Novel Biomarkers for the Diagnosis of Pancreatic Ductal Adenocarcinoma: The Proteomic Approach

J.M.F.TANG

M.Phil **August 2010**

Novel Biomarkers for the Diagnosis of Pancreatic Ductal Adenocarcinoma: The Proteomic Approach

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Master in Philosophy by Joseph Man Fung Tang

August 2010

Preface

I hereby certify that this thesis is the result of work undertaken by myself with supervision and guidance as acknowledged. The material contained within the thesis has not been presented wholly or in part for any other degree or qualification.

The research was undertaken in the Division of Surgery and Oncology, University of Liverpool, 5th Floor UCD Building, Daulby Street, Liverpool, L69 3GA.

Cell culture, Western Blot analysis, ELISA, and all statistical analyses were undertaken by myself. Quantification of serum cytokines, chemokines, and growth factors was performed with the aid and under the direct supervision of Dr Victoria Shaw in accordance with the Good Clinical Laboratory Practice regulations.

Figure 1.1 was used (unmodified) with the kind permission from Encyclopaedia Britannica, Inc., copyright 2003.

Joseph Man Fung Tang

August 2010

Table of Contents

ACKNOWLEDGEMENTS ..... **VIII**

ABBREVIATIONS ..... **IX**

CHAPTER 1- GENERAL INTRODUCTION: PANCREAS AND PANCREATIC CANCER **1**

1.1 ANATOMY OF THE PANCREAS 2

1.2 INTRODUCTION TO PANCREATIC CANCER 3

1.2.1 Epidemiology of Pancreatic Cancer *3*

1.2.2 Pathogenesis of Pancreatic Cancer *5*

1.2.3 Molecular Hallmarks of Pancreatic Cancer *7*

1.2.3.1 K-Ras signalling pathway 7

1.2.3.2 Tumour suppressor genes and pathways 8

1.2.3.3 Embryonic Signalling pathways 8

1.2.4 Symptoms and Signs of Pancreatic Cancer *9*

1.2.5 Diagnosis and Staging of Pancreatic Cancer *10*

1.2.6 Management and prognosis of pancreatic cancer *11*

1.2.6.1 Advanced pancreatic cancer 11

1.2.6.2 Resectable pancreatic cancer 12

1.3 BIOMARKERS FOR PANCREATIC CANCER 13

1.3.1 Introduction to biomarkers *13*

1.3.1.1 The need for diagnostic biomarkers of pancreatic cancer 13

1.3.2 Current biomarker of pancreatic cancer: CA19-9 ..... *14*

1.3.2.1 CA19-9 in screening and diagnosis of pancreatic cancer 14

1.3.2.2 CA19-9 in prognosis, surveillance, and assessment of chemotherapy 15

1.3.3 Other markers of pancreatic cancer *15*

1.3.4 Techniques for biomarker discovery *16*

1.3.4.1 2D-PAGE 16

1.3.4.2 Mass Spectrometry and iTRAQ 17

1.3.4.3 Western blotting 18

1.3.4.4 Enzyme-Linked Immunosorbent Assay (ELISA) 20

1.3.4.5 Multiplex Assays (LUMINEX) 21

1.4 ROLE OF CCGFS IN INFLAMMATION AND CANCER 23

1.4.1 The relationship between inflammation and cancer *23*

1.4.2 Pathways linking inflammation and cancer *24*

1.4.2.1 NF- 24

1.4.2.2 STAT3 24

1.4.2.3 HIF 1-alpha 25
1.4.3 Tumour microenvironment, cytokines and cancer *26*
1.4.3.1 Immune cells, cytokines, and cancer 27
1.4.3.2 Pancreatic stellate cells, CCGFs, pancreatic fibrosis, and cancer 27
1.5 STATISTICAL MODELLING METHODS 29
1.5.1 Background *29*
1.5.1.1 The Stepwise Regression (SR) Model 29
1.5.1.2 The Multinomial Logistic Regression Model 30
1.5.1.3 Artificial Neural Network Model 31
1.6 RESEARCH PROJECT DESIGN 32

CHAPTER 2- THE DIAGNOSTIC POTENTIAL OF VDBP, RBP-4, AND FINC FOR PANCREATIC CANCER 33

2.1 BACKGROUND AND INTRODUCTION 34
2.1.1 Vitamin D binding protein *35*
2.1.2 Retinol-binding protein *36*
2.1.3 Fibronectin *36*
2.2 STUDY AIMS 37
2.3 MATERIALS AND METHODS 37
2.3.1 Study design *37*
2.3.2 Patients and samples *39*
2.3.2.1 Phase I and II Liverpool pancreatic cancer database (LPCD) 39
2.3.2.2 Pre-diagnostic serum samples- UKCTOCS 41
2.3.3 Cancer cell lines, cell culture and lysate preparation *42*
2.3.3.1 Cell lines and cell culture 42
2.3.3.2 Lysate preparation 43
2.3.3.3 Measuring lysate protein concentration- Bradford Assay 43
2.3.4 Western Blot Analysis *43*
2.3.4.1 SDS-polyacrylamide gel preparation 43
2.3.4.2 Sample preparation 44
2.3.4.3 SDS-polyacrylamide gel electrophoresis (SDS-PAGE) 44
2.3.5 Relative quantification of western blots *45*
2.3.6 Statistical Analysis *46*
2.3.6.1 General statistics, univariate, and multivariate analyses 46
2.3.6.2 Correlation analysis 46
2.3.6.3 Diagnostic potential of biomarkers for PDAC 47
2.3.6.4 Software for statistical analyses 47
2.4 RESULTS 48
2.4.1 Phase I- Validation of iTRAQ results by Western Blotting *48*

2.4.2 Phase II- Further Validation by Western Blot (Liverpool samples) 50
2.4.2.1 Patient demographics and clinical characteristics 50
2.4.2.2 Western Blot images for VDBP, RBP-4, and FINC 51
2.4.2.3 The serum level of VDBP in PDAC and Controls 52
2.4.2.4 The serum levels of RBP-4 in PDAC and Controls 53
2.4.2.5 The serum level of FINC in PDAC and Controls 54
2.4.2.6 Accuracy of candidate markers for the diagnosis of PDAC versus HC 55
2.4.2.7 Accuracy of candidate markers for the diagnosis of PDAC against CP 56
2.4.2.8 Accuracy of candidate markers for the diagnosis of PDAC against DC 57
2.4.2.9 Diagnostic accuracy of PDAC against all controls 58
2.4.2.10 Accuracy of candidate markers for the diagnosis of PDAC against HC and CP combined 59
2.4.2.11 Correlation between age and candidate protein markers 61
2.4.3 Phase III- Validation with pre-diagnostic samples (UKCTOCS) 62
2.4.3.1 Characteristics of the UKCTOCS samples 62
2.4.3.2 Example of Western Blot images for VDBP and RBP-4 63
2.4.3.3 Selection of markers for validation in Phase III 64
2.4.3.4 The level of VDBP in pre-diagnosis samples compared to Controls 64
2.4.3.5 Change in the relative serum level of VDBP through time 65
2.4.3.6 The level of RBP-4 in pre-diagnosis samples compared to Controls 66
2.4.3.7 Change in the relative serum level of RBP-4 through time 67
2.4.4 Expression of VDBP, RBP-4, and FINC in cell lines 68
2.4.4.1 Quantification of cell lysate 68
2.4.4.2 Expression of VDBP, RBP-4 and FINC in cancer cell lines 69
2.5 DISCUSSION 70
2.5.1 VDBP, RBP-4, and FINC for the diagnosis of pancreatic cancer 70
2.5.2 VDBP and RBP-4 as screening modalities for pancreatic cancer 71
2.5.3 The roles of VDBP, RPB-4, and FINC in pancreatic cancer 72
2.5.3.1 Vitamin D-Binding protein and cancer 72
2.5.3.2 Retinol-Binding protein and cancer 73
2.5.3.3 Fibronectin and cancer 73
CHAPTER 3- THE DIAGNOSTIC POTENTIAL OF CCGFS FOR PANCREATIC CANCER 75
3.1 INTRODUCTION 76
3.2 STUDY AIMS 77
3.3 MATERIALS AND METHODS 77
3.3.1 Patients and Samples- The Liverpool Pancreatic Cancer Database 77
3.3.1.1 Discovery Phase patient demographics and sample characteristics 79

3.3.1.2 Validation Phase patient demographics and sample characteristics 80
3.3.2 Study design *81*
3.3.3 Quantification of Cytokines, Chemokines, and Growth Factors *83*
3.3.3.1 Sample and Standard Preparation 83
3.3.3.2 Coupled magnetic beads, detection antibody, and Streptavidin-PE preparation 84
3.3.3.3 Assay Procedure 84
3.3.4 Quantification of serum CA19-9 by ELISA *85*
3.3.5 Statistical Analysis *86*
3.3.5.1 General statistics, univariate, and multivariate analyses 86
3.3.5.2 Correlation analysis 86
3.3.5.3 Diagnostic accuracies of CCGFs for PDAC 87
3.3.5.4 Selection of candidate markers 87
3.3.5.5 Generating disease-predicting mathematical algorithms: M-LR and NN 87
3.3.5.6 Diagnostic accuracies of the models 88
3.3.5.7 Software for statistical analyses 88
3.4 RESULTS 89
3.4.1 Basic analysis- CCGFs as individual markers of PDAC *89*
3.4.1.1 Diagnostic accuracy of individual CCGFs for PDAC against HC 89
3.4.1.2 Diagnostic accuracy of individual CCGFs for PDAC against CP 91
3.4.1.3 Diagnostic accuracy of individual CCGFs for PDAC against biliary obstruction (DC) 93
3.4.1.4 Diagnostic accuracy of individual CCGFs for PDAC against all Controls 95
3.4.1.5 Correlation studies to determine the relationship between CCGFs 97
3.4.2 Discovery Phase- Diagnostic potential of CCGFs in combination *99*
3.4.2.1 Selection of CCGFs for combination 99
3.4.2.2 Diagnostic accuracy of the combined CCGF marker 102
3.4.3 Validation Phase- Validation of the disease-predicting algorithms *104*
3.4.3.1 The diagnostic accuracy of the prediction models in the Validation Phase 104
3.4.4 The diagnostic accuracy of CA19-9 compared to CCGFs *106*
3.4.4.1 Quantification of CA19-9 106
3.4.4.2 The diagnostic accuracy of CA19-9 compared to the NN-CCGF algorithm 107
3.4.5 CA19-9 in combination with CCGFs *108*
3.4.5.1 Discovery Phase- accuracy of the combined CCGF-CA19-9 marker 109
3.4.5.2 Validation Phase- validation of the CCGF-CA19-9 algorithms 111
3.4.6 The impact of each biomarker on the M-LR diagnostic algorithm *113*
3.5 IMPACT OF CLINICAL-DEMOGRAPHICAL FACTORS ON THE ACCURACY OF THE DISEASE-PREDICTING MODELS 114
3.5.1 Impact of Patient Age on the serum levels of candidate CCGFs *114*
3.5.2 Impact of diabetes on the serum levels of candidate CCGFs *116*

3.5.3 Impact of smoking on the serum levels of candidate CCGFs *117*
3.6 DISCUSSION 118
3.6.1 CCGFs for the diagnosis of pancreatic cancer *119*
3.6.2 The roles of Il-4, IL-17, G-CSF, and IP-10 in PDAC and pancreatic inflammatory diseases *120*
3.6.2.1 Interleukin 4 120
3.6.2.2 Interleukin 17 120
3.6.2.3 G-CSF 121
3.6.2.4 IP-10 122
CHAPTER 4- FINAL DISCUSSION, LIMITATIONS, AND FUTURE DIRECTIONS 123
4.1 FINAL DISCUSSION AND CONCLUSIONS 124
4.1.1 Conclusions from the current study *126*
4.1.1.1 Validation of iTRAQ results 126
4.1.1.2 Discovery and validation of CCGF markers 127
4.1.2 General Limitations *129*
4.1.2.1 Limitation of sample size 130
4.2 FUTURE DIRECTIONS 131
CHAPTER 5- REFERENCES 132
REFERENCES 133

Acknowledgements

The current research project was made possible by the patience and guidance from all members of the biomarker research team at the Division of Surgery and Oncology, University of Liverpool. I would like to thank all who helped with the planning and implementation of this work.

I wish to thank Dr Eithne Costello, my primary supervisor, for offering me the opportunity to undertake this research project and for her guidance throughout this year. Furthermore, I am grateful to Dr William Greenhalf, my secondary supervisor, and Brain Lane, the departmental bioinformatician, for their help with the statistical analyses performed in this study.

I would especially like to thank Dr Victoria Shaw and Dr Claire Jenkinson for teaching and supervising me in the various techniques used in this study as well as for their help and support in the preparation of this thesis. I am very grateful for Seonaid Murray (PhD student) for sharing her iTRAQ/MS data with me, which formed the foundations of my research study in chapter 2.

Finally, I would like to thank the Wolfson's Foundation for their financial support, which enabled me to undertake this MPhil degree.

Abbreviations

2 Dimension polyacrylamide gel electrophoresis	2D PAGE
5-Fluorouracil	5-FU
Artificial Neural Network	NN
Basic Fibroblast Growth Factor	FGF-b
Carbohydrate Antigen 19-9	CA19-9
Central Office of Research Ethics Committee	COREC
Chronica Pancreatitis	CP
Contrast-Enhanced Computed Tomography	CE-CT
Control group	CP, DC, and HC
Cytokines, Chemokines, and Growth Factors	CCGFs
Cytotoxic T lymphocyte	CTL
Difference Gel Electrophoresis	DIGE
Disease Control/ Biliary Obstruction	DC
Dithiothreitol	DTT
Electro-spray Ionisation Tandem Mass Spectrometry	ESI-MS
Enchanced Chemiluminescence	ECL
Endoscopic Ultrasonography	EUS
Enzyme Immunoassay	EIA
Enzyme-Linked Immunosorbent Assay	ELISA
Epidermal Growth Factor Receptor	EGFR
European Study Group for Pancreatic Cancer	ESPAC
Extracellular Matrix	ECM
Fibronectin	FINC
Fine needle biopsy	FNB
Granulocyte Colony-Stimulating Factor	G-CSF
Granulocyte-Macrophage Colony-Stimulating Factor	GM-CSF
Guanosine triphosphate	GTP
Healthy Control	HC
Horseradish Peroxidase	HRP
Human Embryonic Kideney	HEK
Hypoxia Inducible Factor	HIF

Immobilized pH gradients	IPGs
Inteferon Gamma Inducible Protein-10	IP-10
Inter Cellular Adhesion Molecule	ICAM
Inter-alpha-trypsin-inhibitor heavy chain 4	ITIH4
Interferon Gamma	IFN-γ
Interleukin	IL
Inter-Quartile Range	IQR
Intro-ductal papillary mucinous neoplasm	IPMN
Isobaric Tag for Relative and Absolute Quantification	iTRAQ
Isoelectric Focusing	IEF
Janus Kinase	JAK
KF-kB inhibitors	IkB
Kruskal Wallis	K-W
Liverpool Pancreatic Cancer Database	LPCD
Macrophage Inflammatory Protein-1	MCP-1
Macrophage inhibitory cytokine 1	MIC-1
Matrix Metallopeptidase 9	MMP-9
Monocyte Chemotatic Protein-1	MCP-1
Mucin 1	MUC-1
Mucinous cystic neoplasm	MCN
Multicentre Research Ethics Committee	MREC
Multinomail Logistic Regression	M-LR
National Academy of Clinical Biochemistry	NACB
Nuclear Factor Kappa light chain enhancer of activated B Cells	NF-κB
Pancreatic Ductal Adenocarcinoma	PDAC
Pancreatic intraepithelial neoplasm	PanIN
Pancreatic stellate cells	PSC
Phosphate Buffer Solution- Tween 20	PBST
Platelet Derived Growth Facotr Receptor	PDGFR
Polyacrylamide Gel Elecrophoresis	PAGE
Pre-Pancreatic Cancer	PPC
Pylorus-preserving partial pancreaticoduodenectomy	PP-PPD
Receiver Operator Characteristics	ROC

Receiver Operator Characteristics Area Under Curve	ROC-AUC
Regulated Upon Activation Normal T-Cell Expressed and Secreted	RANTES
Retinol-Binding Protein	RBP
Signal Transducer and Activator of Transcription	STAT
Smooth muscle actin	SMA
Sodium Dodecyl Sulphate	SDS
Stepwise Regression	SR
T-Helper cells	Th
Tissue inhibitors of metalloproteinases	TIMPs
Transabdominal ultrasound scan	USS
Transforming Growth Factor	TGF
Tumour Necrosis Factor	TNF
Tumour, Node, and Metastasis	TNM
Tumour-associated macrophages	TAMs
UK Collaborative Trial of Ovarian Cancer Screening	UKCTOCS
Vitamin D-Binding Protein	VDBP/DBP
Vitamin D-Binding Protein Macrophage Activating Factor	DBP-MAF

Abstract

Novel Markers for the Diagnosis of Pancreatic Ductal Adenocarcinoma: The Proteomic Approach. Joseph Man Fung Tang

Background: Pancreatic ductal adenocarcinoma (PDAC) is a disease of late presentation where the majority of patients present with non-specific symptoms and advanced disease. Current guidelines recommend that patients presenting with symptoms of suggestive of PDAC should be investigated by Contrast-Enhanced Computed Tomography (CE-CT). However, the radiographic features are often similar to benign diseases such as chronic pancreatitis (CP). Evidently, there is a need for a novel diagnostic biomarker, which can accurately identify patients with PDAC thereby reducing the number of otherwise unnecessary invasive procedures.

Aim: The current thesis aimed to determine the potential of a number of serum proteins as diagnostic markers of PDAC.

Method: Two approaches for the discovery and validation of diagnostic markers of PDAC were employed. In Chapter 2, the serum expression of three iTRAQ- Mass Spectrometry identified proteins (vitamin d-binding protein [VDBP], retinol-binding protein 4 [RBP-4], and fibronectin [FINC]) were validated by western blotting in a three-phased study consisting of 20, 60, and 120 serum samples. Their diagnostic potentials as individual and combined markers were assessed statistically. In Chapter 3, the serum concentrations of 27 cytokines, chemokines, and growth factors (CCGFs) in 90 PDAC and 90 controls were quantified using the multiplex cytokines assay and the potential of individual CCGFs for the diagnosis of PDACs were assessed. One-hundred and twenty serum samples were randomly allocated to discovery where stepwise regression was used to select independent CCGF markers of PDAC. These were then combined into a single marker and the diagnostic accuracy for PDAC assessed. Finally, validation utilised the remaining sixty samples to investigate the accuracy of the combined CCGF marker for the diagnosis of PDAC.

Results: Results from Chapter 2 showed that the serum concentrations of VDBP, RBP-4, and FINC were significant decreased in PDAC with ROC-AUCs of >0.74 against CP and healthy volunteers (HC). However, their diagnostic accuracies were decreased (ROC-AUC <0.63) in the presence of individuals with biliary obstruction (disease controls, DC). Combining all three markers increase the diagnostic accuracy for PDAC against HC and CP (ROC-AUC, 0.91) but not against DC (ROC-AUC, 0.74). Further validation using pre-diagnostic serum samples showed that a small subset of patients exhibited a gradual decline in the serum concentration of VDBP and RBP-4 closer to diagnosis.
Results from Chapter 3 showed that fourteen CCGFs were differentially expressed in PDAC compared to controls, of which, IFN-γ was the most significant individual marker of PDAC with comparable accuracy to CA19-9. Discovery analysis identified four independent markers of PDAC: IL-4, IL-17, G-CSF, and IP-10. When combined, an ROC-AUC of 0.99 was achieved. Validation of the combined CCGF marker in yielded encouraging results of ROC-AUC >0.95.

Conclusion: Results indicate that combined VDBP, RBP-4, FINC, as well as IL-4, IL-17, G-CSF, and IP-10 are accurate markers of PDAC. It is possible that their use will improve the current diagnostic process.

