)	50 (19%)	
Yonsei MDACC	46 (2.8%)	5 (4.0%)	6 (3.2%)	8 (3.1%)	
Grade					0.010
G1	21 (6.4%)	3 (7.7%)	2 (3.5%)	6 (6.8%)	
G2	123 (37%)	4 (10%)	14 (25%)	25 (28%)	
G3	164 (50%)	31 (79%)	39 (68%)	52 (59%)	
Other	21 (6.4%)	1 (2.6%)	2 (3.5%)	5 (5.7%)	
Lauren Class					<0.001
Intestinal	589 (53%)	35 (49%)	38 (29%)	98 (62%)	
Diffuse	470 (42%)	32 (44%)	88 (67%)	52 (33%)	
Mixed	52 (4.7%)	5 (6.9%)	5 (3.8%)	7 (4.5%)	
Signet Ring					<0.001
Signet Ring	44 (14%)	2 (5.4%)	27 (39%)	10 (11%)	
No	239 (76%)	31 (84%)	37 (54%)	74 (84%)	
Other	33 (10%)	4 (11%)	5 (7.2%)	4 (4.5%)	
Tumour Location					<0.001
Proximal	526 (51%)	55 (76%)	51 (47%)	60 (36%)	
Distal	495 (48%)	16 (22%)	55 (50%)	107 (64%)	
Whole	19 (1.8%)	1 (1.4%)	3 (2.8%)	1 (0.6%)	
Chemotherapy	753 (73%)	50 (69%)	61 (60%)	93 (57%)	<0.001
Radiation	97 (23%)	13 (27%)	21 (26%)	21 (19%)	0.6

[1]Median (IQR); n (%)

[2]Kruskal-Wallis rank sum test; Pearson's Chi-squared test; Fisher's Exact Test for Count Data with simulated p-value (based on 2000 replicates)

Table D5-3: Demographic data for ACRG Subtypes

Variable	**EMT**, N = 278[1]	**MSI**, N = 372[1]	**MSS TP53-**, N = 850[1]	**MSS TP53+**, N = 702[1]	p-value[2]
Age	58 (49, 66)	64 (55, 72)	62 (53, 70)	60 (50, 68)	<0.001
Stage					0.016
I	26 (12%)	56 (20%)	108 (17%)	68 (13%)	
II	49 (23%)	88 (32%)	187 (29%)	160 (30%)	
III	88 (41%)	88 (32%)	217 (33%)	197 (37%)	
IV	51 (24%)	45 (16%)	140 (21%)	110 (21%)	
Sex					<0.001
Female	100 (37%)	150 (42%)	272 (33%)	192 (28%)	
Male	169 (63%)	210 (58%)	547 (67%)	492 (72%)	
Study					<0.001
ACRG	46 (17%)	68 (18%)	107 (13%)	79 (11%)	
Kosin	18 (6.5%)	12 (3.2%)	40 (4.7%)	39 (5.6%)	
KUGH	13 (4.7%)	13 (3.5%)	38 (4.5%)	29 (4.1%)	
MDACC	5 (1.8%)	8 (2.2%)	15 (1.8%)	12 (1.7%)	
Samsung	21 (7.6%)	52 (14%)	182 (21%)	177 (25%)	
Shanghai	9 (3.2%)	12 (3.2%)	31 (3.6%)	18 (2.6%)	
Singapore	42 (15%)	46 (12%)	86 (10%)	74 (11%)	
TCGA	62 (22%)	77 (21%)	163 (19%)	110 (16%)	
Yonsei	54 (19%)	76 (20%)	162 (19%)	141 (20%)	
Yonsei MDACC	8 (2.9%)	8 (2.2%)	26 (3.1%)	23 (3.3%)	
Grade					<0.001
G1	6 (8.0%)	4 (4.3%)	7 (3.5%)	15 (10%)	
G2	9 (12%)	23 (25%)	82 (41%)	52 (36%)	
G3	56 (75%)	61 (66%)	100 (50%)	69 (48%)	
Other	4 (5.3%)	5 (5.4%)	13 (6.4%)	7 (4.9%)	
Lauren Class					<0.001
Intestinal	63 (34%)	138 (58%)	322 (56%)	237 (50%)	
Diffuse	115 (62%)	89 (37%)	228 (40%)	210 (44%)	
Mixed	7 (3.8%)	11 (4.6%)	22 (3.8%)	29 (6.1%)	
Signet Ring					<0.001
Signet Ring	34 (38%)	10 (11%)	26 (13%)	13 (9.6%)	
No	48 (53%)	74 (84%)	154 (79%)	105 (77%)	
Other	8 (8.9%)	4 (4.5%)	16 (8.2%)	18 (13%)	
Tumour Location					0.002
Proximal	79 (48%)	95 (42%)	272 (49%)	246 (55%)	
Distal	77 (47%)	130 (58%)	272 (49%)	194 (43%)	
Whole	7 (4.3%)	1 (0.4%)	6 (1.1%)	10 (2.2%)	
Chemotherapy	99 (60%)	139 (63%)	372 (69%)	347 (78%)	<0.001
Radiation	21 (21%)	27 (20%)	57 (23%)	47 (27%)	0.5