Chapter 1

General Introduction:

Pancreas and Pancreatic Cancer

1.1 Anatomy of the pancreas

The pancreas is a retroperitoneal organ, which can be anatomically divided into the head, neck, body, and tail (***Figure 1.1***[1])[2-4]. The head of the pancreas lies within the curvature of the duodenum and overlies the body of the second lumbar vertebra and the aorta[3]. The neck represents a constriction, which connects the head to the body of the pancreas[3]. It can be identified by the superior mesenteric vessels, which pass over the uncinate process and then posteriorly behind the neck of the pancreas[4]. The tail of the pancreas extends towards the spleen and connects with the splenic flexure of the colon[4]. The pancreas receives its blood supply from the lineal and the pancreaticoduodenal branches of the hepatic and superior mesenteric arteries and it is drained by the lineal and superior mesenteric veins[2-4].

This text box is where the unabridged thesis included the following third party copyrighted material:

Encyclopaedia Britannica. Anatomy of the Pancreas. [Image] [2nd July, 2010]. Available at: http://www.britannica.com/EBchecked/topic/440971/pancreas.

Figure 1.1-Anatomy of the pancreas[1]. *By courtesy of Encyclopaedia Britannica, Inc., copyright 2003; used with permission.* ***The pancreas is a retroperitoneal organ that is situated in front of the third lumbar vertebrae and is in close proximity to a number of important organs and vessels including the liver, stomach, small intestines, spleen, bile duct, aorta, vena cava, and superior mesenteric artery and vein.***

Structurally, the pancreas is composed of lobules connected by areola tissue[3-4]. Each lobule consists of multiple acini each receiving one of the ultimate ramifications of the main pancreatic duct (***Figure 1.1***[1]). Histologically, two broad types of cells can be found in the pancreas parenchyma: exocrine and endocrine cells. Pancreatic exocrine cells are responsible for secreting digestive enzymes into the lumen of the acini, which in turn drains into the main pancreatic duct and ultimately into the duodenum[2-4]. Endocrine cells of the pancreas form clusters (islets of langerhans), which are embedded within the exocrine tissue and are responsible for secreting hormones such as insulin and glucagon into the systemic circulation[2-4].

1.2 Introduction to Pancreatic Cancer

1.2.1 Epidemiology of Pancreatic Cancer

Pancreatic cancer is the eleventh most commonly diagnosed cancer in the United Kingdom and it has an incidence rate of over 7,500 new cases per annum[5]. Furthermore, it is the sixth leading cause of cancer mortality in 2007 with a reported 5-year overall survival rate of less than 5%[6]. The lifetime risk of developing pancreatic cancer for both men and women is 1/86 with the majority of cases occurring in patients over 65 years of age. Interestingly, studies have observed that pancreatic cancer is more frequent in the black population compared to the Caucasian and Asian populations[7]. Whilst the reasons for this difference remain unclear, studies have identified a higher prevalence of risk factors such as smoking, diabetes, obesity, and vitamin D insufficiency in the black population, which may explain this observation[7-8].

In addition to demographical risk factors, a number of medical, genetic, and environmental factors have also been linked to an increased risk of pancreatic cancer (***Box 1.1***)[7]. Indeed several studies have associated benign medical conditions such as chronic pancreatitis, obesity, and diabetes with an increased risk of developing pancreatic cancer. Chronic pancreatitis is one of the most frequently reported risk factors for the development of pancreatic cancer[9-11] with various studies reporting a 10-fold increased risk of pancreatic cancer in individuals with chronic pancreatitis compared to healthy controls [10, 12].

Another frequently identified disease-related risk factor is obesity. Despite various studies reporting an increased risk of pancreatic cancer in obese individuals, the exact role of obesity in the development of pancreatic cancer remains unclear[13-15]. However, some studies have suggested that obesity may be indirectly linked to cancer via inflammatory responses[16].

The role of diabetes mellitus in pancreatic cancer is somewhat difficult to evaluate because although some studies demonstrated that patients with >10-year history of type 2 diabetes are 1.5 times more likely to develop pancreatic cancer[17-19], other studies indicated that new onset diabetes is an early symptom in up to a third of all pancreatic cancer patients[20].

Genetic predisposition plays an important role in the development of pancreatic cancer. Studies have reported that the relative risk of pancreatic cancer is increased by as much as 57-fold in families with four or more affected members[21]. Furthermore, a number of studies have reported that various germline diseases such as familial Peutz-Jeghers syndrome and hereditary pancreatitis are associated with a very high risk of pancreatic cancer development[7, 22]. Indeed, Giardiello *et al.* reported that individuals with familial Peutz-Jeghers syndrome are 132 times more likely to develop pancreatic cancer[23].

There are several environmental factors, which may increase the risk pancreatic cancer development. In particular, the association between pancreatic cancer and cigarette smoking has been frequently reported[7-8, 19, 21, 24]. Results from a study by Iodice *et al.*[8] indicated that the relative risk of pancreatic cancer in current smokers is approximately 1.7 times greater than non-smokers and this risk remains elevated for at least 10 years after cessation.

***Box 1.1**- Summary of Factors Associated with Pancreatic Cancer*[7]

- Pancreatic cancer typically occur in patients >65 years of age.
- The black population has an increased risk of pancreatic cancer compared to Caucasians and Asians.
- 20-25% of pancreatic cancers are attributable to cigarette smoking, which is the most frequent but also the most preventable risk factor for pancreatic cancer.
- Benign pancreatic diseases such as chronic pancreatitis and type II diabetes mellitus are independent risk factors for developing pancreatic cancer
- 5-10% of pancreatic cancers are associated with a germline disease
- Non-O blood type has been associated with an increased risk of pancreatic cancer[25].
- Several hallmark genetic mutations have been identified in pancreatic cancer including the KRAS2 oncogene, which is present in 90-95% of all pancreatic cancers[26].

1.2.2 Pathogenesis of Pancreatic Cancer

The majority of pancreatic cancers are ductal adenocarcinomas (PDACs)[27]. Microscopically, this type of cancer is characterised by a glandular structure with a ductal appearance and varying degrees of cellular atypia and differentiation[27]. Whilst the development of PDAC is generally regarded as sporadic, some studies have proposed that PDAC may arise from precursor lesions, which are cells with an atypical but non-cancerous cellular morphology that are frequently observed in association with PDAC[27-28]. Three types of PDAC precursor lesions have been proposed: pancreatic intraepithelial neoplasm (PanIN), mucinous cystic neoplasm (MCN), and intra-ductal papillary mucinous neoplasm (IPMN) [28-30].

PanINs are a relatively common finding in the elderly population[27]. This type of precursor lesion was initially associated with PDAC by post-mortem studies, which showed an increased incidence of PanINs in patients with PDAC[31]. The relationship between PanINs and PDAC was later reinforced by molecular profiling studies showing an increasing number of common genetic alterations between higher grade PanINs and invasive PDAC[29]. PanINs can be graded from stage I to III according to the degree of dysplastic growth (***Box 1.2***)[32-33]. The transformation of high-grade PanINs into PDAC is marked by the invasion of intra ductal carcinoma beyond the basement membrane[32-33].

Box 1.2- *Grading of Pancreatic Intraepithelial Neoplasms*[32-33]	
PanIN Grade	**Description**
PanIN-1A	Flat mucinous epithelium without cellular atypia
PanIN-1B	Papillary mucinous epithelium without cellular atypia
PanIN-2	increasing signs of cellular atypia and a prevalence of papillary architecture
PanIN-3	Carcinoma in situ/ intra-ductal carcinoma

MCNs are large mucin-producing epithelial cystic lesions that have a distinctive ovarian-type stroma with a variable degree of epithelial dysplasia and focal regions of invasion[27, 30, 34-39]. The majority of MCNs arise from the body and tail of the pancreas and do not communicate with the pancreatic ductal system except in the presence of erosions or fistulous tracts[27, 30, 34]. The association between MCN and PDAC is based on observational studies demonstrating the presence of invasive tubular/ductal adenocarcinoma in approximately one-third of all resected MCNs and on studies showing several common genetic mutations between MCN and PDAC (including the *KRAS-2* oncogene, *TP53, and SMAD4*)[28]. However, it should be noted that there are some controversies regarding the role of MCNs as precursors of pancreatic cancer due to the lack of direct evidence demonstrating the progression of MCNs to PDAC. This together with the fact that individuals with invasive ductal adenocarcinoma from resected MCNs have a much better prognosis (5-year survival, ~60%) compared to the reported prognosis for sporadic PDACs (5-year survival, <25%), suggest that pancreatic cancer associated with MCNs should be regarded as a separate entity compared to sporadic PDAC[35].

IPMNs account for 3-5% of all pancreatic masses and are the most common type of pancreatic cystic lesions [27, 40] and they are defined by the presence of mucin-filled cystic lesions ≥1cm in the main pancreatic duct and/or its secondary branches[27]. Histologically, IPMNs are characterised by tall, columnar mucin-secreting epithelial cells that form papillae with fibro-vascular core[40]. In addition, IPMNs with intestinal or pancreaticobiliary type differentiation often involve the main pancreatic duct and show moderate or high-grade dysplasia[27, 40]. Studies have demonstrated that approximately 20-50% of IPMNs are associated with the presence of invasive adenocarcinomas, which may be of either mucinous type or ductal type [40-44]. Interestingly, studies have shown that whilst adenocarcinoma associated with the ductal type of IPMN is morphologically identical and confers a similar prognosis to

non-IPMN associated PDACs, the adenocarcinomas associated with the mucinous type IPMN is characterised by neoplastic epithelial cells "suspended" in large pools of extracellular mucin and is associated with a better prognosis[42-44]. Similar to PanINs, the association between IPMNs and PDAC is only based on histological observational studies and studies indicating that IPMNs and PDACs share a number of genetic mutations (e.g. *KRAS2*, *p16*, *TP53*, and *SMAD4*)[28]. Therefore, in absence of evidence directly demonstrating the progression of IPMNs to PDAC, the role of IPMN in the development of PDAC will remain a controversial topic.

1.2.3 Molecular Hallmarks of Pancreatic Cancer

Pancreatic carcinogenesis is a complex process involving dynamic changes in the genome and molecular pathways, which together drive the progression from precursor lesions to invasive cancer[45]. Although the exact mechanism underlying pancreatic cancer formation is yet to be fully understood, research efforts in the past decade has significant improved our understanding of this disease. Recent studies have identified several key genetic mutations and signalling pathways, which have been found to be essential in pancreatic cancer tumourigenesis[26-27].

1.2.3.1 K-Ras signalling pathway

A number of studies have demonstrated that the mutation of the *K-Ras2* oncogene is present in 75-90% of all pancreatic cancers[26, 46-47]. The *K-RAS* gene is the cellular homologue of the *RAS* gene of Kristen murine sarcoma virus. It encodes for a 21-kDa membrane-bound GTP-binding protein (KRAS protein), which is involved in growth factor-mediated signal transduction[48]. The *K-RAS* signalling pathway plays an important role in promoting cell cycle progression, cell proliferation, and resistance to apoptosis[26, 47]. Point mutations of the *K-RAS* gene often result in an impaired GTPase activity, which means that the KRAS protein is locked in the GTP-bound (activated) state and therefore, the permanent activation of downstream signalling cascades[26, 47]. In addition, the *K-RAS* signalling pathway can be activated through the over expression or activation of its upstream receptor molecules, such as epidermal growth factor receptor (EGFR)[26]. Interestingly, *K-RAS* mutations are also found in patients with chronic pancreatitis and are therefore not exclusive to malignant cells of the pancreas[27]. Furthermore, activation of the *K-RAS* signalling

pathway alone is insufficient to cause malignant transformation of pancreatic cells[26-27, 47]. It has been proposed, therefore, that co-existing aberrations such as epigenetic silencing of tumour-suppressors or activation of other oncogenic pathways must also be present[26-27, 47].

1.2.3.2 Tumour suppressor genes and pathways

The deletion of the p16 INK4A gene locus in up to 95% of cases and the alteration or deletion of the p53 gene locus in 50-75% of cases are the most frequently reported mutations observed in pancreatic cancer[27]. Tumour suppression by both of these genes is made possible through the inactivation of CDK4/6 and CDK2 thereby inhibiting the phosphorylation of the retinoblastoma protein and subsequently preventing cell cycle progression through the G1-S checkpoint[26-27]. Furthermore, TP53 also contributes to tumour suppression by regulating cell cycle and promoting apoptosis in cells when DNA damage is sustained.

Another common mutation is the deletion of the Smad4 gene, which is found in 55% of pancreatic cancers[26-27]. Smad4 is an important downstream mediator for the Transforming Growth Factor (TGF-β) signal pathway (an inhibitory pathway for PDAC) and is responsible for the transmission of TGF-β signals into the nucleus thereby regulating the expression of cancer-associated genes[26-27]. In addition, studies have demonstrated that the disruption of TGF-β signalling pathway facilitates cancer cell growth, differentiation, and migration[26-27].

1.2.3.3 Embryonic Signalling pathways

A number of studies have reported that embryonic signalling pathways such as Hedgehog and Notch are reactivated in pancreatic cancer. In particular, over-expression of the Indian and/or sonic Hedgehog ligands have been associated with enhanced tumour progression[26-27, 49]. Moreover, recent studies have demonstrated that the expression of the sonic Hedgehog ligand in transgenic mice results in the formation of PanIN-like lesions[50].

Notch is an embryonic signalling pathway, which controls cellular differentiation, proliferation, and apoptosis[51-53]. Although it is not usually active in the pancreas, up-regulation of Notch target genes has been observed in pre-neoplastic lesions and

invasive pancreatic cancer[52]. In addition, Notch signalling has been shown to promote neo-vascularisation of tumours[51, 53].

1.2.4 Symptoms and Signs of Pancreatic Cancer

For the majority of patients, pancreatic cancer remain undiagnosed until it is at an advanced stage and at initial presentation, symptoms are often non-specific (***Box 1.3***)[21]. Unlike other cancers, for example prostate and breast cancer, the anatomical location of the pancreas means that pancreatic tumours cannot usually be felt on physical examination[3]. Therefore, the suspicion of pancreatic malignancy relies heavily on systemic symptoms and signs, which are often indications of disparate disease[54-55].

The majority of symptoms can be explained in terms of compression of anatomical structures within or adjacent to the pancreas. In particular, tumours arising from the head of the pancreas (approximately 70%) can cause obstruction of the duodenum, pancreatic duct, and common bile duct[54-55]. When the main pancreatic duct is obstructed, activation of digestive enzymes secreted by the pancreas may lead to auto-digestion of the pancreatic parenchyma and subsequently pancreatitis whereas bile duct obstruction is likely to result in cholestasis and jaundice[54-55]. Furthermore, pancreatic cancer may cause the dysfunction of pancreatic endocrine cells leading to dysglycaemia[20, 54-55]. It is therefore important for clinicians to consider pancreatic cancer as a differential diagnosis in patients presenting with acute pancreatitis or new onset diabetes[20].

Clearly, the biggest problem encountered in the diagnosis of pancreatic cancer is that all the symptoms associated with pancreatic cancer can also be associated with non-malignant diseases (***Box 1.3***)[54]. In particular, jaundice, which is a commonly observed symptom in pancreatic cancer, is also frequently reported in benign pancreaticobiliary diseases such as biliary obstruction and chronic pancreatitis.

Box 1.3*[21, 54] *- Signs and symptoms of pancreatic cancer and confounding diseases	
Symptom in PDAC	**Examples of confounding diseases exhibiting the same symptom**
Anorexia	Bacterial/ viral infections; **most cancers**, **Chronic Pancreatitis**; **gallstone related biliary obstruction**
Weight Loss	**Most cancers**; acute infections; diabetes; **Chronic Pancreatitis**
Jaundice	**Biliary obstruction**, **acute and chronic pancreatitis**; liver failure
Hepatomegaly	Hepatitis, heart failure; liver cirrhosis
Peripheral lymphadenopathy	Heart failure, renal failure; **other cancers**; medications; malnutrition
Abdominal Pain	**Chronic Pancreatitis; gallstone related Biliary obstruction**; peptic ulcers; inflammatory bowel disorders; GI cancers
Anaemia	**Most cancers;** autoimmune diseases; B12 deficiency; malabsorption; GI bleeding
Fatigue	**Most cancers;** diabetes; obesity heart failure; anaemia; depression
Ascites	Liver **metastasis**; **most cancers;** malnutrition; cirrhosis (liver or biliary); heat failure
Acute pancreatitis	Alcoholic pancreatitis; acute exacerbation of **chronic pancreatitis**; cancer of biliary tree or duodenum; **gallstone-related biliary obstruction**

1.2.5 Diagnosis and Staging of Pancreatic Cancer

In patients with suspected pancreatic cancer, subsequent investigations are designed to provide information regarding the presence, location, staging, and resectability of the disease[21, 26, 54, 56]. A number of modalities have been developed for the diagnosis and staging of pancreatic cancer. The initial investigation may be a simple trans-abdominal ultrasound scan (USS), which may be able to identify signs of late pancreatic cancer such as biliary dilatation and liver metastasis[57-58]. However, USS is not useful in the diagnosis of early pancreatic cancer[57] and therefore, the preferred diagnostic investigation is contrast enhanced multi-slice computed tomography (CE-CT) scan, which is able to assess the location, size and sometimes the type of lesion in addition to providing evidence for the staging and resectability of the tumour[57, 59]. In general, contrast-enhanced CT scans are 80-90% accurate in predicting surgical resectability[59]. Other investigations such as endoscopic ultrasonography (EUS), fine needle biopsy (FNB), and Laparoscopy may be useful in confirming the presence of smaller or equivocal lesions seen on CT scan[60]. Nevertheless, the final diagnosis of PDAC can only be made histologically[21].

Pancreatic cancers are staged using the Tumour, Node, and Metastasis (TNM) classification system[61]. Each tumour is scored according to three criteria: Tumour size/extent (T), Lymph node involvement (N) and the presence of distant metastasis (M) *(see **Box 1.4**)*. In practice, patients with TNM status equal to or less severe than T3, N1, and M0 are considered resectable[61]. However, studies have demonstrated

that patients with positive lymph node involvement (N1, staged 2B or above) have considerably poorer survival compared to patients with N0 status[62-66].

Box 1.4- TNM classification system and staging for pancreatic cancer[61]			
TNM	**Description**	**Stage with N0**	**Stage with N1**
Tx	Primary tumour cannot be assessed	-	
T0	No evidence of primary tumour	-	
Tis	Carcinoma in situ	-	
T1	Tumour limited to pancreas, 2cm or less in greatest dimension	1A	2B
T2	Tumour limited to pancreas, more than 2cm in greatest dimension	1B	2B
T3	Tumour extends beyond pancreas, but without involvement of celiac axis or superior mesenteric artery	2A	2B
T4	Tumour involves celiac axis or superior mesenteric artery	3	3
Nx	Regional lymph nodes cannot be assessed		
N0	No regional lymph node metastasis		
N1	Regional lymph node metastasis		
M0	No distant metastasis		
M1	Distant metastasis	4	4

1.2.6 Management and prognosis of pancreatic cancer

1.2.6.1 Advanced pancreatic cancer

Over 75% of pancreatic cancers are inoperable due to localised advanced disease, metastases, or performance status[67-68]. The treatment for this group of patients is therefore directed at symptom control[57]. Pain is one of the most commonly reported symptoms of inoperable pancreatic cancer and it is usually controlled by oral opiate preparations[69]. Recent studies have suggested that celiac plexus block may improve pain control in selected patients but it has an insignificant effect on the quality of life and survival[70-72]. Furthermore, patients with advanced pancreatic cancer invariably develop weight loss due to pancreatic exocrine insufficiency, obstruction of the common bile duct, or cancer-associated cachexia[26]. Whilst there is no treatment for the latter condition, the former two can be effectively treated by pancreatic enzyme supplements and stenting of the bile duct[26, 73].

Chemotherapy plays an important role in improving the survival and quality of life of patients with advanced pancreatic cancer[74-75]. Unlike other cancers, PDACs are highly resistant to chemotherapy with a relatively low response rate of 10-25%[74]. However, studies have shown that treatment with chemotherapeutic agents such as 5-Fluorouracil (5-FU) and Gemcitabine can significantly improve the median survival of patients with advanced pancreatic cancer[74]. Evidence from a randomised

controlled trial by Burris *et al.* [74] reported not only that treatment with Gemcitabine significantly improves the 1 year survival rate compared to 5-Fluorouracil (18% versus 2%, respectively), this trial also observed a milder toxicity and a better clinical response (24% versus 5%) with Gemcitabine. This subsequently saw the replacement of 5-FU with Gemcitabine as the preferred drug[26]. In addition, a number of randomised controlled trials have demonstrated that for patients with advanced pancreatic cancer, the combination of Gemcitabine with newer chemotherapeutic agents such as Capecitabine (a 5-Fluorouracil pro-drug) and Oxaliplatin is associated with a better prognosis[75-80]. In particular, a phase III randomised control trial by Cunnigham *et al.*, reported a significant improvement in progression-free survival for patients with advanced PDAC treated with Gemcitabine -Capecitabine combined chemotherapy compared to Gemcitabine alone (hazard ratio 0.78, p=0.004)[80].

1.2.6.2 Resectable pancreatic cancer

In the remaining 25% of pancreatic cancer patients, where surgical resection with intention-to-treat is deemed possible, the aim of surgery is to achieve complete clearance of the tumour both macroscopically and microscopically (R0 resection)[81-82]. However, in practice, a large proportion of patients have incomplete resection of the tumour (R1, microscopically; R2, macroscopically) [82]. The most commonly employed surgical procedure for the removal of pancreatic tumours located in the head of the pancreas is pylorus-preserving partial pancreaticoduodenectomy (PP-PPD) whereas tumours located in the body or tail undergo distal pancreatectomy with resection of the spleen and hilar lymph nodes[67].

Despite radical resection of the primary tumour, the reported 5-year survival rate remains low (approximately 10%), mainly due to cancer recurrence. Furthermore, pancreatic cancer patients (resected or otherwise) will invariably develop metastatic disease, typically of the liver or lung. Various prognostic markers have been identified, of which, lymph node status, tumour size and tumour grade are the most important predictors of post-operative survival. Evidence from randomised controlled trials by Neoptolemos *et al.* [83-85] and Oettle *et al.* [86] independently reported that the use of adjuvant chemotherapy improves the 5-year survival from 9-12% (resection alone) to 21-29% (resection with chemotherapy)[87-89]. Further evidence from the European Study Group for Pancreatic Cancer 3 (ESPAC-3) trial[90] supported

this finding and, in addition, reported comparable survival rates in patients treated with adjuvant 5-FU compared to adjuvant Gemcitabine .

1.3 Biomarkers for pancreatic cancer

1.3.1 Introduction to biomarkers

A biomarker is defined by the National Institute of Health[91] as "a characteristic that is objectively measured and evaluated as an indicator of normal biologic process, pathogenic process, or pharmacologic responses to a therapeutic intervention". The abundance or scarcity of cancer biomarkers in cancer relative to non-cancer conditions (e.g. inflammatory diseases and in health) may be an indication of changes to cellular biology in carcinogenesis. Therefore, it is the aim of cancer biomarker studies to identify these differentially expressed molecules and to assess their clinical usefulness as a screening, diagnostic, and/or prognostic modality for cancer. There are many samples, which can be used for the discovery of biomarkers including blood derivatives, pancreatic juice, tissue, saliva, and urine[92]. Of these, blood plasma or serum is most widely used in biomarker studies because they are readily accessible, minimally invasive to collect, generally acceptable to patients, and are potentially rich sources for most types of biomarkers[92].

1.3.1.1 The need for diagnostic biomarkers of pancreatic cancer

As previously discussed, the differential diagnosis of pancreatic cancer is based entirely on non-specific symptoms, signs, and first-line investigation findings that are at best suggestive of a disease of pancreaticobiliary origin and/or the presence of metastatic disease[54-55]. Moreover, under the current recommendations, patients suspected of having pancreatic cancer will undergo a series of relatively invasive procedures including CE-CT, EUS, and/or FNB[57]. Clearly, there is a need for novel, accurate, and less invasive methods for the diagnosis of pancreatic cancer such as a blood-based protein biomarker. Aside from the fact that a blood-based biomarker would be less invasive compared to current diagnostic techniques, there are two other major advantages: the number of patients undergoing unnecessary invasive investigations and the time required to reach a diagnosis of PDAC would be greatly

reduced. There is also the possibility that biomarkers can be employed to detect the presence of early pancreatic cancer in otherwise asymptomatic individuals. Subsequently, this would mean that patients with PDAC are diagnosed earlier thereby increasing their chances of having operable disease and therefore improving the prognosis.

1.3.2 Current biomarker of pancreatic cancer: CA19-9

The tumour associated antigen, Carbohydrate Antigen 19-9 (CA19-9), was first described in pancreatic cancer by Koprowski *et al.* in 1981[93]. Approximately 95% of the general population are able to synthesise CA19-9 while the remaining 5-10% of the population, due to genetic differences, have a Lewis^{a-b-} phenotype meaning that they are unable to synthesise CA19-9[94].

1.3.2.1 CA19-9 in screening and diagnosis of pancreatic cancer

CA19-9 was initially evaluated as a potential diagnostic marker exclusively for pancreatic cancer; however, studies in the past two decades have reported elevated levels of CA19-9 in other malignant tumours including gastric, ovarian, hepatocellular, and colorectal cancers as well as benign pancreaticobiliary diseases such as chronic pancreatitis, cholangitis, and choledocholithiasis[23, 95-99]. Furthermore, a recent study by Morris-Stiff *et al.* demonstrated a direct correlation between serum levels of CA 19-9 and bilirubin [100]. This finding, together with evidence from clinical studies demonstrating that CA19-9 is not sufficiently sensitive for the detection of early or small-diameter pancreatic cancer, suggest that CA19-9 should not be used alone as a screening modality for pancreatic cancer[96, 101-102].

Evidence from systematic reviews have suggested that with a median sensitivity of 79% (70-90%) and a median specificity of 82% (58-91%), CA19-9 is not sufficiently accurate as a standalone diagnostic marker of pancreatic cancer[97, 99]. Indeed various expert groups including the European Group on Tumour Markers, the National Academy of Clinical Biochemistry (NACB), and the National Cancer Comprehensive Network have stated that CA19-9 should only be used in conjunction with other diagnostic modalities such as CE-CT and EUS[96, 101-102].

1.3.2.2 CA19-9 in prognosis, surveillance, and assessment of chemotherapy

Mounting evidence shows that serum concentrations of CA19-9 correlate with the prognosis of resected pancreatic cancer patients[103-106]. In a retrospective study by Ferrone *et al.*, it was reported that a CA19-9 level of <200 kU/l or a decrease in CA19-9 levels following surgical resection are independently associated with better prognosis[107]. Recently, the NACB guidelines recommend that whilst CA19-9 should be considered for risk stratification in patients with pancreatic cancer and that high concentrations are indicative of poor outcome, the guidelines emphasised that CA19-9 is only one of many factors influencing the prognosis and treatment planning of pancreatic cancer[102].

CA19-9 is also used in postoperative surveillance of pancreatic cancer[96, 108]. Several studies have shown that sequential measurements of CA19-9 may be able to detect recurrent/metastatic pancreatic cancer before clinical or radiological evidence[95-96, 108]. In addition, there is a consensus that a declining CA19-9 level following initiation of chemotherapy is associated with a better outcome compared to no decline[96]. Based on this evidence, the NACB recommends that serial measurements of CA19-9, along with radiological imaging at regular intervals may be used for both post-operative surveillance and the monitoring of therapy[102].

1.3.3 Other markers of pancreatic cancer

A number of potential diagnostic markers for pancreatic cancer have been proposed in the past decade[96, 109]. In 2007, Grote *et al.*[109] highlighted 16 novel blood-based markers for pancreatic cancer in their review article including Mucin 1 (MUC-1), macrophage inhibitory cytokine 1 (MIC-1), inter-alpha-trypsin-inhibitor heavy chain 4 fragments (ITIH4 fragments), and Apolipoprotein A-II, which have shown relatively high sensitivities (range, 0.71-0.90) and specificities (range, 0.92-0.96) for pancreatic cancer against chronic pancreatitis and healthy controls compared to CA19-9. Similarly, results from a review by Bussom *et al.*[110] identified several novel biomarkers e.g. PAM4 and carcinoembryonic antigen-related cell adhesion molecule-1 (CEACAM-1), which have shown promising results for the detection of early-stage pancreatic cancer. Moreover, several studies have demonstrated the potential of combining novel markers with CA19-9 to improve the diagnostic

accuracy compared to CA19-9 alone[111-112]. However, considering the relatively small sample size in these studies and in view of recent evidence demonstrating the confounding effects of biliary obstruction on the diagnostic accuracies of some proteomic biomarkers, further validation must be performed before these novel markers can replace the role of CA19-9 in pancreatic cancer.

1.3.4 Techniques for biomarker discovery

There are two main approaches to biomarker discovery: The genomic approach, which focuses on identifying genetic mutations or changes in gene expression on micro RNA levels and the proteomics approach, which mainly examines the difference in protein levels between PDAC and benign conditions. The following sections will describe the various techniques used in the discovery of proteomic biomarkers including 2D polyacrylamide gel electrophoresis (2D PAGE) and isobaric Tag for Relative and Absolute Quantification (iTRAQ) in addition to common techniques used in the validation of biomarkers such as western blotting and enzyme-linked immunosorbent assay (ELISA). Furthermore, the current section will describe the use of a microsphere-based multiplex cytokines assay, for both the discovery and validation of biomarkers for pancreatic cancer.

1.3.4.1 2D-PAGE

The technique of two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) was first described by O'Farrel *et al.* and Klose in 1975[113-116]. The original technique of 2D-PAGE described by these authors consisted of protein separation by carrier-ampholyte-generated pH gradients in the first dimension (isoelectric focusing, IEF) followed by separation by protein molecular weight in the second dimension (SDS-PAGE) [113-116].

2D-PAGE has a number of desirable properties and potential applications including its ability to separate proteins into their individual polypeptide components, compare protein expression profiles of paired samples (e.g. cancer versus control), detect global protein behavioural in responses to a change in conditions, and more importantly, its potential capacity to simultaneously resolve hundreds to thousands of proteins[116-119]. This technique was widely applied throughout the 1980s[118], however,

it was not long before researchers recognised that the original 2D-PAGE method suffered from a number of limitations including the lack of reproducibility, low resolution, inability to separate very acidic and/or very basic proteins, and limited sample loading capacity[116, 118]. In an effort to overcome these limitations, Görg *et al.*[120] introduced a new gradient for first dimension separation in the 1980s- the immobilized pH gradients (IPGs) [120-122]. The use of IPG enabled an extremely stable pH gradient to be generated, which subsequently improved isoelectric focusing and the reproducibility of the technique[118]. Later studies further enhanced this technique by introducing narrow-overlapping IPGs, which enabled a higher resolution as well as permitted the detection of lower abundance proteins and proteins with isoelectric points ranging from pH 2.5 to pH 12[123-125].

Research on 2D-PAGE and its related technologies in the past decade has been focused on improving the solubilisation and separation of hydrophobic proteins, the display of low abundance proteins, and achieving more reliable protein quantification by either fluorescent dyes or isobaric tags[118]. Indeed, the recent development of the difference gel electrophoresis (DIGE) technology has enabled mixed samples to be analysed on a single 2-DE gel via differential fluorescent dye labelling[126].

1.3.4.2 Mass Spectrometry and iTRAQ

Isobaric Tag for Relative and Absolute Quantification (iTRAQ) is a chemical labelling multiplexing technique, which quantifies the concentration of proteins using mass spectrometry[127-129]. iTRAQ coupled with electro-spray ionisation tandem mass spectrometry (ESI-MS/MS) is becoming increasingly popular over the past 5 years in the field of biomarker research due to its ability to identify and quantify hundreds of proteins in a single experiment[128]. Indeed, a previous study from the Division of Surgery and Oncology, University of Liverpool, have identified over 300 differentially expressed proteins in the serum of pancreatic cancer patients compared to controls[129]. In Chapter 2, the accuracies of three iTRAQ/MS identified proteins as diagnostic markers for pancreatic cancer is provided.

The technique of iTRAQ-MS relies on the fact that proteins can be digested to a unique set of different tryptic peptides. The iTRAQ part of the technique consists of the digestion of proteins in a sample into their constituent peptides followed by the labelling of these peptides by isobaric tags. Each sample group (i.e. HC, CP, and

PDAC) is labelled with a different tag with a unique reporter group of a specific mass, which is released during mass spectrometry through collision-induced dissociation thus allowing the association of a peptide with a specific sample group. The detection part of iTRAQ-MS involves the vaporisation and ionisation of the labelled peptides through an electromagnetic field. The resulting trajectory data and mass-to-charge ratio data from the MS analysis can then be used to identify the protein origins of these peptides. The different tags allow the relative quantification of peptides between samples, for example, a given peptide labelled with the disease-specific tag could be four times more abundant than the same peptide labelled with the control specific tag, indicating that the peptide is more abundant in the disease than the controls.

Mass spectrometry based methods such as iTRAQ offers the identification and quantification of numerous proteins in a single experiment. This is clearly advantageous for biomarker studies, where several potential biomarkers can be identified upon data comparison between the disease and control groups. However, there are two major drawbacks for this technique: iTRAQ requires a large amount of sample and the sample preparation stage may span require weeks to complete especially in the case of serum, where abundant protein depletion is necessary. An acceptable solution to this problem is the use of pooled samples but this is not without its own disadvantages. Although pooled samples present an 'average' profile for the disease group in question, this "average" is very susceptible to skewing by an outlier with unusually high or low expression of a particular protein. Therefore, it is important to validate the results from pooled samples by other proteomic methods such as western blotting or ELISA.

1.3.4.3 Western blotting

Western blotting (also called Protein Immunoblotting) is a laboratory technique used to detect the presence of specific proteins in a given sample and it is frequently employed to validate the serum expression of proteins identified by iTRAQ. The method for western blotting was first described by Burnette in 1981[130] and surprisingly little has changed in the original method over the past 20 years. The most prominent changes have been in the development of newer apparatus, specific

antibodies, and better labelling/detection methods. Indeed, novel detection methods for western blots such as infrared labelling have been of increasing interest to researchers. In a recent study comparing tradition chemiluminescence with infrared detection[131], the authors described many potential advantages to infrared detection such as the ability to simultaneously detect two proteins and the loss-less preservation of signal long periods of time. Due to its higher detection sensitivity however, the resulting blot may in practice show more non-specific bands compared to traditional chemiluminescence and therefore may not be desirable.

The technique of western blotting involves two phases: sodium dodecyl sulphate (SDS) polyacrylamide gel electrophoresis (PAGE) and immunofluorescence detection. The purpose behind SDS-PAGE is to separate uniformly charged, denatured proteins according to their molecular weight through the use of an electric current. Briefly, denatured proteins are reduced by SDS (i.e. given a uniform charge) and are subjected to an electric current. Proteins with a lower molecular weight will travel through the polyacrylamide gel at a faster rate compared to heavier proteins, therefore, with time, lighter proteins will travel a greater distance compared to heavier proteins. The proteins in the polyacrylamide gel are then transferred and immobilised on a nitrocellulose membrane in preparation for immuno-detection. In the detection phase, a primary antibody is used to bind to the protein of interest and a horseradish peroxidase (HRP) secondary antibody is used to bind to the primary antibody. This is necessary because the HRP on the secondary antibody allows chemiluminescence detection by catalyzing the reaction between two luminescence substrates.

Western blotting remains one of the most favoured techniques in modern proteomic research for the validation of the findings in biomarker discovery projects. This is largely because of its high sensitivity for detecting the presence of specific proteins meaning that it is heavily relied upon in cell-related proteomic analyses. Furthermore, in biomarker research where samples are often very precious, western blotting has the advantage of being able to detect and quantify proteins using a considerably smaller amount of sample compared to other proteomic techniques such as ELISA or mass spectrometry. However, there are a number of potential pitfalls in using western blotting for the quantification of proteins. Firstly, there is often a variable

loss of proteins during the transfer step and therefore direct comparison between blots is not recommended unless internal standards are rigorously used. Secondly, the primary antibody, which is engineered to recognise a specific amino acid sequence unique to the protein of interest, may have a variable ability in recognising the same sequence after the protein has been denatured. Finally, the quantification of bands from scanned x-ray films must be interpreted with care because the intensities of x-ray scanned bands follow a sigmoid-like correlation with the actual concentrations of the protein of interest in a sample[132]. This means that the difference in band intensity is likely to be less distinct when compared to the difference in actual protein concentration.

1.3.4.4 Enzyme-Linked Immunosorbent Assay (ELISA)

Enzyme-linked immunosorbent assay (ELISA), sometimes referred to as enzyme immunoassay (EIA), is another technique frequently employed in the quantification of blood-based proteins. Although it is widely used in the research setting, it is also frequently employed in clinical settings especially in diagnostic medicine. Interestingly, the technique of ELISA is the result of synthesized knowledge from studies published by a number of researchers from 1960 to 1971[133-135]. As with western blotting, very little has changed in the original method since 1971 other than the development of better antibodies and more sophisticated/automated apparatus, which play a pivotal role in minimised human error and improving data reproducibility. There are a number of variations to the ELISA technique including direct, indirect, sandwich, competitive, and multiplex assays however, the principle behind the variations remains the same.

The technique of ELISA is somewhat similar to the chemiluminescence phase of western blotting. In a typical sandwich ELISA, the protein of interest (antigen) is first immobilised to the surface of the wells by a capture antibody. A primary antibody then binds to the antigen, which in turn is bound by an antibody-HRP conjugate. The HRP in the secondary antibody catalyses a chromogenic substrate to cause a shift in colour that is directly correlated with the concentration of the protein of interest. This colour shift is detected by an ELISA plate reader and the resulting data can be used to quantify the concentration of the protein of interest.

There are many advantages in using ELISA for protein quantification, for example, the assay itself is simple and relatively quick to perform, each analysis simultaneously quantifies over 40 samples, and the resulting data are readily reproducible. However, as with western blotting, one of its basic requirements is the availability of an antibody specific to the antigen of interests. In addition, the serum sample requirement for a typical ELISA (25 µL) is over 60 times greater than a typical western blot analysis (0.4µL) and it is usually more expensive (per sample) to perform.

1.3.4.5 Multiplex Assays (LUMINEX)

Researchers and biomedical companies in the past decade have sought to overcome the various disadvantages of conventional ELISAs[136-137]. One of the ways in which this is achieved is by combining the ELISA technology with existing biochemical techniques such as microspheres, flow cytometry, and laser detection[136-137]. In chapter 3 of the current MPhil thesis, I described the use of the LUMINEX multiplex assay to analyse 27 different cytokines, chemokines, and growth factors in serum samples. A brief description of this technique is given below (see ***section 3.3.3*** for detailed protocol).

The technique for the LUMINEX assay is similar to conventional sandwich ELISA in many ways. However, instead of using an antibody-based capture system, the LUMINEX assay employs a solid-phase microsphere-based capture system where each microsphere population are designed to bind to a specific analyte of interest. Furthermore, each microsphere population is internally dyed to emit a unique wavelength in the red to infrared spectrum upon laser excitation, which allows the identification of its corresponding analyte. As in the case of ELISA, samples are incubated with a mixture of microspheres to allow binding of the analyte to its corresponding microsphere population. Next, the microsphere-substrate complex is allowed to bind to a detection antibody, which facilitates the binding of streptavidin. Finally, the LUMINEX system employs a dual laser technology whereby a red laser is used to identify the microsphere population, and a green laser is used to detect the fluorescence intensity of each microsphere. Data from the assay standards can then be used to convert the fluorescence intensities into concentration (pg/ml).

The LUMINEX system is a relatively new technology with the obvious advantage of being able to quantify up to 100 analytes in up to 39 samples simultaneously and is therefore significantly more sample-efficient than a conventional ELISA. The substitution of enzyme-catalysed chromogenic substrate (ELISA) with immunofluorescence and laser detection (LUMINEX) meant that a wider range of concentrations could be determined. In addition, results from previous studies have shown that data obtained from multiplex assays are reproducible and have an acceptable accuracy. Furthermore, it is a simple and quick assay to perform with an experimental duration of approximately 4 hours. However, as with other immunoassays, the integrity of the assay is dependent upon the sensitivity and specificity of the primary antibody. Moreover, one major disadvantage is the cost incurred by each assay, which can be up to 5 times the cost of an ELISA plate. Nevertheless, considering its ability to quantify multiple analytes simultaneously, the LUMINEX assay can be considered as sufficiently cost-effective as an alternative method for the discovery of potential biomarkers.