[1]Median (IQR); n (%)

[2]Kruskal-Wallis rank sum test; Pearson's Chi-squared test; Fisher's Exact Test for Count Data with simulated p-value (based on 2000 replicates)

Table D5-4: Demographic data for TME Subtypes

Variable	**High**, N = 731[1]	**Low**, N = 1,471[1]	p-value[2]
Age	60 (51, 68)	62 (53, 69)	0.008
Stage			<0.001
I	137 (21%)	121 (12%)	
II	214 (33%)	270 (26%)	
III	202 (31%)	388 (38%)	
IV	105 (16%)	241 (24%)	
Sex			0.4
Female	250 (35%)	464 (33%)	
Male	471 (65%)	947 (67%)	
Study			<0.001
ACRG	71 (9.7%)	229 (16%)	
Kosin	26 (3.6%)	83 (5.6%)	
KUGH	22 (3.0%)	71 (4.8%)	
MDACC	8 (1.1%)	32 (2.2%)	
Samsung	334 (46%)	98 (6.7%)	
Shanghai	10 (1.4%)	60 (4.1%)	
Singapore	88 (12%)	160 (11%)	
TCGA	99 (14%)	313 (21%)	
Yonsei	60 (8.2%)	373 (25%)	
Yonsei MDACC	13 (1.8%)	52 (3.5%)	
Grade			0.7
G1	9 (7.5%)	23 (5.9%)	
G2	35 (29%)	131 (33%)	
G3	68 (57%)	218 (55%)	
Other	8 (6.7%)	21 (5.3%)	
Lauren Class			0.8
Intestinal	322 (53%)	438 (51%)	
Diffuse	261 (43%)	381 (44%)	
Mixed	29 (4.7%)	40 (4.7%)	
Signet Ring			0.007
Signet Ring	9 (8.0%)	74 (19%)	
No	96 (85%)	285 (72%)	
Other	8 (7.1%)	38 (9.6%)	
Tumour Location			0.3
Proximal	270 (48%)	422 (51%)	
Distal	274 (49%)	399 (48%)	
Whole	13 (2.3%)	11 (1.3%)	
Chemotherapy	454 (81%)	503 (62%)	<0.001
Radiation	40 (25%)	112 (23%)	0.5

[1]Median (IQR); n (%)

[2]Wilcoxon rank sum test; Pearson's Chi-squared test; Fisher's Exact Test for Count Data with simulated p-value (based on 2000 replicates)