1.4 Role of CCGFs in inflammation and cancer

The role of cytokines, chemokines, and growth factors (CCGFs) in pancreatic inflammatory diseases (e.g. acute and chronic pancreatitis) has been reported by many studies[138-144]. More recently, evidence from experimental and epidemiological studies have shown that CCGFs play a pivotal role in mediating cancer-related inflammation[144-150]. Indeed, a recent review by Colotta *et al.* suggested that cancer-related inflammation may be considered as the seventh hallmark of cancer[150] and that an inflammatory tumour microenvironment contributes to the proliferation, angiogenesis, survival and metastasis of certain cancers[143, 151-154].

1.4.1 The relationship between inflammation and cancer

The association between chronic inflammation and cancer was initially proposed by Rudolf Virchow over 100 years ago[145]. Later studies have given evidence in support of this hypothesis and it was estimated that 20% of all cancers are attributable to chronic infection and inflammation[145, 155]. Epidemiological studies have suggested that the used of anti-inflammatory agents may reduce the risk and mortality of certain cancers[156-159]. More importantly for pancreatic cancer, studies have demonstrated a correlation between the activation of certain oncogene signalling pathways (e.g. *Ras*, and TGF-β) and inflammation[26, 46, 160]. In 2008, Mantovani *et al.* described two pathways linking inflammation and cancer: the intrinsic and extrinsic pathways[155]. The intrinsic pathway is activated by the mutation of certain cancer-related genes (e.g. activation of oncogenes or inactivation of tumour suppressor genes) thereby resulting in the production of inflammatory mediators and ultimately the formation of an inflammatory tumour microenvironment. In contrast, the extrinsic pathway refers to the promotion of tumour formation in the presence of an underlying inflammatory condition (such as chronic pancreatitis)[155]. Both pathways lead to the activation of transcription factors such as Nuclear Factor Kappa light-chain-enhancer of activated B cells (NF-κB), Hypoxia Inducible Factor 1-alpha (HIF-1α), and Signal Transducer and Activator of Transcription-3 (STAT3), which in turn induce the secretion of inflammatory mediators (e.g. cytokines, chemokines, and growth factors) by immune cells[145, 161].

1.4.2 Pathways linking inflammation and cancer

1.4.2.1 NF-κB

NF-κB is a family of transcription factors consisting of five members including NF-κB1, NF-κB2, RelA, RelB, and c-Rel[162-163]. Members of the NF-κB family share a 300 amino acid region, which gives NF-κB the characteristic ability to form dimmers and to bind with other NF-κB proteins, DNA, and NF-κB inhibitors (IκB)[164]. In normal eukaryotic cells, NF-κB is usually in a quiescence state; however, NF-κB can be activated by a number of stimuli including cellular stress, inflammatory mediators (e.g. TNF-α and IL-1β), and bacterial/viral antigens[162-163]. The activation of NF-κB promotes the transcription of a number of inflammatory mediators including IL-1α, IL-1β, and TNF-α, which may in turn stimulate the NF-κB pathway in a positive feedback loop and may lead to a longer and more severe inflammatory response[162-163].

Aberrant activation of the NF-κB pathway has been associated with a number of cancers[162-163, 165-168]. It has been proposed that the continuous activation of NF-κB promotes oncogenesis by enhancing the transcription of genes encoding for cytokines (e.g. TNF-α), chemokines (e.g. IL-8), and growth factors (e.g. Vascular Endothelial Growth Factor, VEGF) as well as anti-apoptosis genes[162-163, 166]. Furthermore, evidence from a number of *in vitro* and *in vivo* studies has demonstrated that NF-κB can promote tumour metastasis by regulating the expression of cell adhesion molecules (e.g. ICAM-1), cell surface proteases (e.g. MMP-9), and plasminogen activators (e.g. urokinase-type plasminogen activator).

1.4.2.2 STAT3

STAT3 is a member of the signal transducer and activator of transcription (STAT) family of proteins, which are responsible for the transduction of cytoplasmic signals from extracellular stimuli, the regulation of genes involved in tumour proliferation, survival, angiogenesis, and metastasis, and the induction of tumour-promoting inflammatory mediators especially cytokines (e.g. IL-17) and growth factors (e.g. VEGF)[169]. Interestingly, whilst some members of the STAT family (e.g. STAT1) exhibit anti-tumour properties, others (e.g. STAT3) induce cancer-promoting inflammation[170-174]. Furthermore, studies have shown that STAT3 is commonly

activated in malignant cells and it plays a crucial role in regulating the expression of genes associated with cancer-related inflammation in the tumour microenvironment[169, 175]. STAT3 can be activated via a number of intrinsic and extrinsic mechanisms. In particular, STAT3 can be activated by a number of upstream receptors including cytokine receptors (e.g. IL-6R, IL-10R) and growth factor receptors (e.g. Platelet-derived growth factor (PDGF) receptor and EGFR)[169, 174]. Indeed, recent studies have demonstrated that IL-6 can activate STAT3 through the activation of Janus Kinase (JAK), which ultimately leads to the up-regulation of anti-apoptotic genes and tumour cell survival[170, 173, 176-178]. Interestingly, IL-10 has been shown to exhibit anti-tumour activity by inhibiting the NF-κB pathway[179]. However, recent evidence has also indicated that IL-10 may play a dual role in cancer by being an activator of STAT3 thereby indirectly promoting cancer proliferation, angiogenesis, survival, and metastasis through promoting the transcription of anti-apoptotic genes and growth factors[169, 180].

Furthermore, in view of the common role of STAT3 and other cancer-associated transcription factors in cancer-associated inflammation, it is somewhat unsurprising that the two transcription factor pathways should interact with each other on many levels. For example, recent studies have suggested that the activation of NF-κB promotes the expression of IL-6 gene, which in turn is an activator of STAT3. In contrast, the activation of STAT3 prevents RELA of the NF-κB family from leaving the nucleus thereby contributing to the persistent NF-κB activation in cancer[181].

1.4.2.3 HIF 1-alpha

One of the major factors dictating tumour growth, proliferation, and survival is the availability of oxygen[182]. Cancer associated hypoxia results when the demand for oxygen from the rapidly growing tumour exceeds its vascular supply. Prolong hypoxia leads to cellular necrosis, which is often observed in solid tumours. Indeed, evidence from computed tomography and intratumoural oxygen tension studies indicated that pancreatic cancers are characterised by an avascular appearance and that they are unusually hypoxic compared to other solid tumours[183].

HIF are a family of transcription factors consisting of six members (HIF-1α, HIF-1β, HIF-2α, HIF-2β, HIF-3α, and HIF-3β), which are responsible for promoting neo-angiogenesis and wound healing[182]. One of the best-characterized HIFs is the heterodimer protein, hypoxia-inducible factor-1. HIF-1 consists of two subunits, HIF-1α and HIF-1β[183-184]. Whilst both subunits are expressed in all cells, HIF-1α is virtually undetectable in well-oxygenated cells due to its rapid degradation by ubiquitination[183-184]. In contrast, during cellular hypoxia, there is an accumulation of HIF-1α, which undergoes dimerization with HIF-1β and ultimately promoting the transcription of various angiogenesis-related proteins including vascular endothelia growth factor (VEGF)[183-184].

Although the exact role of HIF-1 is not yet fully understood, a number of studies have demonstrated that HIF-1α expression is positively correlated to VEGF expression, tumour size, and tumour stage (particularly stage III and IV)[183]. Furthermore, there is increasing evidence suggesting that HIF-1α may enhance the expression of motility factors in pancreatic cancer cells and may therefore play a role in promoting metastasis[183].

1.4.3 Tumour microenvironment, cytokines and cancer

There are three broad categories of cells within the tumour microenvironment: immune cells, cancer cells, and stromal cells[185-186]. The communication within and between these cells relies on direct contact or through signalling molecules such as cytokines, chemokines, and growth factors (CCGFs), which act in autocrine and paracrine manners to control tumour growth, proliferation, migration, and metastasis[186-187]. CCGFs and their modulators play a paradoxical role in cancer where their expression and abundance within the tumour microenvironment dictate the balance between tumour promoting inflammation and anti-tumour immunity[179]. When host-mediated anti-tumour immunity is stronger than tumour-induced immunosuppression, there is a net elimination of tumour cells[179]. By contrast, in established tumours, the balance is shifted towards tumour-associated inflammation and there is a net promotion of tumour growth and proliferation[179]. Furthermore, in advanced cancer, there is minimal anti-tumour activity and therefore tumour regression rarely occurs without therapeutic intervention [179, 188].

1.4.3.1 Immune cells, cytokines, and cancer

Tumour-associated macrophages (TAMs) are a major source of inflammatory cytokines the tumour microenvironment[189-191]. In contrast to normal macrophages, these tumour-recruited phagocytes are unresponsive to their normal regulatory mechanisms have been associated with six tumour-promoting extrinsic traits: chronic inflammation, matrix remodelling, tumour cell invasion, intravasation (invasion of blood vessels), angiogenesis, and distant metastasis[192]. Indeed, studies on the transition of carcinoma in-situ to invasive cancer have demonstrated the presence of TAMs at points of basement-membrane breakdown[192-193]. This finding, together with evidence from multi-photon imagining studies, suggests that tumours can manipulate TAMs to facilitate invasion and migration through the surrounding stroma[192, 194].

TAMs are also an important producer of VEGF, a key component for neoangiogenesis in tumours[179, 193, 195]. It has been proposed that cytokines produced by hypoxic cancer cells are responsible for the recruitment of macrophages[193, 196]. The induction of hypoxic inducible factor 2 alpha (HIF-2a) in recruited macrophages promotes the expression of vascular endothelial growth factor (VEGF) and angiogenesis[193, 195]. Furthermore, VEGF also acts as a chemo-attractant for macrophages thus creating a positive feedback loop for rapid vascularization in tumours[193].

In contrast to TAMs, T cells may play a promoting or suppressing role in cancer depending on their effector functions[186]. Clinical and experimental evidence indicates that the anti-tumour function of T lymphocytes is mediated by both cytotoxic mechanisms and cytokines[186]. Correspondingly, studies have demonstrated that increased activated cytotoxic T lymphocytes (CTLs) and Type 1 T-helper cells (Th1) are associated with better prognosis in pancreatic cancer. However, evidence has also suggested that several subsets of T cells (e.g. CD8, IFNγ-producing Th1 cells, and Th2 cells) are involved in tumour promotion, progression, or metastasis[186].

1.4.3.2 Pancreatic stellate cells, CCGFs, pancreatic fibrosis, and cancer

An important stromal component of pancreatic tumour microenvironment is pancreatic stellate cells (PSCs)[185]. PSCs are vitamin A containing spindle-shaped cells found in the pancreas that is capable of producing extracellular matrix (ECM)

proteins such as collagen, fibronectin, and desmin[186]. In normal pancreas, quiescent PCSs produce little ECM proteins and have undetectable levels of cytoplasmic alpha-smooth muscle actin (α-SMA). However, in response to oxidative stress from pancreatic injury, PSCs are transformed in to myofibroblast-like cells that express α-SMA[154]. In chronic pancreatitis and pancreatic cancer-associated inflammation, cytokines and growth factors released by immune cells stimulate PSC growth (e.g. PDGF) and the production of ECM proteins (TGF-β, IL-1β)[140]. In addition, activated PSCs in turn produce their own inflammatory mediators including MCP-1, IL-8, and RANTES, which generate a positive feedback for inflammation[185]. Interestingly, PSCs produce both ECM degrading enzymes (e.g. MMPs) and their inhibitors (tissue inhibitors of metalloproteinases, TIMPs)[185-186]. Indeed, the suppression of MMP-3 and MMP-9 by the TGF-β pathway and the promotion of MMP-2 expression (causes fibre deposition) by IL-6 and TGF-β1 have been observed in areas of pancreatic fibrosis[186]. Therefore, the regulation of pancreatic fibrosis is dependent on a delicate balance between ECM proteins production and degradation. Most importantly, recent studies have proposed that ECM remodelling is a key process, which facilitates the transition of dormant cancer cells to growth and in addition, angiogenic factors released by the remodelled ECM triggers the angiogenic switch thereby resulting in tumour growth and metastasis[197-198].

1.5 Statistical Modelling Methods

1.5.1 Background

This section is written with the aid of an article titled "Classification of Breast Cancer Cells Using JMP" by Dr Marie Gaudard *et al.* in 2009[199]. This article was later published in Chapter 9 the book *Visual Six Sigma: Making Data Analysis Lean*. However, the original article can still be viewed online at http://www.northhavengroup.com/documents/BreastCancer_WhitePaper_Current.pdf.

In Chapters 2 and 3 of the current thesis, a number of statistical modelling methods were employed to select and then combine candidate biomarkers into a single marker including stepwise regression model, multinomial logistic regression model, and artificial neural network model.

1.5.2 The Stepwise Regression (SR) Model

The SR model was used in Chapter 3 to select independent markers of PDAC amongst a large number of potential markers, which were statistically significant on univariate analysis. The SR model utilises the Wald/Score statistics to select independent predictors of outcome. The Wald/Score test in the SR model serves two purposes. Firstly, it compares the predictive capabilities of each variable with each other and secondly, the predictive capabilities are assessed in consideration of variables already entered into the SR mode. Therefore, whilst a variable may be a significant predictor of the outcome, it will not be entered into the SR model if previously selected variables can already identify the same samples and subsequently, this means that only variables with independent predictive abilities are selected by the SR model.

There are three variations of the SR model including Forward, Backward, and Combined. In forwards stepwise regression, the variable with the highest significance is entered into the SR model at each step providing it has a Wald/Score significance within the pre-set parameters (typically <0.05) and this would be iterated until no more variables can enter the model. In backwards stepwise regression, all variables are entered into the SR model initially and the variable with the lowest significances is removed from the SR model at each step if it has a Wald/Score statistic outside the pre-set parameters (typically >0.05). In combined

stepwise regression (used in Chapter 3), each step consists of a forwards component followed by a backwards component. This means that at each step, the variable with the highest significance will be entered into the SR model then a backwards step will be performed to remove any variables, which then became insignificant.

1.5.3 The Multinomial Logistic Regression Model

Logistic (or Binary) regression is a mathematical method used to predict a dichotomous outcome e.g. PDAC versus Controls. It is widely employed by researches to predict a dependent variable based on a number of continuous and/or categorical independent variables.

In the current thesis, M-LR was used to estimate the probability of PDAC (dependent variable) using the serum concentrations of candidate CCGFs selected by the stepwise regression model (continuous independent variables). This is achieved through the use of the logistic formula:

$$f(z) = Probability\ of\ PDAC = \frac{1}{1 + e^{z}}$$

Where

$$z = \beta_0 + \beta_1 \times CCGF_1 + \beta_2 \times CCGF_2 + \beta_3 \times CCGF_3 + \beta_4 \times CCGF_4$$

And

- β_0 represents the intercept constant
- β_1, β_2, β_3, and β_4 are regression coefficients for the corresponding CCGF
- This algorithm will generate, using serum concentrations of candidate CCGFs in a given sample, a probability value of PDAC ranging from 0 to 1
- A probability value of 0 represents a likely control sample where as a probability value of 1 represents a likely PDAC sample

1.5.4 Artificial Neural Network Model

The neural network model is so named because of its ability to "learn" in a similar fashion to neurons in the human body. There are three layers of neurons arranged in nets within the NN model: input layer, hidden layer (where processing occurs), and output layer. Each layer of neurons can be excited to a range of degrees (i.e. not binary). The input layer can be considered as sensory neurons, which react to a stimulus (CCGF concentrations) and generates an output signal that reflects the intensity of the input stimulus. This signal then enters the hidden layer where a predefined number of hidden nodes will again generate a weighted pattern of stimulus (i.e. a linear function [the weighted part] of logistic functions generated at each node), which is carried to the output layer. Finally, the inputs are summed (i.e. the estimated probability of PDAC) and compared to a threshold value (optimal cut-off) to determine their output (predicted PDAC or Control). Most importantly, perhaps the most important feature in NN model is its ability to learn from its mistakes. This is achieved by comparing the predicted out-put with the actual classification and by propagating the degree of error back through the whole network with the incorrectly classified connections down-weighted and correctly classified connections strengthened.

Again, this model will estimate the probability of a sample being PDAC based on the serum concentrations of the four CCGFs where a probability of 0 suggests a likely control sample whereas a probability of 1 suggests a likely PDAC sample.

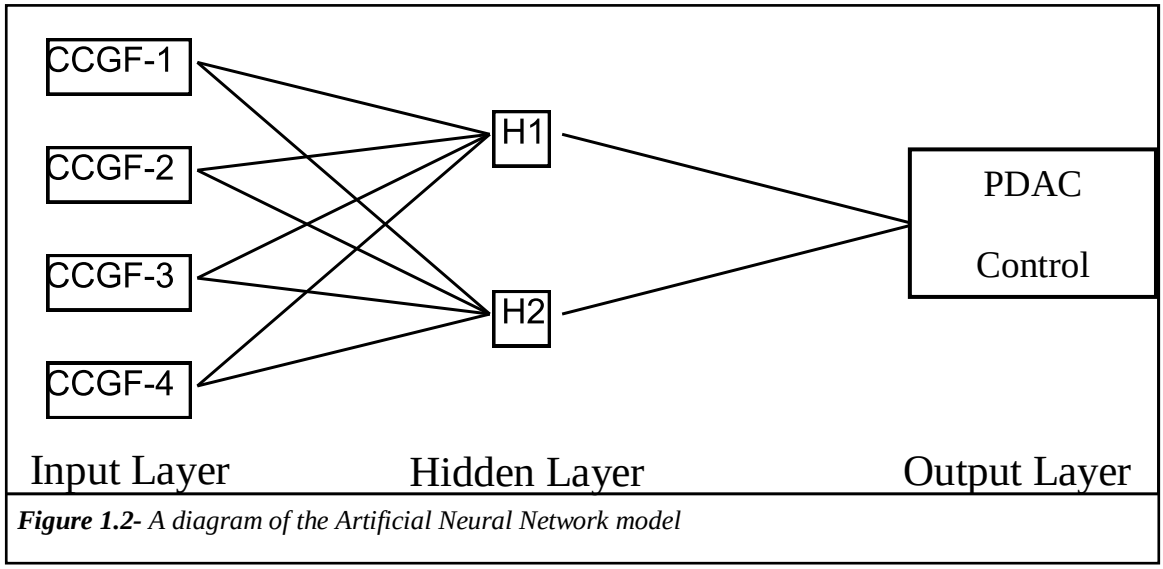


Figure 1.2- *A diagram of the Artificial Neural Network model*

1.6 Research Project Design

This research project takes a two-fronted approach to the discovery of proteomic biomarkers for pancreatic cancer whereby ***two independent studies are described in Chapter 2 and Chapter 3 of this thesis*** (***see Figure 1.3***).

Chapter 2 aim to determine the diagnostic potential of three serum-based candidate markers identified from a previous mass spectrometry experiment are determined by western blot analyses using pancreatic cancer and control samples from two database: Liverpool Pancreatic Cancer database (LPCD) and UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS) (***See section 2.3.2.2***).

Chapter 3 describes the use of multiplex cytokine assays for the discovery of novel biomarkers for pancreatic cancer. In this Chapter, the diagnostic potential of cytokines is examined as individual and combined markers. The results are validated using an independent validation sample set. (***See section 3.3.2***)

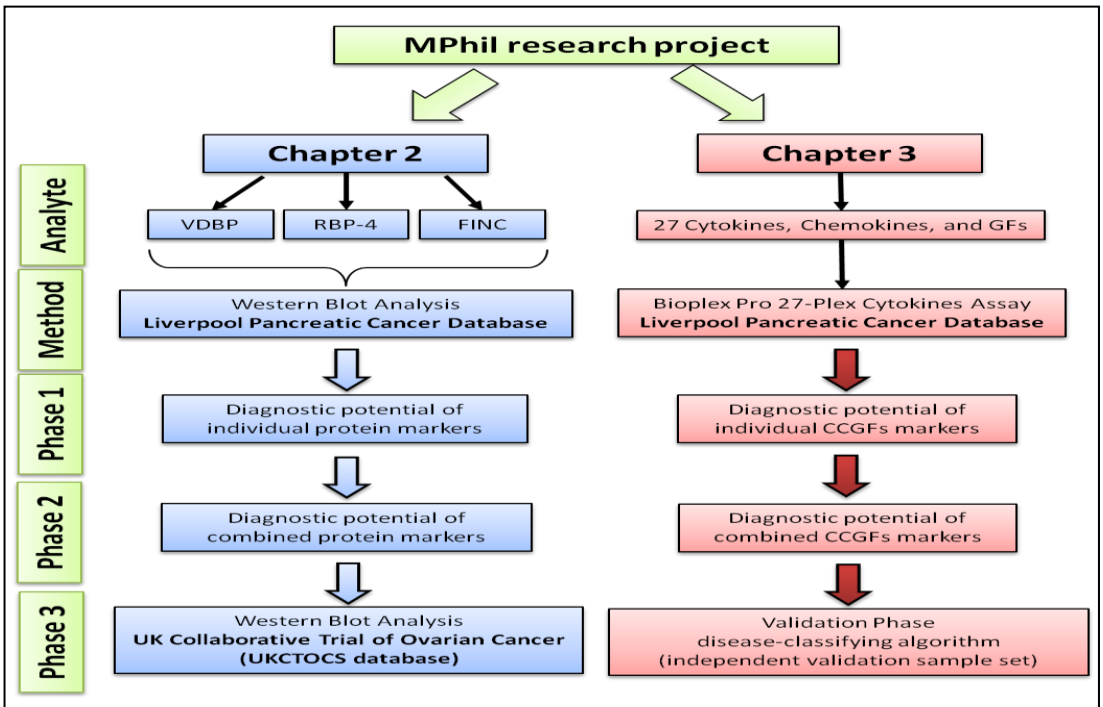


Figure 1.3- Design of the current MPhil Project. *The current thesis consists of two separate experimental studies. Each study is described in an individual Chapter and each study consists of three phases. The aim of* ***Chapter 2*** *is to validate the expression of three proteins identified by a previous iTRAQ experiment conducted in the department by western blotting: Vitamin D-Binding Protein (**VDBP**), Retinol-Binding Protein (**RBP-4**), and Fibronectin (**FINC**). In **Chapter 3** the use of a multiplex cytokine assay for the quantification of 27 serum cytokines, chemokines, and growth factors (**CCGFs**) is described. In addition, a disease-predicting algorithm was generated in the Discovery Phase and then validated by an independent sample set in the validation phase.*

Chapter 2

The diagnostic potential of VDBP, RBP-4, and FINC for pancreatic cancer

2.1 Background and Introduction

Blood serum is one of the most preferred sources for proteomic biomarker research because it offers a rich source of proteins and it is relatively non-invasive to acquire. However, despite much effort in the past few decades on the discovery of a novel serum diagnostic marker of pancreatic cancer, there has been little success in identifying a biomarker that is both sufficiently sensitive and specific to pancreatic cancer. Nonetheless, recent developments in proteomic technology such as iTRAQ and mass spectrometry have given new hopes to the ongoing search for a diagnostic marker or marker panel for use in the detection of pancreatic cancer.

Previous data from an iTRAQ experiment performed in the Division of Surgery, University of Liverpool, by Seonaid Murray and colleagues have identified a number of serum-based proteins, which are differentially observed in PDACs compared to Controls (chronic pancreatitis [CP] and healthy controls [HC]). Results from this experiment identified 254 proteins and quantified 234 proteins with greater than 95% confidence. A detailed description of the iTRAQ experimental protocol can be found in a previously published article by Tonack *et al.* [129]. Of the quantified proteins, forty-eight proteins showed a greater than three-fold difference in relative quantities between the early-stage pancreatic cancer group (EC) and control groups (CP and HC).

This chapter focuses on the validation of three proteins including Vitamin-D Binding Protein (VDBP, >3-fold ↓ in EC), Retinol Binding Protein-4 (RBP-4, >4-fold ↓ in EC), and Fibronectin (FINC, >3-fold ↓ in EC), which were found to be down regulated in the iTRAQ analysis (***Figure 2.1***). the serum levels of these proteins in individual patient samples were ascertained by western blotting. Furthermore, the accuracy of the three proteins for the diagnosis of pancreatic cancer against benign pancreaticobiliary disease (chronic pancreatitis and benign biliary obstruction, disease control [DC]) and healthy controls is demonstrated. In addition, the expression of these proteins was examined in pre-diagnostic serum samples collected from patients with PDAC up to 6 years before the confirmed diagnosis of PDAC was made.

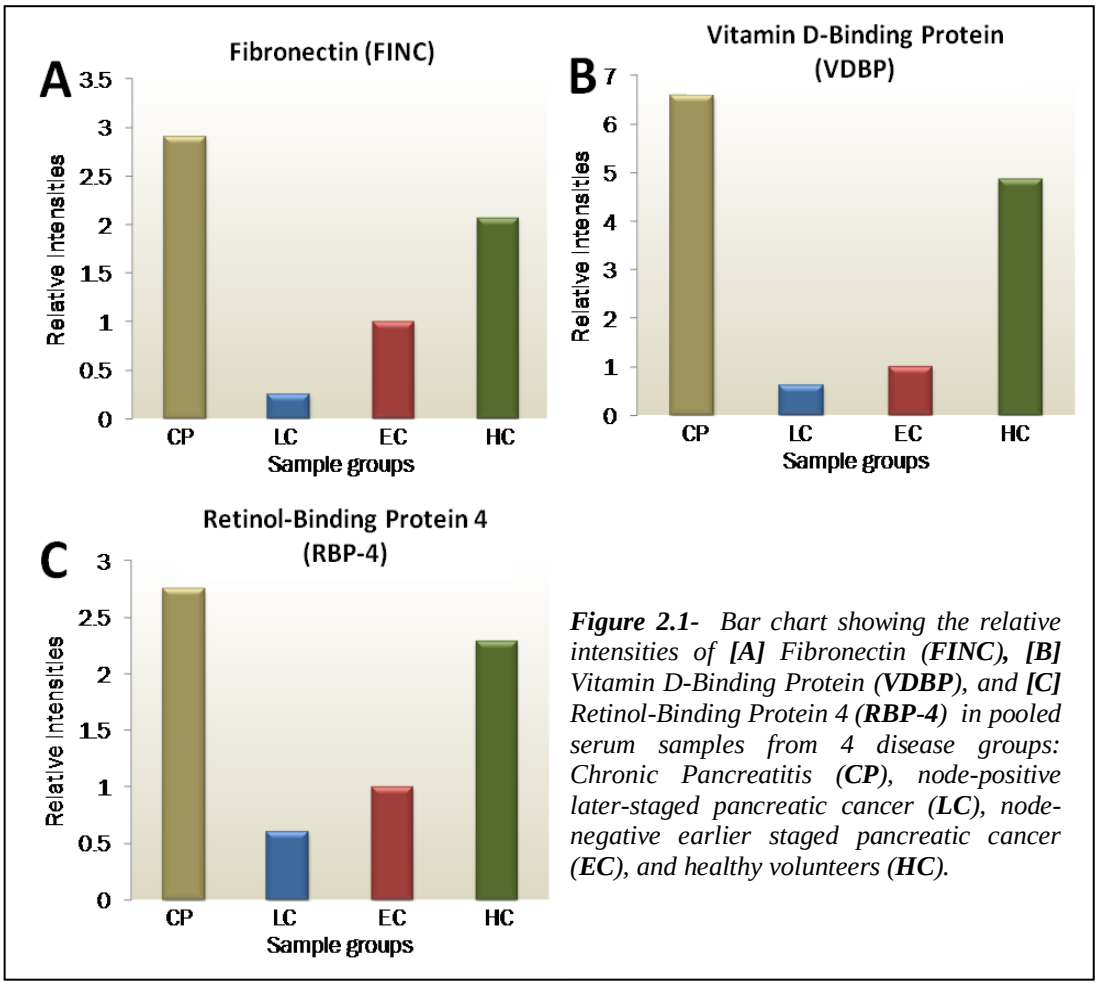


Figure 2.1- *Bar chart showing the relative intensities of* ***[A]*** *Fibronectin* ***(FINC), [B]*** *Vitamin D-Binding Protein* ***(VDBP)****, and* ***[C]*** *Retinol-Binding Protein 4* ***(RBP-4)*** *in pooled serum samples from 4 disease groups: Chronic Pancreatitis* ***(CP)****, node-positive later-staged pancreatic cancer* ***(LC)****, node-negative earlier staged pancreatic cancer* ***(EC)****, and healthy volunteers* ***(HC)****.*

2.1.1 Vitamin D binding protein

Vitamin D-binding protein is a member of the albumin superfamily and plays a number of physiological roles including the transportation of vitamin D and its metabolites, the binding of actin, neutrophil chemotaxis, and macrophage activation. VDBP was first identified as a group-specific component (Gc) protein by Hirschfeld in 1959[200] but was later given the name DBP after the discovery of its role in the transportation of vitamin D analogues[201]. In 1996, Yamamoto *et al.* [202] demonstrated in an animal study that VDBP also has a stimulatory effect on macrophage activity through its metabolite, which was named DBP-Macrophage Activating Factor (DBP-MAF). VDBP is a 58kDa glycosylated α-globulin consisting of three domains: two repeated homologous domains and a shorter domain [203-205]. It is the unique

orientations of these domains and the presence of oxygen-linked carbohydrate chains that give VDBP its unique physiological functions[203-205].

The plasma concentration of VDBP usually ranges from 0.2-0.5g/l and remains stable from birth[204]. Several physiological states have been shown to influence the serum level of VDBP including prolonged fasting, pregnancy, and high oestrogen states[206-207]. In addition, recent studies have demonstrated an inverse association between serum VDBP concentration and liver-related diseases such as liver cirrhosis, acute liver failure, and hepatocellular carcinoma[208-210].

2.1.2 Retinol-binding protein

Vitamin A has a number of biological roles including vision, maintenance of differentiated epithelia, mucus secretion, and growth[211]. The majority of vitamin A is stored in the liver and is usually transported by RBP in the systemic circulation as retinol (lipid alcohol form of vitamin A)[211-212]. RBP is a 21KDa molecule that consists of 182 amino acids arranged in a single polypeptide chain with three disulphide bonds[213-214]. It is produced by the liver and excreted by the kidneys and the plasma level of RBP typically ranges between 40-50 µg/ml[213-214]. However, a number of studies have shown that the systemic levels of RBP can be influenced by various diseases including malnutrition, liver diseases, and chronic renal diseases[215-220].

2.1.3 Fibronectin

Fibronectin is a multi-domain glycoprotein with a molecular weight of approximately 440kDa, which usually exists as a dimer composed of two identical monomers [221-223]. FINC has been shown to bind to a variety of biologically important molecules including a number of clotting-related molecules (such as heparin, collagen, and fibrin) and in addition, FINC is capable of binding to cell surfaces through integrins[223]. Fibronectin can be sub-classified into two broad categories: plasma fibronectin and cellular fibronectin. Hepatocytes are the main source of circulating plasma fibronectin, which has been reported to be in a closed and non-active state. In contrast, the active form of FINC (cellular fibronectin) is produced by multiple cell types (e.g. fibroblasts, epithelial cells, and macrophages) and is typically found in its insoluble form as a part of the extracellular matrix[222-223].

2.2 Study Aims

The aims of the current chapter:

1. To validate the difference in the serum levles of VDBP, RBP-4, and FINC between PDAC and Controls observed in the original iTRAQ experiment performed by Seonaid Murray (Division of Surgery and Oncology, UOL)

2. To determine the diagnostic potential of VDBP, RBP-4, and FINC for pancreatic cancer

3. To determine the expression of VDBP and RBP-4 in pre-diagnostic serum samples
4. To determine the expression of VDBP, RBP-4, and FINC in pancreatic cancer cell-lines

2.3 Materials and Methods

2.3.1 Study design

The expression of VDBP, RBP-4, and FINC in pancreatic cancer and control subjects was examined in three phases (***Figure 2.2***). Each phase involved western blotting followed by relative quantification of the blots through densitometry analysis. In Phase-I, the levels of VDBP, RBP-4, and FINC in the pooled iTRAQ samples (***section 2.3.2.1***) were validated using the individual samples. In Phase-II, the expressions of VDBP, RBP-4, and FINC were assessed using an independent sample set consisting of 60 samples from the LPCD (***section 2.3.2.1***). In addition, the diagnostic potential of these proteins as individual and combination markers for early-staged pancreatic cancer were determined. Protein markers, which were differentially expressed in Phase-II were further assessed for their potential as a screening marker using the UKCTOCS pre-diagnostic serum samples (Phase-III, ***2.3.2.2***). Western blot analyses in the current study were performed ***in triplicate*** to maximise the accuracy of the results.

Supplementary to this experiment, western blot analyses were performed to determine whether VDBP, RBP-4, and FINC are expressed in five pancreatic cancer cell lines, one hepatocellular carcinoma cell line and one human embryonic kidney cell line (***section 2.3.3***).

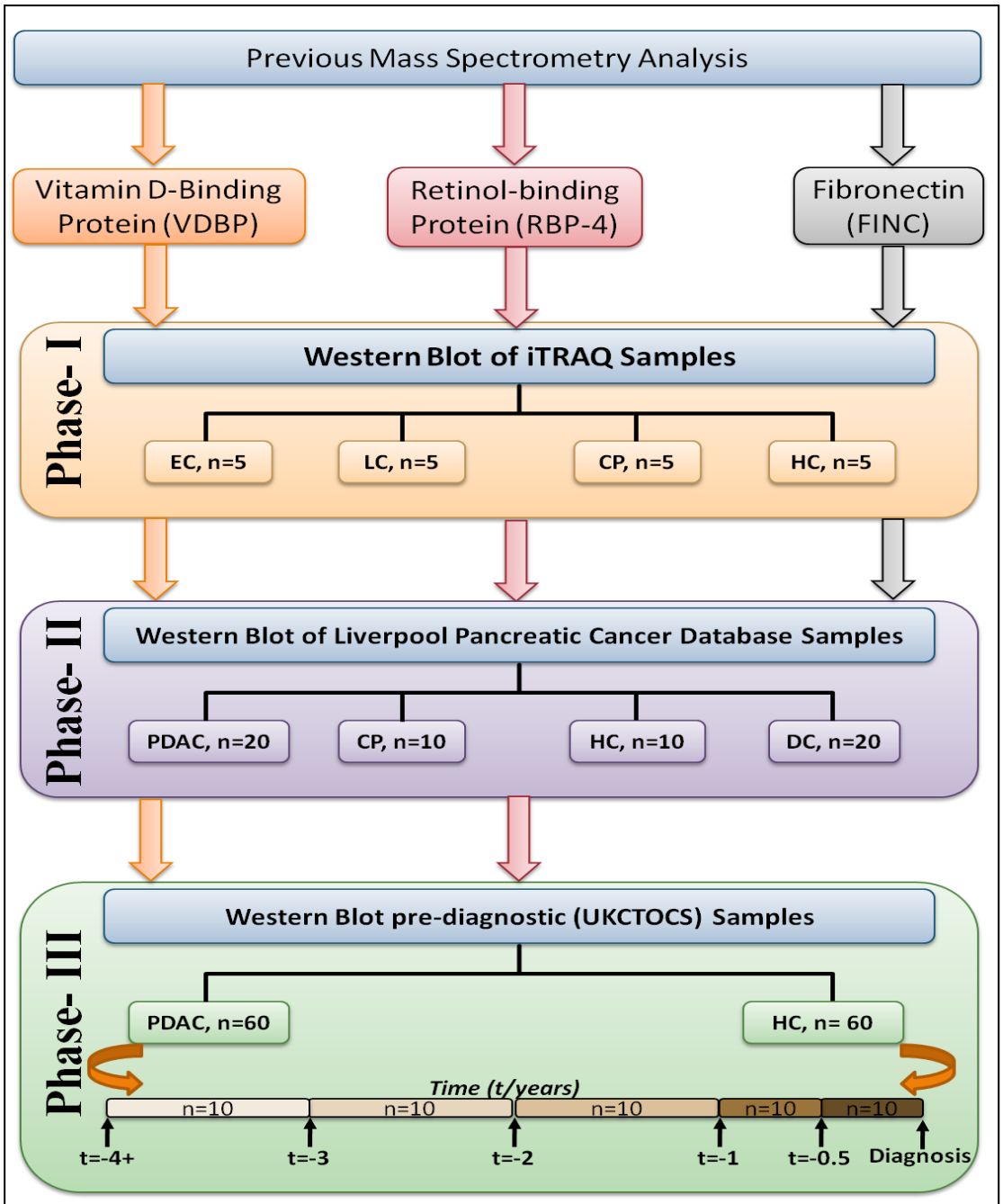


Figure 2.2- *Study design for Chapter 3 of the current thesis. The current study consists of three phases and each phase involves western blot analysis of a unique sample set. In* ***Phase-I****, the expression of VDBP, RBP-4, and FINC were validated using 5 samples from of the following disease groups: node-negative earlier-staged PDAC (****EC****), node-positive later-staged PDAC (****LC****), chronic pancreatitis (****CP****), and healthy volunteers (****HC****).* ***Phase-II*** *of the study aimed to validate the results in the original iTRAQ experiment and Phase-I by using 60 samples from individuals with PDAC, CP, HC, and disease controls (****DC****).* ***Phase-III*** *aimed to determine the value of significant markers (VDBP and RBP-4) in screening for pancreatic cancer using samples from the UKCTOCS database.*

2.3.2 Patients and samples

Approval for the current study was obtained from the relevant research ethics committee including the multicentre research ethics committee (MREC) for the use of the serum samples from the Liverpool Pancreatic Cancer Database and the Central Office of Research Ethics Committee (COREC) for the use of UKCTOCS samples. Informed consent was obtained from all individuals involved.

2.3.2.1 Phase I and II Liverpool pancreatic cancer database (LPCD)

Pre-operative serum samples were prospectively collected at the Royal Liverpool University Hospital from patients with resectable PDAC, chronic pancreatitis, benign biliary obstruction, and healthy volunteers between 1996 and 2010. Serum samples were collected in Sarstedt Monovette tubes (Sarstedt Ltd, Leicester, UK) and allowed to clot at 4℃ for 15 minutes. The serum fraction was acquired by centrifugation at 800-x g for 10 minutes and was then aliquoted into cryotubes (Nunc GmbH & co KG., Thermo Fisher Scientific, Langenselbold, Germany). All samples were stored at minus 80℃ until further use.

Phase-I patient demographical data (iTRAQ)

Demographical data of the twenty individuals involved in Phase-I of the current study is presented in ***Table 2.1***. There were sixteen males, 3 females, and one individual whose gender was not recorded. The median ages for patients with node negative earlier-stage PDAC (EC), node-positive later-stage PDAC (LC), and CP were 66, 71, and 63, respectively. There were four individuals with a confirmed history of diabetes in the cancer groups (EC and LC) and none in the control groups. However, some demographical data were not available due to privacy and confidentiality reasons.

Table 2.1- *Summary of patient Demographics in Phase-I (iTRAQ)*

Parameters	**Disease Groups**			
	CP (n=5)	**LC (n=5)**	**EC (n=5)**	**HC (n=5)**
Age (median/year)	63	71	66	>50
Gender				
Male/Female (n)	4/0	2/3	5/0	5/0
Not recorded	1	0	0	0
History of Diabetes (n)				
Yes/No (n)	0/4	1/3	3/0	-/-
Unknown	1	1	2	5

Abbreviations*: Chronic pancreatitis (**CP**); Node positive later-staged pancreatic cancer (**LC**); Node negative earlier-staged pancreatic cancer (**EC**); Healthy volunteers (**HC**).*

Phase II patient demographical data (LPCD)

Demographical data of the 60 individuals involved in Phase-II of the current study is presented in ***Table 2.2.*** Thirty-two male and twenty-eight female participants were involved in the current Phase, of which, 20 had PDAC and 10 had CP, 20 had biliary obstruction (DC), and 10 were HC. The median ages were 69, 51, 64, and 36 years for the PDAC, CP, DC, and HC groups, respectively. Again, some demographical data were not available due to privacy and confidentiality reasons.