6. Intratumour heterogeneity analysis

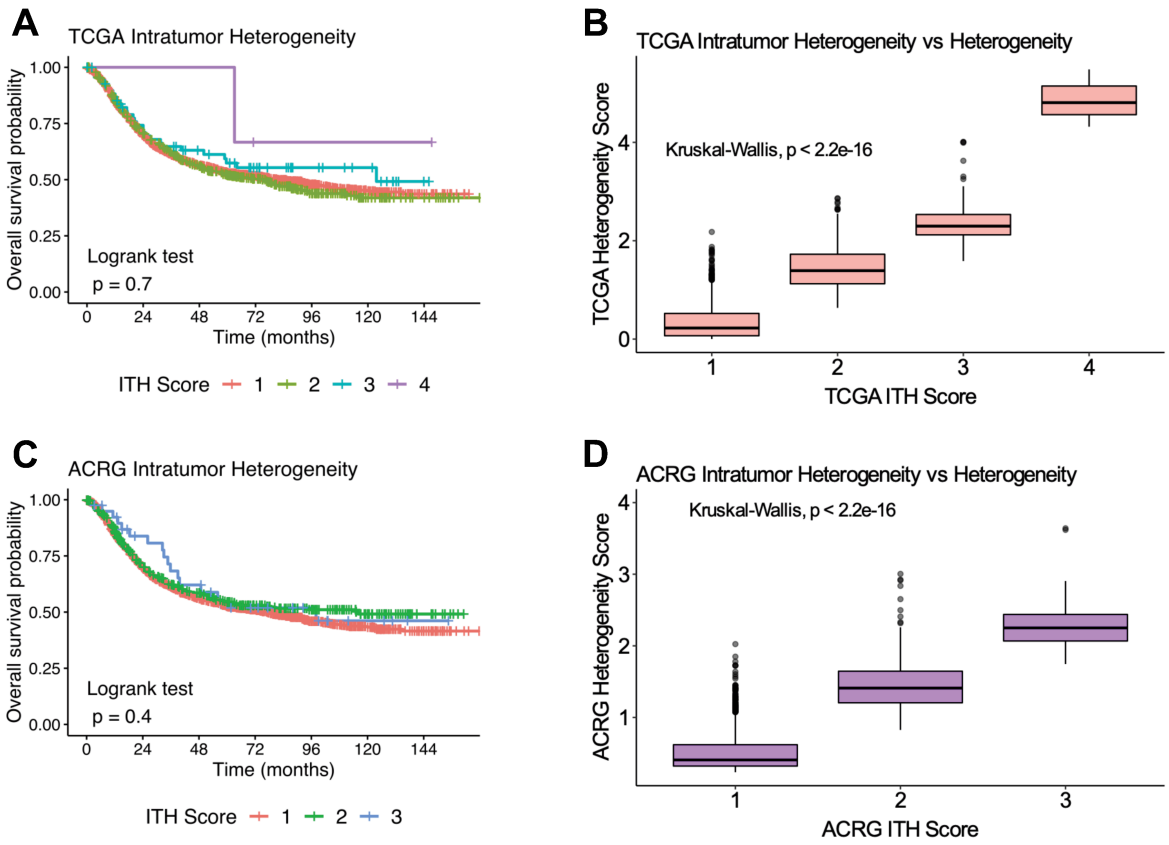


Figure D6-22 Analysis of intratumour heterogeneity score (ITH Score) from Marisa et al. versus variance heterogeneity. **A.** Kaplan Meier curves for ITH score in TCGA subtypes. **B.** Boxplots of TCGA sc-varHet versus ITH score. **C.** Kaplan Meier curves for ITH score in ACRG subtypes **D.** Boxplots of ACRG sc-varHet versus ITH score.

7. Survival Analysis

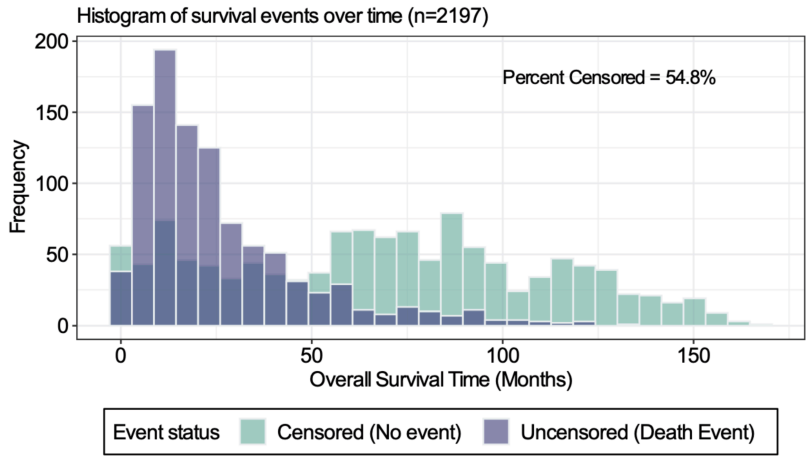


Figure D7-23 Histogram of survival events in publicly available gastric cancer patients.

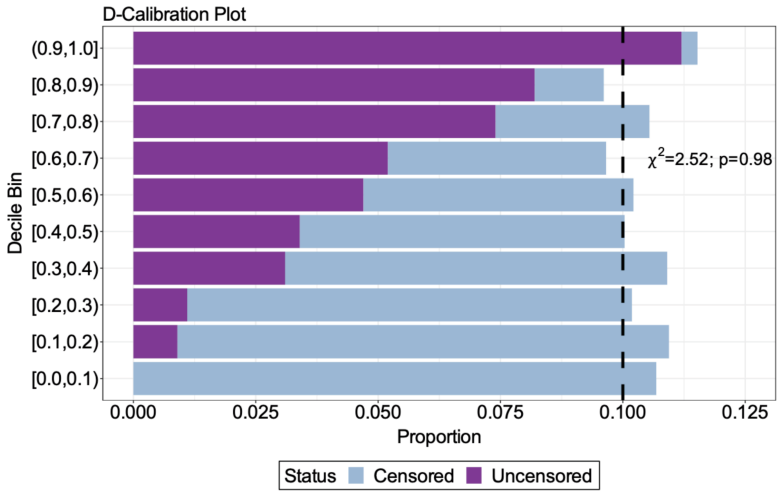


Figure D7-24 D-Calibration plot for MTLR model. The proportion of censored and uncensored patients approximates 0.1 in each decile of survival probability. The corresponding Chi-square test is annotated in plot.