Table 2.2- *Summary of patient Demographics in Phase-II (Liverpool database)*

Parameters	**Disease Groups**			
	PDAC (n=20)	**CP (n=10)**	**DC (n=20)**	**HC (n=10)**
Age (median/year)	69	51	64	36
Gender				
Male/Female (n)	10/10	2/8	15/5	5/5
Median CA19-9 (range)	307 (10-10061)	12 (7-50)	21 (0-1402)	6 (2-12)
History of Diabetes (n)				
Yes/No (n)	4/15	0/9	2/16	-/-
Unknown	1	1	2	10

Abbreviations*: Chronic pancreatitis (**CP**); Node positive later-staged pancreatic cancer (**LC**); Node negative earlier-staged pancreatic cancer (**EC**); Healthy volunteers (**HC**).*

2.3.2.2 Pre-diagnostic serum samples- UKCTOCS

The UK collaborative trial of ovarian cancer screening (UKCTOCS) is a multicentre trial involving 13 regional centres aimed to establish the impact of ovarian cancer screening on ovarian cancer mortality by comparing disease mortality in the screened and control groups[224]. Serum samples were prospectively collected from 50,000 postmenopausal women aged 50-74 with no past medical history of malignant disease or familial predisposition to ovarian cancer in the screened group[224]. Participants in this group were sub-classified into low, intermediate and high risk for the development of ovarian cancer according to their serum CA-125 levels and were followed up prospectively at regular intervals: annually for low risk, 3-monthly for intermediate risk, and referral for further investigations for participants with elevated CA-125[224]. During each follow-up session, a serum sample was collected and the development of malignancies was recorded. Serum samples were collected in Greiner gel tubes (Greiner Bio-One 455071, Stonehouse, UK) and transported overnight at ambient temperature to the central laboratory where samples were processed within 56 hours of venepuncture (otherwise discarded)[224]. Separation of the serum layer was achieved by centrifugation at 4,000 rpm for 10 minutes. Excess serum was aliquoted into 500µl straws, which were then heat-sealed and stored in liquid nitrogen until transfer to a cryonic biorepository[224].

At the University of Liverpool, we collaborated with the biomarker group at the University College London to obtain over 300 pre-diagnostic (up to 6 years pre-diagnosis) serum samples from UKCTOCS participants who developed pancreatic cancer within the trial period accompanied by an equal number of healthy control samples matched for time before sample processing and storage time.

Phase III patient demographics (UKCTOCS)

One-hundred and twenty serum samples from the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS) were analysed in Phase III ***(Table 2.3)***. Sixty pre-diagnostic serum samples from post-menopausal women who developed pancreatic cancer during the course of the UKCTOCS trial (pre-pancreatic cancer, PPC) and sixty-paired healthy control samples (HC) matched for gender, storage, and processing time were used in this Phase. Samples were stratified into six time-categories according to the duration between sample collection and diagnosis: 0-0.5, 0.5-1, 1-2, 2-3, 3-4, and over 4 years. Each time group therefore, consists of 10 PPC samples and 10 matched HC samples.

Table 2.3- *Summary of demographics for UKCTOCS samples stratified by time-category*

Parameters	Time Category/ year											
	0-0.5		0.5-1		1-2		2-3		3-4		4+	
	HC	PPC	HC	PPC	HC	PPC	HC	PPC	HC	PPC	HC	PPC
Number (n)	10	10	10	10	10	10	10	10	10	10	10	10
Age/ years												
Median	61	65	62	70	60	61	61	61	58	61	61	65
Range	52-75	57-78	53-72	57-73	53-72	53-72	51-73	52-71	52-71	57-75	51-72	52-73
Delay in sample processing/ hr												
Median	22	22	20	21	23	23	24	24	23	23	20	20
Range	19-25	20-26	19-24	19-24	19-46	20-46	18-25	18-25	21-43	21-44	4-26	2-22

Abbreviations: *Healthy control (**HC**); pre-pancreatic cancer (**PPC**)*

2.3.3 Cancer cell lines, cell culture and lysate preparation

2.3.3.1 Cell lines and cell culture

Five human pancreatic ductal adenocarcinoma cell lines (PANC-1, ASPC-1, BXPC-3, CFPAC-1, and FAMPAC), one hepatocellular carcinoma cell line (HEPG-2), and one human embryonic kidney cell line (HEK) were were screened for mycoplasma contamination before cell culturing commenced. All cell lines except HEPG-2 were cultured in T-75 flasks at 37℃ and 5% CO_2 in 20 ml of RPMI 1640 media supplemented with 10% foetal calf serum, 2mM glutamine, 100U/ml penicillin, and 100U/ml streptomycin (Sigma, Poole, UK). For culture of HEPG-2 cell lines Dulbecco's Modified Eagle Medium supplemented with 10% foetal calf serum, 2mM glutamine, 100U/ml penicillin, and 100U/ml streptomycin (Sigma, Poole, UK) was used instead of RPMI. When confluent, cancer cells were detached from the flask using 2ml trypsin (sigma, Poole, UK). Once fully detached, 8ml of culture media was added to the flask and the culture was split in the following ratio: 1ml for further culture, 9ml for cell lysate preparation.

2.3.3.2 Lysate preparation

Cells collected from culture splitting were centrifuged for 5 minutes at 1,200 rpm. The resulting pellet was washed three times in PBS solution and lysed using 100µl of lysis buffer (0.15 M NaCl, 5mM EDTA, 10mM Tris-Cl pH 7.4, 1% Triton X-100, and protease inhibitor cocktail; all from Sigma-Aldrich) at 4℃ for 10 minutes then sonicated. The protein concentration of the lysate was then determined using the Bradford Assay. The lysate was then stored at minus 80℃ until required.

2.3.3.3 Measuring lysate protein concentration- Bradford Assay

The Bradford reagent assay (Bio-Rad, Hemel Hempstead, UK) was employed to determine the protein concentration of cancer cell lysates. The Bradford reagent contains a dye (brilliant blue G), which forms a complex with the proteins in the reconstituted samples resulting in a shift in the absorption wavelength of the dye from 465 to 595 nm. Therefore, the amount of absorption is directly proportional to the protein concentration in the solution.

Briefly, six reference standards ranging from 1 µg/µl to 10 µg/µl were constituted using bovine serum albumin (Sigma Aldrich, Gillingham, UK). Then 2 µl of standard or sample were added to 798µl of MilliQ water in a 1.5ml Eppendorf microfuge tube followed by 200µl of Bradford Reagent (Bio-Rad, Hemel Hempstead, UK). The mixture was incubated at room temperature for 10 minutes then analysed by spectrophotometry at a wavelength of 595nm. The protein concentration of each sample was calculated using a linear equation constructed from the reference standards, which converts Bradford assay values to protein concentration (µg/µl). Finally, the volume of lysate required for 15µg of protein was calculated for each lysate.

2.3.4 Western Blot Analysis

2.3.4.1 SDS-polyacrylamide gel preparation

The composition of SDS-polyacrylamide gel required for the analysis of VDBP, RBP-4, and FINC are 12%, 15%, and 6%, respectively. The running gel solution was prepared according to ***Table 2.4***, and immediately poured into a gel cast (BioRad, Hemel Hempstead, UK) and allowed to set. A stacking gel solution was added to the cast followed by a 10-well or 15-well comb and allowed to set.

Table 2.4- *Recipe for 10 ml of Resolving Gels and 4ml of Stacking Gel for SDS-PAGE*

Components required /ml	**Gel Density**			
	6% (FINC)	**12%** (VDBP)	**15%** (RBP-4)	**5%** (Stacking)
H_20	5.3	3.3	2.3	2.7
30% acrylamide mix	2.0	4.0	5.0	0.67
1.5M Tris (pH 8.8)	2.5	2.5	2.5	0.5
10% SDS	0.1	0.1	0.1	0.04
10% ammonium persulphate	0.1	0.1	0.1	0.04
TEMED	0.01	0.006	0.005	0.005

Abbreviations*: Sodium dodecyl sulphate (**SDS**); Polyacrylamide Gel Electrophoresis (**PAGE**); fibronectin (**FINC**); retinol-binding protein 4 (**RBP-4**); vitamin D-binding protein (**VDBP**); Tetramethylethylenediamine (**TEMED**)*

2.3.4.2 Sample preparation

Four micro-litres of diluted serum samples/standards (***Table 2.5***) or 10µg of cell lysate were denatured and ionised by adding 3µl of 5x reducing sample buffer (300mM Tris pH 6.8, 50% Glycerol, 10% SDS, 0.05% Bromophenol blue, 5% Dithiothreitol [DTT]) and made up to a final volume of 10µl using 1x reducing sample buffer in a 0.5ml Eppendorf microfuge tube. The mixture was heated at 95℃ for 15 minutes then cooled on ice for 3 minutes. Finally, the mixture was flash spun at 13,000g to bring down condensation prior to gel loading.

Table 2.5- *Preparation of samples and standards for western blot*

Analyte	**Protein Dilution**	**Sample Volume (µl)**	**Standard volume (µl)**		
			Std 1	**Std 2**	**Std 3**
VDBP	1:100	4	2	**4***	6
RBP-4	1:10	4	2	**4***	6
FINC	1:10	4	4	6	**10***

*******The volume of serum standard was used in the densitometry analysis of serum samples*

2.3.4.3 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

Prior to performing polyacrylamide gel electrophoresis, the wells were flushed with running buffer to remove any un-polymerised polyacrylamide. Five micro-litres of PageRuler™ Pre-stained Protein Ladder (Fermentas, York, UK; ***Figure 2.3***) was loaded into the first well and 10µl of prepared standard/samples were loaded to each subsequent well. Electrophoresis was performed using the Mini-PROTEAN Tetra Electrophoresis System (Bio-Rad, Hemel Hempstead, UK) set at 25-50mA constant (depending on the number of gels) until the dye front completely left the gel (approximately 1-1.5 hour).

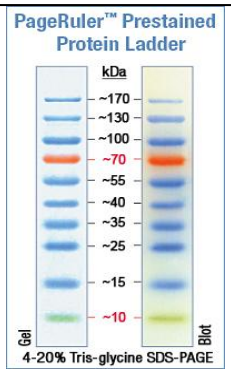


Figure 2.3- *Pre-stained protein ladder used in western blot analysis*

Proteins in the gel were transferred onto a nitrocellulose membrane using the Mini-Transblot pack (Bio-Rad, Hemel Hempstead, UK) set at 110V (constant voltage) for 1 hour. The membrane was incubated at room temperature with 20ml of 5% milk diluted in phosphate buffer solution with 1% Tween-20 (5% milk-PBST solution) for 2 hours followed by incubation with the appropriate concentration of primary antibody diluted in 5% milk-PBST solution for 16 hours at 4℃ (***Table 2.6***). The membrane was then washed six times at 10-minute intervals (1 hour) with PBST followed by incubation with the appropriate secondary antibody-HRP diluted in 5% milk-PBST for 1 hour. After a further six washes at 10-minute intervals, the membrane was incubated with Enhanced Chemiluminescence (ECL) substrate (Western Lightning® Plus, PerkinElmer, Cambridge, UK) for 4 minutes followed by chemiluminescence detection by x-ray films (Fisher Scientific, Leicestershire, UK) or the Kodak digital imaging machine. For subsequent western blotting of β-Actin in protein lysates, the membrane was incubated at 60℃ for 30 minutes in stripping buffer, rinsed with PBST, and then incubated with primary β-Actin antibody (1:10,000 dilution) for 16 hours. Washes, secondary antibody incubation, and detection were carried out as described above. Each gel was repeated twice to obtain three sets of results per gel, per antibody.

Table 2.6*- Summary of sample preparation, primary antibody, and secondary antibody information*

Analyte	**MW** (kDA)	**Primary anti-analyte antibody**			**Secondary anti-primary antibody-HRP**		
		Company	**Dilution**	**Description**	**Company**	**Dilution**	**Description**
VDBP	53	Abcam	1:30,000	Rabbit Monoclonal	Dako	1:3,000	Polyclonal Goat Ant-rabbit Ig/HRP
RBP-4	25	Santa Cruz	1:5,000	Mouse Monoclonal	Dako	1:3,000	Polyclonal Goat Ant-mouse Ig/HRP
FINC	220	Santa Cruz	1:20,000	Mouse Monoclonal	Dako	1:3,000	Polyclonal Goat Ant-mouse Ig/HRP
Beta-Actin	42	Sigma-Aldrich	1:10,000	Mouse Monoclonal	Dako	1:3,000	Polyclonal Goat Ant-mouse Ig/HRP

Abbreviations: *vitamin D-binding protein (**VDBP**); retinol-binding protein (**RBP-4**); fibronectin (**FINC**); Immunoglobulin (**Ig**); expected molecular weight (**MW**)* ***Note:*** *5% Milk in Phosphate Buffer Solution with 1%-Tween 20 used as blocking solution;*

2.3.5 Relative quantification of western blots

The Kodak digital imaging machine and Kodak molecular imaging software (Carestream molecular Imaging, Woodbridge, USA) were used to analyse the Western Blot images. The relative quantity of protein in each band was assessed using densitometry analysis. To define the profiles a rectangular box was defined arbitrarily with the same width as the widest band on the film and longer than the

largest band. This same box was used to measure all bands in the image, resulting in a mean intensity value for all bands. To account for gel-to-gel variation relative intensities for each sample were calculated by dividing sample mean intensity by the internal control mean intensity.

2.3.6 Statistical Analysis

2.3.6.1 General statistics, univariate, and multivariate analyses

All continuous data (e.g. Age, serum analyte levels) were classified as non-parametric data and were summarised using median and range. The Kruskal-Wallis test was used to determine whether there is a significant difference in non-parametric distribution between multiple (>2) groups (e.g. age versus PDAC, CP, HC, and DC). For univariate analysis, the Wilcoxon's test was used to assess any difference in the distribution of a non-parametric variable (e.g. age) between two groups and the Fisher's Exact test was used to assess the difference between two categorical variables (e.g. gender versus PDAC/Control). Variables that were significant on univariate analysis were further tested using multivariate analysis to identify any independent variables. For univariate and multivariate analyses, a p-value of <0.05 was considered statistically significant.

In Hierarchical Cluster Analysis, data from each variable was first standardised by mean and standard deviation to give a value between -1 and 1. The standardised values were then analysed by Ward's minimal variance method and represented graphically in a heat map.

Data from the current study were also graphically represented using box plots, mosaic plots, logistic plots, and receiver-operator characteristic (ROC) curves.

2.3.6.2 Correlation analysis

The correlation between age and the relative serum concentration of VDBP, RBP-4, and FINC were assessed using Kendall Tau (τ) multivariate correlation analysis for non-parametric tied data. The Kendall τ coefficient ranges from -1 to 0 to 1 where a τ coefficient of -1 or 1 represents a perfect negative or positive correlation, respectively, and a τ coefficient of 0 indicates that there is no correlation. The associated p-value represents how confident the test is that the actual τ value would

be more negative or more positive than then calculated τ value. In the current study, a Kendall Tau coefficient of >0.5 was regarded as significant.

2.3.6.3 Diagnostic potential of biomarkers for PDAC

The diagnostic accuracies of biomarkers were assessed using ROC Area Under Curve (AUC) and a ROC-AUC of >0.70 was considered statistically significant. In ***sections 2.4.2.6 to 2.4.2.10***, the accuracies of VDBP, RBP-4, and FINC for diagnosing PDAC were independently assessed against the each of the following groups: HC, CP, DC, and all controls (HC, CP, and DC together). Additionally for each group, the diagnostic accuracy of the three proteins as a single combined marker of pancreatic cancer was assessed.

The Multinomial Logistic Regression (M-LR) model was used to combine candidate markers into a single marker by generating a disease-predicting algorithm, which estimates the probability of PDAC based on the relative serum concentrations of VDBP, RBP-4, and FINC. The estimated probability of PDAC ranges from 0-1 where an estimated PDAC probability value of 0 indicates a likely control sample whereas a value closer to 1 indicates a likely pancreatic cancer sample.

2.3.6.4 Software for statistical analyses

All statistical analyses were performed using JMP version 8.02 (SAS, Buckinghamshire, UK). In addition, Microsoft Excel 2007 (Microsoft Limited, Berkshire, UK), was used to graphically represent the data.

2.4 Results

2.4.1 Phase I- Validation of iTRAQ results by Western Blotting

The aim of Phase I is to validate the serum expression of VDBP, RBP-4, and FINC in individual samples used by the original iTRAQ experiment (which consisted of pooled samples from 5 patients per disease group) by Western Blotting. An example of Western Blot image for VDBP, RBP-4, and FINC is shown in ***Figure 2.4*** and the expected molecular weight of each protein can be found in ***Table 2.6***. Results from densitometry analyses were expressed as a ratio relative to the intensity of the control reference band. The relative intensities (**RI**) of FINC, RBP-4, and VDBP for each disease group in the original iTRAQ experiment were comparable to the Western Blot results (***Figure 2.5A-F***). Furthermore, the RIs of all three proteins were decreased in the cancer groups (EC and LC) compared to the control groups (CP and HC). In particular, the pattern of FINC expression in the western blot analysis is almost identical to the original iTRAQ results. The box plots in ***Figure 2.6*** confirm this pattern and therefore suggest that the results were not due to the effects of outliers. However, results from Phase-I were not sufficiently powered to show any significance on further statistical analysis (***Figure 2.6***). Therefore, all three proteins were analysed in Phase-II of the current study, which involves a larger sample set from the Liverpool Pancreatic Cancer Database.

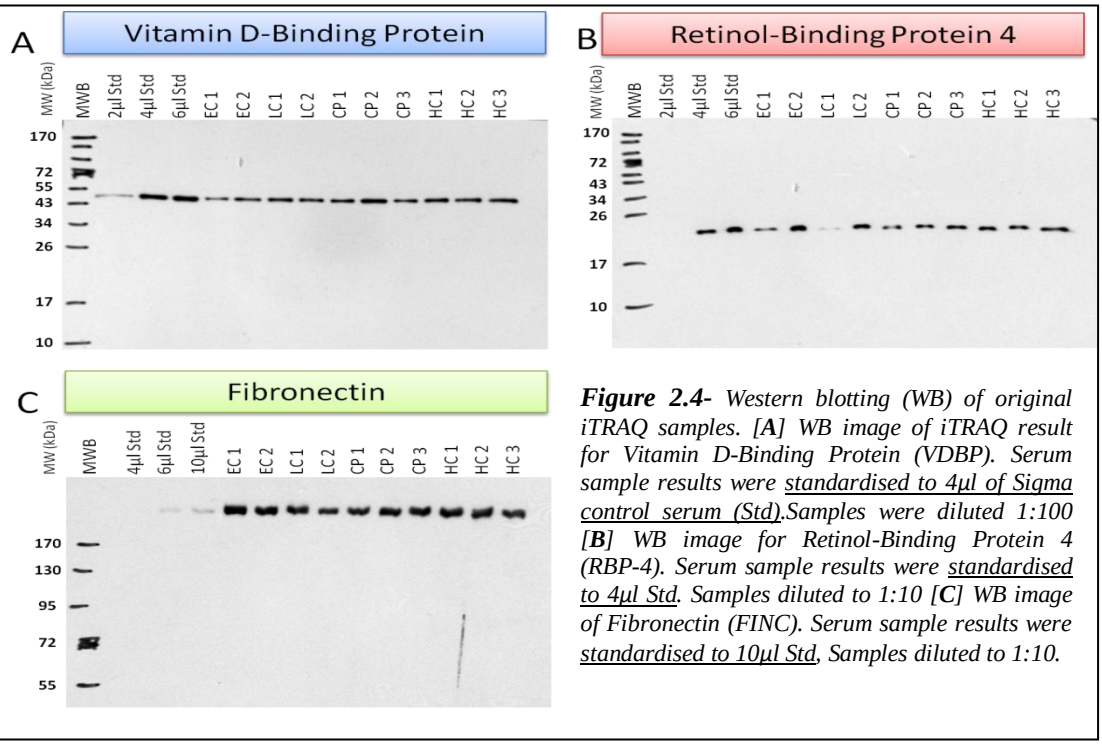
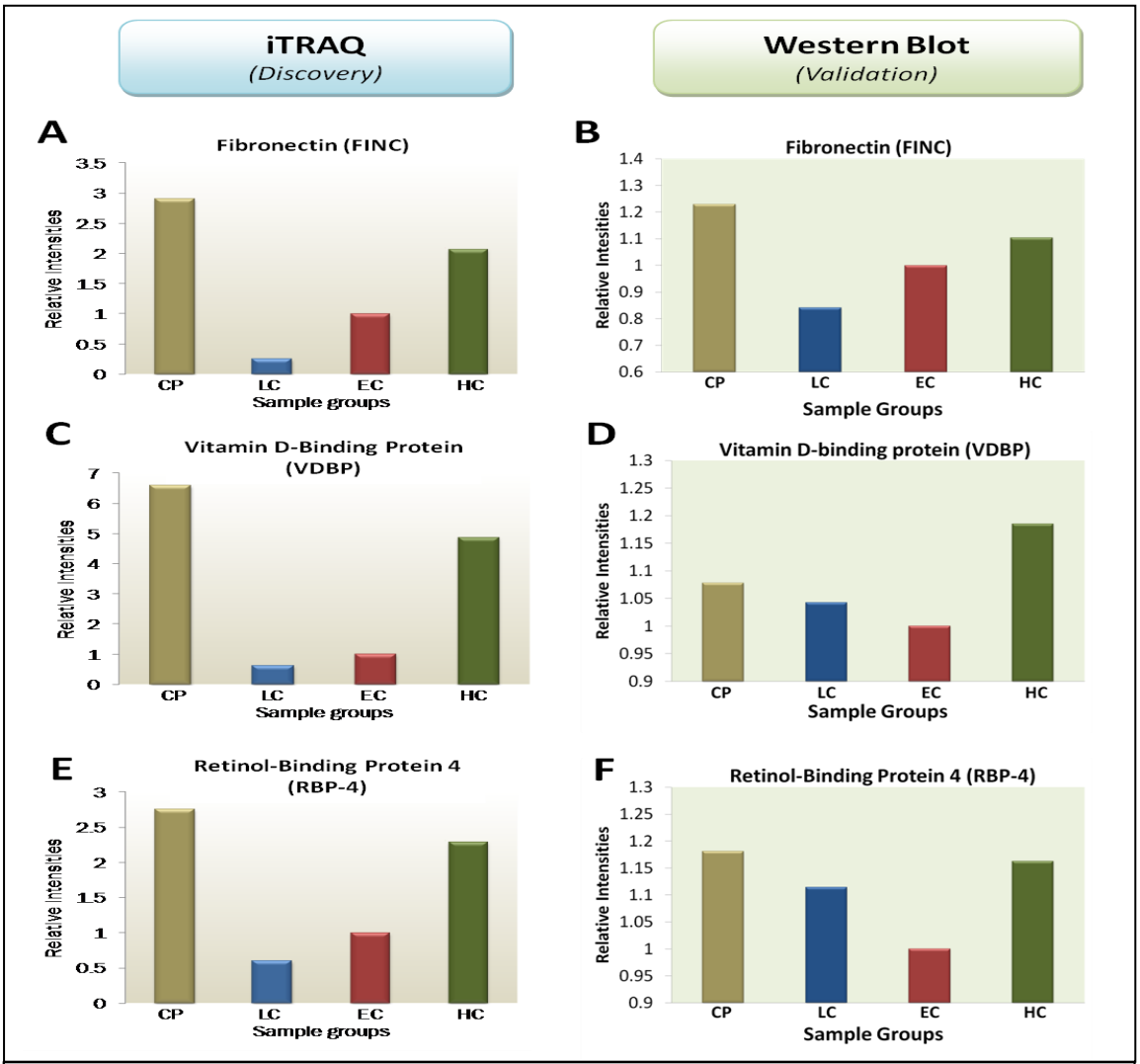


Figure 2.4- *Western blotting (WB) of original iTRAQ samples. [**A**] WB image of iTRAQ result for Vitamin D-Binding Protein (VDBP). Serum sample results were standardised to 4µl of Sigma control serum (Std).Samples were diluted 1:100 [**B**] WB image for Retinol-Binding Protein 4 (RBP-4). Serum sample results were standardised to 4µl Std. Samples diluted to 1:10 [**C**] WB image of Fibronectin (FINC). Serum sample results were standardised to 10µl Std, Samples diluted to 1:10.*



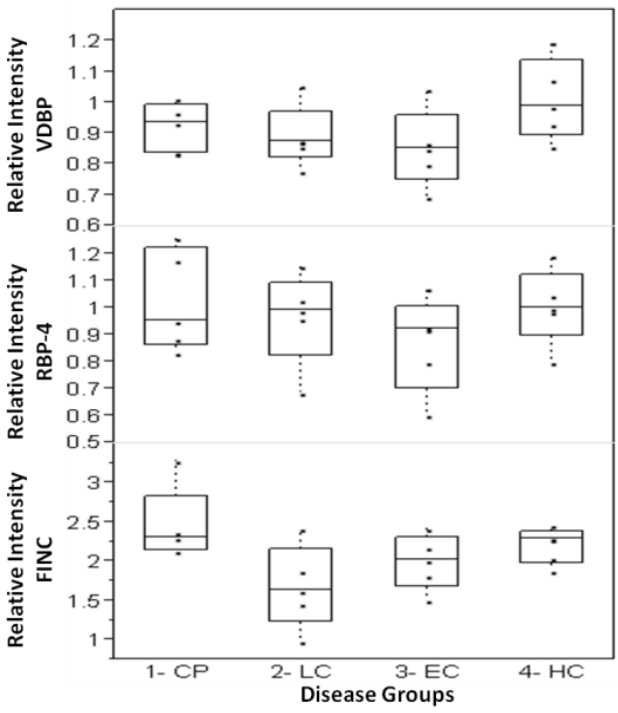


Figure 2.5 (Above)- *[**A-F**]Bar charts showing the mean relative intensity of the chronic pancreatitis (**CP, n=5**), lymph node positive later-staged pancreatic cancer (**LC, n=5**), lymph node negative earlier-staged pancreatic cancer (**EC, n=5**), and healthy volunteers (**HC, n=5**) in the original iTRAQ experiment (**A,C, and E;** consists of pooled samples from the 5 patients in each disease group) and in Phase I (**B,D, and F,** samples were analysed individually by Western Blotting).*

Figure 2.6 (left)- *Box plots showing the relative intensities of VDBP, RBP, and FINC in each disease group. Relative intensities were obtained following densitometry scan of Western Blot images then comparing the density of each sample with an internal control on the same gel.*

2.4.2 Phase II- Further Validation by Western Blot (Liverpool samples)

2.4.2.1 Patient demographics and clinical characteristics

The aim of Phase II is to further verify the serum expression of VDBP, RBP-4, and FINC in a larger sample set consisting of 20 PDAC, 10 CP, 20 DC, and 10 HC individuals (***Table 2.7***). As expected, the serum level of CA19-9 in the PDAC group is significantly higher than the control group (Wilcoxon's Test, p=0.0003; ***Table 2.7***). Furthermore, a significant difference in age between the PDAC group and the Control group was observed (69 years versus 53 years; Wilcoxon's Test, p=0.0018; ***Figure 2.7***). However, we did not find any statistical difference in gender between the two groups (Fisher's Exact Test, P>0.05). In view of the fact that the current study involves a relatively small sample size and that there is a large number of missing data for diabetes especially in the HC group, the correlation between PDAC and diabetes cannot be fully elucidated.

Table 2.7- Summary of patient Demographics in Phase-II (Liverpool database)

Parameters	Disease Groups				P-value
	PDAC (n=20)	CP (n=10)	DC (n=20)	HC (n=10)	
Age (median/year)	69	51	64	36	**0.0018**
Gender					
Male/Female (n)	10/10	2/8	15/5	5/5	N.S.
Median CA19-9 (range)	307 (10-10061)	12 (7-50)	21 (0-1402)	6 (2-12)	**0.0003**
History of Diabetes (n)					
Yes/No (n)	4/15	0/9	2/16	-/-	N.S.
Unknown	1	1	2	10	

***Abbreviations**: Chronic pancreatitis (**CP**); Node positive later-staged pancreatic cancer (**LC**); Node negative earlier-staged pancreatic cancer (**EC**); Healthy volunteers (**HC**).*

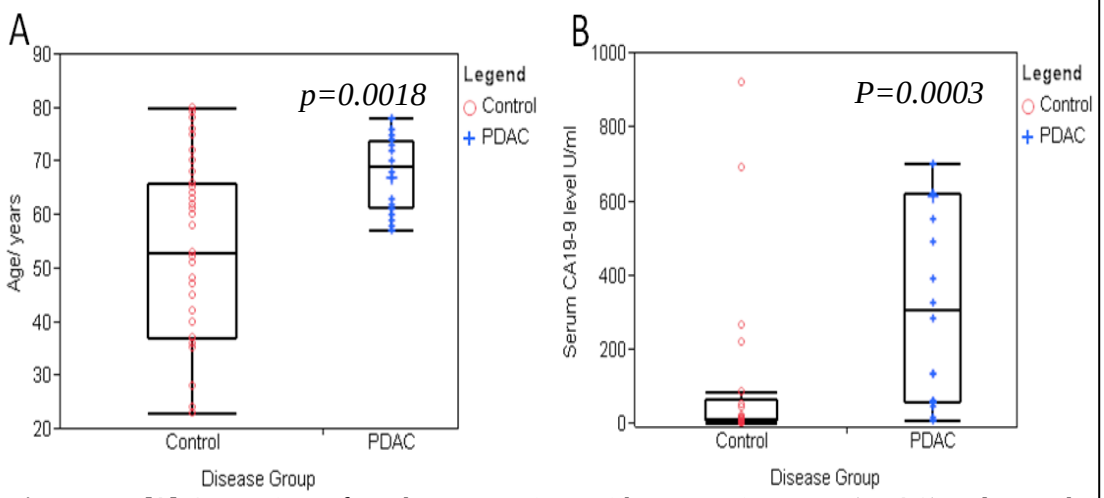


***Figure 2.7- [A]** Comparison of age between patients with pancreatic cancer (**PDAC**) and control patients (inc. healthy volunteers and chronic pancreatitis; Wilcoxon's Test, p=0.0018)); **[B]** Comparison of serum CA19-9 levels between the PDAC and control groups (Wilcoxon's Test, p=0.0003)*

2.4.2.2 Western Blot images for VDBP, RBP-4, and FINC

Examples of Phase II Western Blot images are shown in ***Figure 2.8. All gels were analysed by Western Blot in triplicates.***

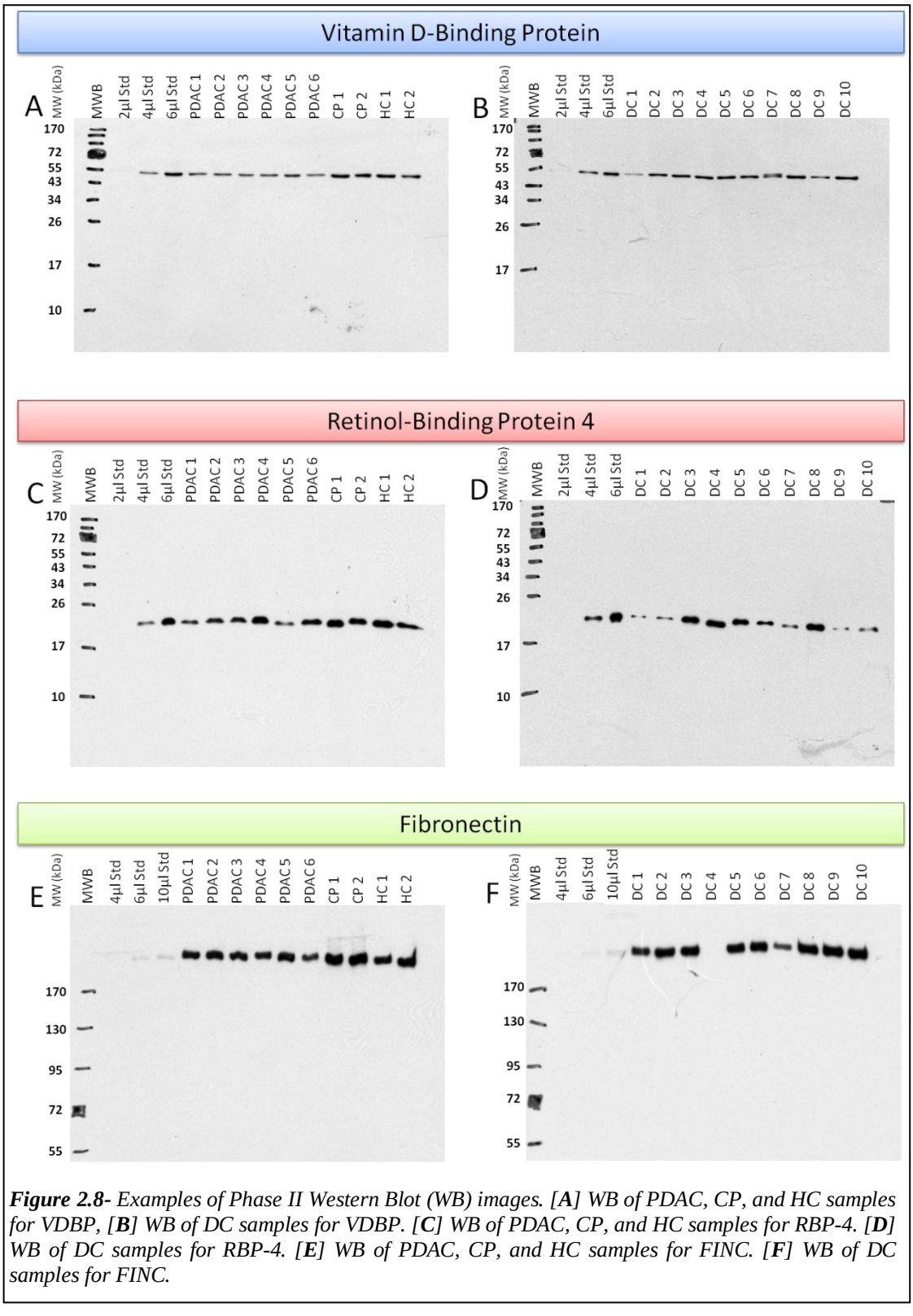


Figure 2.8- *Examples of Phase II Western Blot (WB) images. [**A**] WB of PDAC, CP, and HC samples for VDBP, [**B**] WB of DC samples for VDBP. [**C**] WB of PDAC, CP, and HC samples for RBP-4. [**D**] WB of DC samples for RBP-4. [**E**] WB of PDAC, CP, and HC samples for FINC. [**F**] WB of DC samples for FINC.*

2.4.2.3 The serum level of VDBP in PDAC and Controls

The serum level of VDBP was significantly reduced in patients with pancreatic cancer compared to individuals with chronic pancreatitis and healthy volunteers (Wilcoxon's Test, p=0.004 and p=0.006, respectively; ***Figure 2.9A-B; Table 2.8***). However, the expression of VDBP was similar between individuals with PDAC and biliary obstruction (DC) (Wilcoxon's Test, p=0.465; ***Figure 2.9B***). Further analyses stratified patients into PDAC and non-cancer groups (***Figure 2.9C***). Statistical analysis by Wilcoxon's Test showed that there was significant difference in the serum level of VDBP between the PDAC group and the combined HC and CP group (p<0.001). However, when the DC group was added to the combined control group (i.e. PDAC against HC, CP, and DC) this significance is lost (Wilcoxon's, p=0.113). This may be explained by the apparently low relative intensity of VDBP in both the PDAC and DC groups ***Figure 2.9A.***

Table 2.8*- Relative quantification of serum VDBP by densitometry analysis*

VDBP	**Chronic Pancreatitis** Median (Range)	**Disease control** Median (Range)	**Healthy Control** Median (Range)	**Pancreatic Cancer** Median (Range)
Relative intensity*	1.350 (1.128-1.922)	1.066 (0.894-1.274)	1.338 (1.156-1.520)	1.066 (0.892-1.562)

*Note: *relative intensity calculated from densitometry results: sample mean region of interest (ROI) intensity divided by the standard reference mean ROI Intensity*

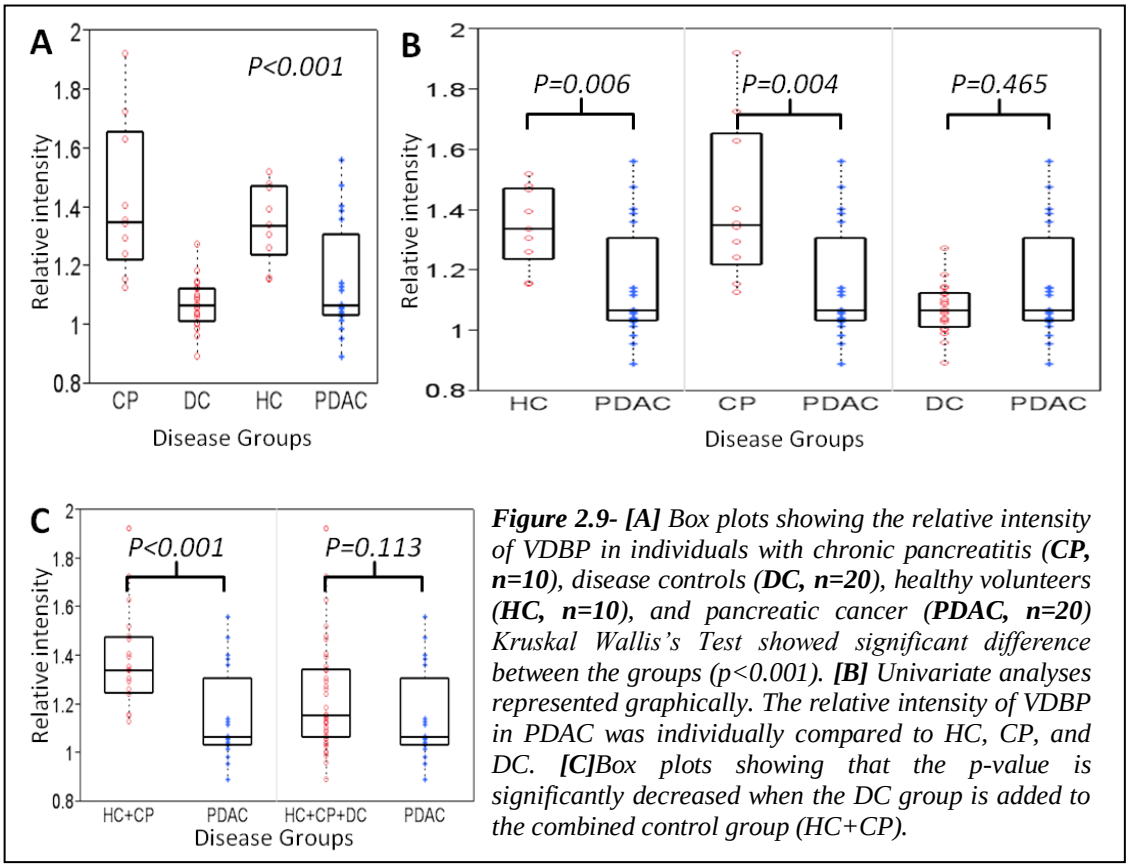


Figure 2.9- [A] *Box plots showing the relative intensity of VDBP in individuals with chronic pancreatitis (**CP, n=10**), disease controls (**DC, n=20**), healthy volunteers (**HC, n=10**), and pancreatic cancer (**PDAC, n=20**) Kruskal Wallis's Test showed significant difference between the groups (p<0.001).* ***[B]*** *Univariate analyses represented graphically. The relative intensity of VDBP in PDAC was individually compared to HC, CP, and DC.* ***[C]****Box plots showing that the p-value is significantly decreased when the DC group is added to the combined control group (HC+CP).*

2.4.2.4 The serum levels of RBP-4 in PDAC and Controls

Results from the Kruskal Wallis Test showed that there was a statistically significant decrease in the serum level of RBP-4 between the disease groups (p=0.043; ***Table 2.9; Figure 2.10A***). On univariate analysis, the expression of RBP-4 was significantly decreased in the PDAC group compared to the HC and the CP groups but not the DC group (Wilcoxon's Test, p=0.011, p=0.031, p=0.176; ***Figure 2.10B)***. Further analysis stratified the groups into PDAC and control groups (***Figure 2.10C***). Results suggest that there was a significant difference in serum level of RBP-4 between PDAC and the combined HC/CP control group (Wilcoxon's Test, p=0.004)). This significance was maintained even when DC was added to the combined control group (Wilcoxon's Test, p=0.014).