E. Chapter 5 Appendix

1. Immunofluorescence and immunohistochemistry protocols

Protocol	Primary Antibody	Secondary antibody	Antigen Retrieval
1	Mouse Anti-pan cytokeratin, 1:25, overnight at 4 degrees Celsius, Abcam, ab7753	Goat anti-mouse IgG Alexa Fluor 568, 1:50, 30 minutes room temperature, Invitrogen, A-11004	Sodium Citrate
	Rabbit Anti-vimentin, 1:300, overnight at 4 degrees Celsius, Abcam, ab92547	Goat anti-rabbit IgG Alexa Fluor 488, 1:200, 30 minutes room temperature, Abcam, ab150077	Sodium Citrate
2	Mouse anti-MUC5AC, 1:250, overnight at 4 degrees Celsius, Invitrogen, MA5-12178	Biotinylated goat anti-mouse IgG, 1:200, 30 minutes room temperature, Jackson ImmunoResearch, 115-065-003	Sodium Citrate
3	Rabbit Anti-pepsinogen II/PGC, 1:100, overnight at 4 degrees Celsius, Abcam, ab255826	Biotinylated goat anti-rabbit IgG, 1:200, 30 minutes room temperature, Vector Laboratories, BA-1000	Tris-EDTA

2. Reagents table for human gastric cancer organoids

Reagent	Source	Catalog Number
Advanced DMEM/F12	Gibco	12634010
L-Glut	Sigma	G8540
HEPES	Fisher Scientific	BP310
Penicillin Streptomycin	Gibco	15140122
Primocin	InvivoGen	Ant-pm-2
N2 Supplement	Gibco	17502011
B27 Supplement	Gibco	17504044
N-acetylcysteine	Sigma	A9165
Gastrin	Sigma	G9145
Nicotinamide	Sigma	N0636
SB202190/p 38 inhibitor	Sigma	S7067
A83-01 (ALK4/5/7 inhibitor)	Sigma	SML0788
Human EGF	Gibco	PHG0313
Dispase II	Sigma	D4693
Collagenase IX	Sigma	C9407
Y-27632	Sigma	Y0503

3. Concordance of tumour and tumour organoid molecular subtypes

Concordance for TCGA Subtypes					
		Tumour Organoid			
		CIN	EBV	GS	MSI
Tumour	CIN	8	0	0	0
	EBV	1	0	0	0
	GS	1	0	0	0
	MSI	1	0	0	1

Concordance for TME Subtypes			
		Tumour Organoid	
		High	Low
Tumour	High	0	1
	Low	1	10

4. Dose response assay validation

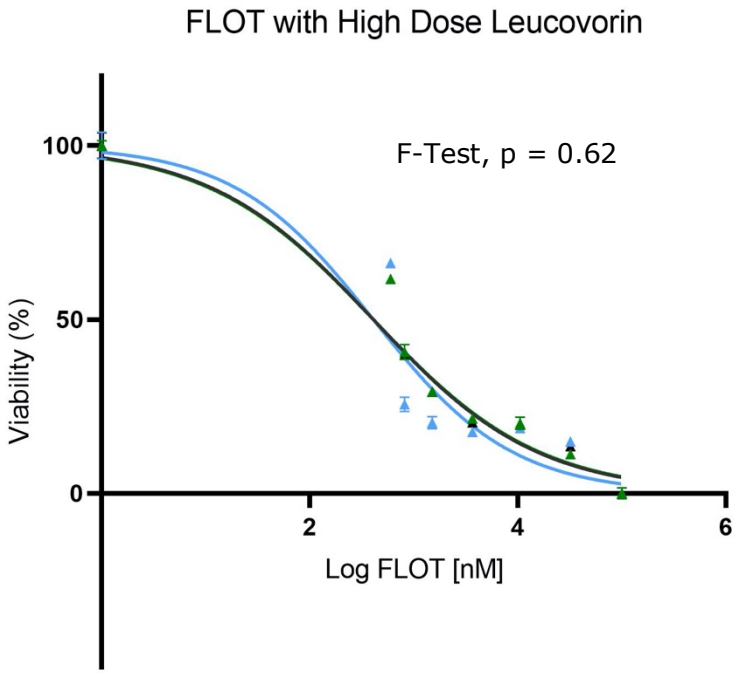


Figure E4-25 FLOT in-vitro drug assay validation with AGS cell line. Three independent dose-response assays were performed. Cell viability was assessed using a CCK-8 assay. Variable slope least squares linear regression was used to calculate dose-response curves. A sum-of-squares F-Test demonstrated that the was best characterized by a single dose-response curve.