Table 2.9*- Relative quantification of serum RBP-4 by densitometry analysis*

RBP-4	**Chronic Pancreatitis** Median (Range)	**Disease control** Median (Range)	**Healthy Control** Median (Range)	**Pancreatic Cancer** Median (Range)
Relative intensity*	1.260 (0.996-1.387)	1.177 (0.857-1.446)	1.240 (1.076-1.646)	1.066 (0.666-1.410)

*Note: *relative intensity calculated from densitometry results: sample mean region of interest (ROI) intensity divided by the standard reference mean ROI Intensity*

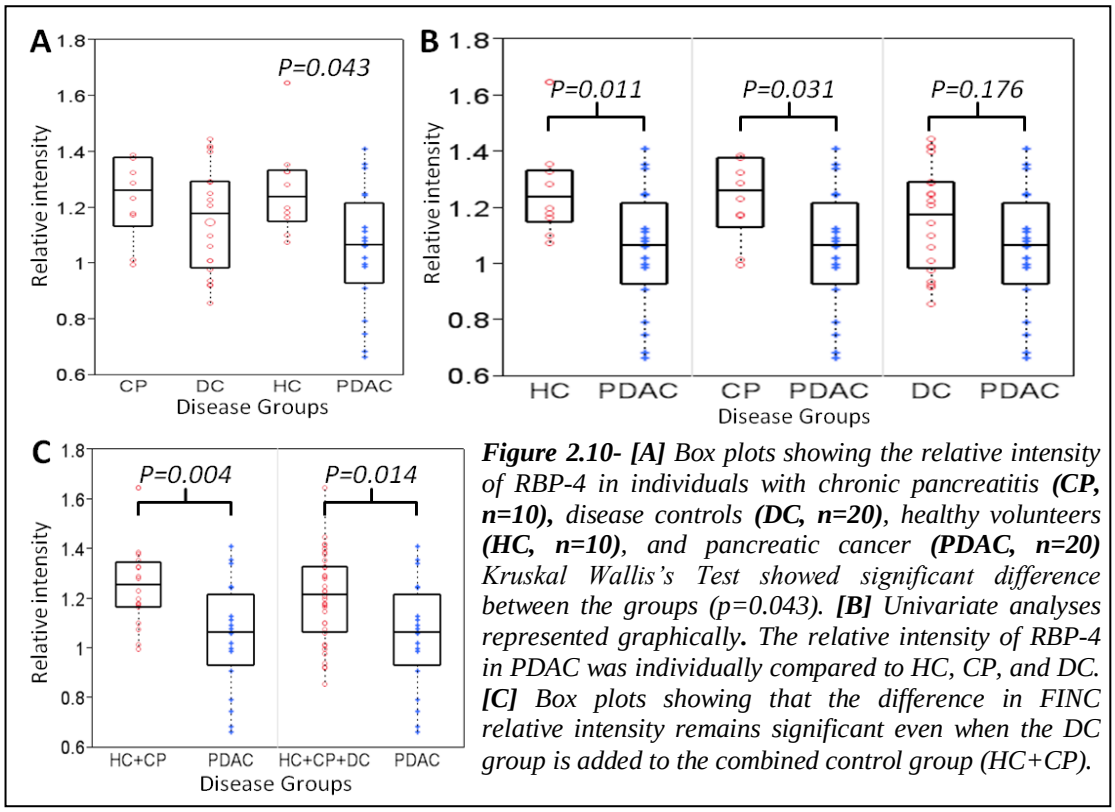


Figure 2.10- [A] *Box plots showing the relative intensity of RBP-4 in individuals with chronic pancreatitis* ***(CP, n=10),*** *disease controls* ***(DC, n=20)****, healthy volunteers* ***(HC, n=10)****, and pancreatic cancer* ***(PDAC, n=20)*** *Kruskal Wallis's Test showed significant difference between the groups (p=0.043).* ***[B]*** *Univariate analyses represented graphically****.*** *The relative intensity of RBP-4 in PDAC was individually compared to HC, CP, and DC.* ***[C]*** *Box plots showing that the difference in FINC relative intensity remains significant even when the DC group is added to the combined control group (HC+CP).*

2.4.2.5 The serum level of FINC in PDAC and Controls

General analysis by Kruskal Wallis Test showed that the there is insignificant difference in serum FINC level between the disease groups (p=0.146; ***Figure 2.11A; Table 2.10***). However, on univariate analysis, the expression of FINC was significantly decreased in the PDAC group compared to the HC and the CP groups but not the DC group (Wilcoxon's Test, p=0.031, p=0.035, p=0.176; ***Figure 2.11B***). Further analysis stratified the groups into PDAC and control groups (***Figure 2.11C***). Again, there seemed to be a significant difference in serum level of FINC between PDAC and the combined control group (HC+CP; Wilcoxon's Test, p=0.009). This significance was maintained when DC was added to the combined control group (Wilcoxon's Test, p=0.012).

Table 2.10- *Relative quantification of serum FINC by densitometry analysis*

FINC	**Chronic Pancreatitis** Median (Range)	**Disease control** Median (Range)	**Healthy Control** Median (Range)	**Pancreatic Cancer** Median (Range)
Relative intensity*	1.989 (1.224-2.428)	2.018 (0.918-2.999)	1.941 (1.707-2.170)	1.754 (0.919-2.125)

*Note: *relative intensity calculated from densitometry results: sample mean region of interest (ROI) intensity divided by the standard reference mean ROI Intensity*

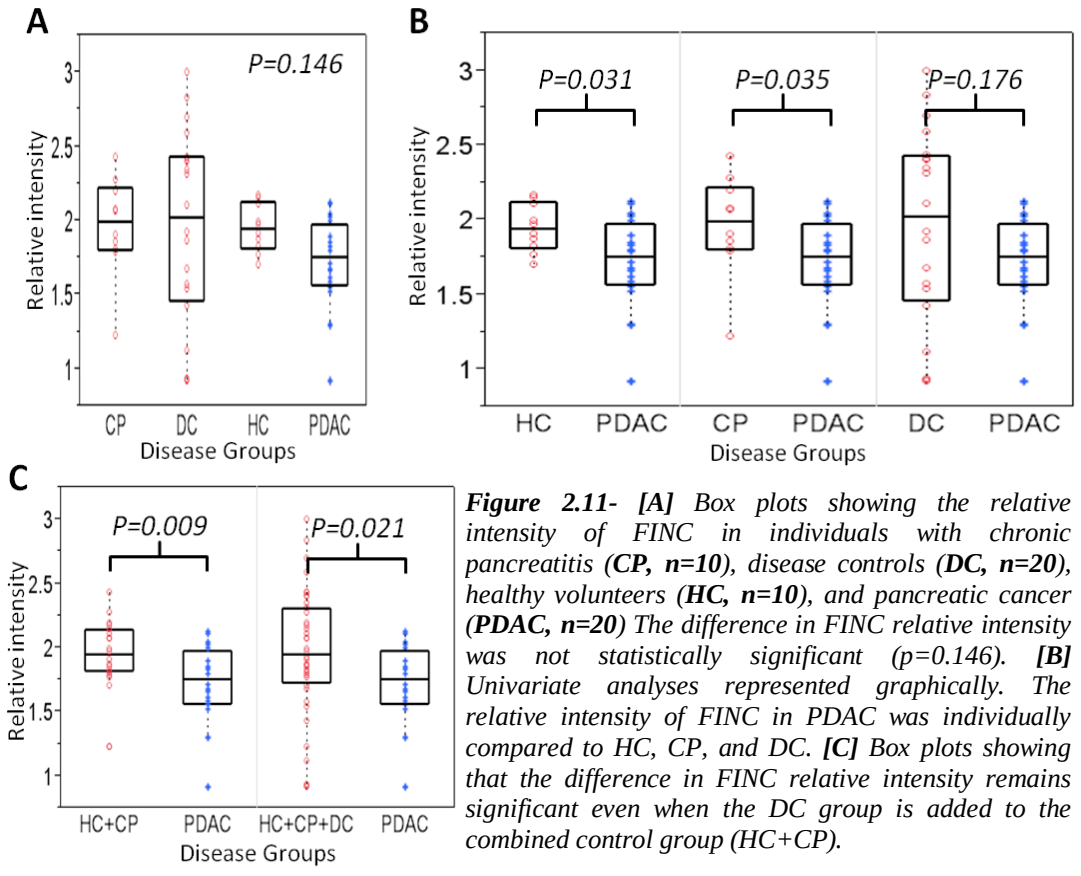


Figure 2.11- [A] *Box plots showing the relative intensity of FINC in individuals with chronic pancreatitis (**CP, n=10**), disease controls (**DC, n=20**), healthy volunteers (**HC, n=10**), and pancreatic cancer (**PDAC, n=20**) The difference in FINC relative intensity was not statistically significant (p=0.146).* ***[B]*** *Univariate analyses represented graphically. The relative intensity of FINC in PDAC was individually compared to HC, CP, and DC.* ***[C]*** *Box plots showing that the difference in FINC relative intensity remains significant even when the DC group is added to the combined control group (HC+CP).*

2.4.2.6 Accuracy of candidate markers for the diagnosis of PDAC versus HC

Results from the diagnostic accuracy analyses for PDAC against healthy controls are summarised in ***Table 2.11***. On univariate analysis, VDBP, RBP-4, and FINC were significant discriminators of PDAC against HC (Wilcoxon's test, all $p<0.05$). However, this significance was not maintained on multivariate analysis i.e. the three markers were not independent to each other. Nevertheless, results from ROC analysis showed that VDBP, RBP-4, and FINC were all relatively accurate at discriminating between PDAC against HC with ROC-AUCs of 0.82, 0.80, and 0.75 respectively. In particular, VDBP achieved a relatively high sensitivity (0.75) and specificity (1.00) at the optimal cut-off of 1.15 (***Figure 2.12***). A preliminary analysis was performed to assess the impact of combining the three markers on the accuracy for the diagnosis of PDAC against HC. Interestingly, the diagnostic accuracy was increased to 0.89 when the three markers were combined by the M-LR model with relative sensitivity and specificity of 0.85 and 0.90 at a cut-off of 0.70 (***Figure 2.12***). Although results from this analysis would require verification in an independent sample set.

Table 2.11- *The diagnostic accuracy of VDBP, RBP-4, and FINC for PDAC against HC*

Biomarker	**Pancreatic Cancer versus Healthy Control**					
	Univariate	Multivariate	ROC-AUC	C/O	Sensitivity	Specificity
VDBP	**0.006**	0.05	0.815	1.15	0.75	1.00
RBP-4	**0.011**	0.09	0.790	1.13	0.75	0.80
FINC	**0.031**	0.16	0.745	1.68	0.45	1.00
Combined	-	-	0.890	0.70	0.85	0.90

Abbreviations: *Vitamin D-Binding Protein (**VDBP**); Retinol-Binding Protein 4 (**RBP-4**); Fibronectin (**FINC**); M-LR combined marker consisting of VDBP, RBP-4, and FINC (**Combined**); Receiver Operator Characteristics Area Under the curve (**ROC-AUC**); Cut-off for the reported sensitivity and specificity (**C/O**)*

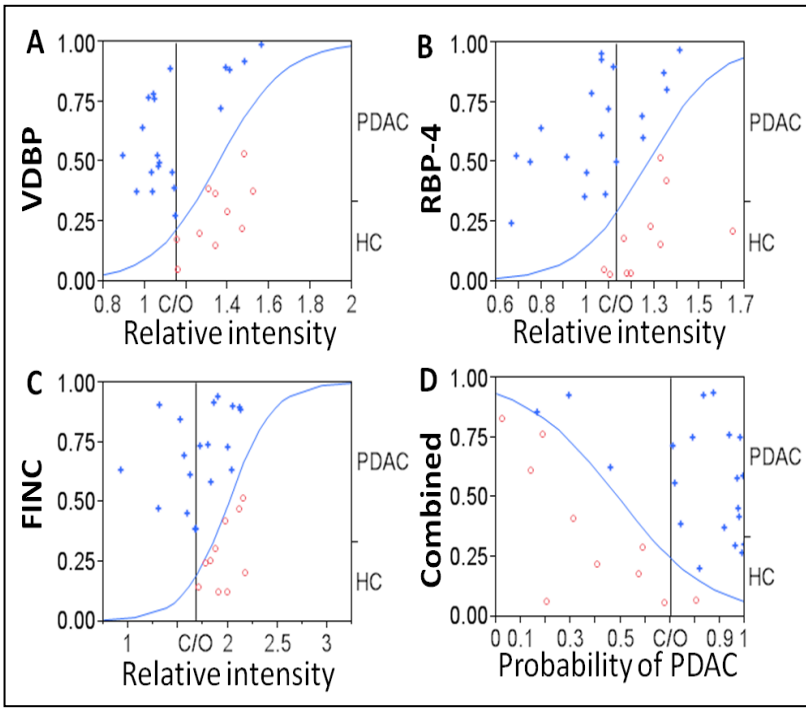


Figure 2.12- *Logistic Plots showing the diagnostic accuracy of **[A]** VDBP ,**[B]** RBP-4, and **[C]** FINC as individual markers of PDAC against HC. **[D]** Logistic plot showing the diagnostic accuracy when all three proteins were combined. Individuals with PDAC are coloured **Blue** HC in **red**. The optimal cut-off value (**C/O**) was determined by ROC and is shown on the logistic plots as a solid black line. Although results from this preliminary analysis was encouraging, further analysis would be required to verify the accuracy of the M-LR model*

2.4.2.7 Accuracy of candidate markers for the diagnosis of PDAC against CP

Results from the diagnostic accuracy analyses for PDAC against chronic pancreatitis are summarised in ***Table 2.12***. On univariate analysis, VDBP, RBP-4, and FINC were significant discriminators of pancreatic cancer against chronic pancreatitis (Wilcoxon's test, all $p<0.05$). Furthermore, this significance was maintained on multivariate analysis i.e. the three markers were independent markers of pancreatic cancer against CP. In addition, results from ROC analysis showed that VDBP, RBP-4, and FINC were all relatively accurate at discriminating between PDAC against CP with ROC-AUCs of 0.83, 0.75, and 0.74 respectively. Again, VDBP appeared to be the best individual maker with a relatively high sensitivity (0.75) and specificity (0.90) at the optimal cut-off of 1.15 (***Figure 2.13***). Similar to the HC results, the accuracy of combining all three markers for the diagnosis of PDAC against CP was assessed (***Figure 2.13***). Results showed that there is a dramatic increase in the diagnostic accuracy with relative sensitivity and specificity of 0.85 and 1.00 at a cut-off of 0.70. Again, despite these encouraging results, further validation on an independent samples set would be necessary.

Table 2.12- *The diagnostic accuracy of VDBP, RBP-4, and FINC for PDAC against CP*

Biomarker	Pancreatic Cancer versus Chronic Pancreatitis					
	Univariate	Multivariate	ROC-AUC	C/O	Sensitivity	Specificity
VDBP	**0.004**	**0.002**	0.825	1.15	0.75	0.90
RBP-4	**0.031**	**0.03**	0.745	1.13	0.75	0.80
FINC	**0.035**	**0.029**	0.740	1.72	0.50	0.90
Combined	-	-	0.920	0.70	0.85	1.00

Abbreviations: *Vitamin D-Binding Protein (**VDBP**); Retinol-Binding Protein 4 (**RBP-4**); Fibronectin (**FINC**); M-LR combined marker consisting of VDBP, RBP-4, and FINC (**Combined**); Receiver Operator Characteristics Area Under the curve (**ROC-AUC**); Cut-off for the reported sensitivity and specificity (**C/O**)*

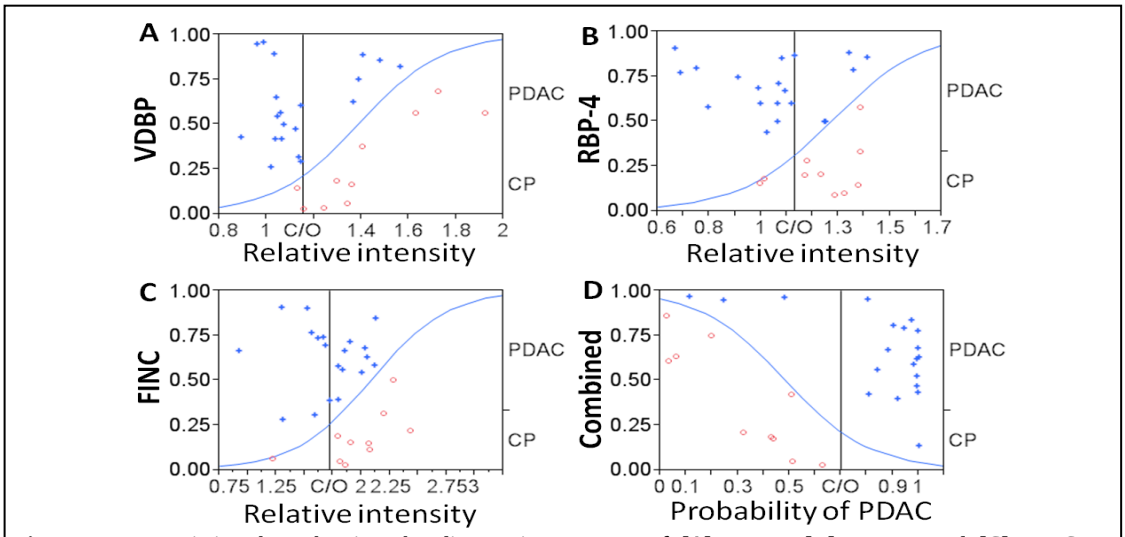


Figure 2.13- *Logistic Plots showing the diagnostic accuracy of **[A]** VDBP ,**[B]** RBP-4, and **[C]** FINC as individual markers of PDAC against chronic pancreatitis. **[D]** Logistic plot showing the diagnostic accuracy when all three proteins were combined. Individuals with PDAC are coloured **Blue** and chronic pancreatitis are in **red**. The optimal cut-off value (**C/O**) was determined by ROC and is shown on the logistic plots as a solid black*

2.4.2.8 Accuracy of candidate markers for the diagnosis of PDAC against DC

Results from the diagnostic accuracy analyses for PDAC against disease controls are summarised in ***Table 2.12***. On univariate analysis, VDBP, RBP-4, and FINC were not significant discriminators of pancreatic cancer against individuals with biliary obstruction (DC) (Wilcoxon's test, all p>0.05). Subsequently, no multivariate analysis was performed. Furthermore, results from ROC analysis showed that VDBP, RBP-4, and FINC were not sufficiently accurate at discriminating between PDAC against DC with ROC-AUCs of 0.57, 0.63, and 0.63 respectively. Moreover, the diagnostic accuracy was not significant (AUC 0.74) when the three markers were combined by multinomial logistic regression (M-LR) with relative sensitivity and specificity of 0.55 and 0.90 at a cut-off of 0.62 (***Figure 2.14)***.

Table 2.13- *The diagnostic accuracy of VDBP, RBP-4, and FINC for PDAC against DC*

Biomarker	**Pancreatic Cancer versus Disease Control**					
	Univariate	Multivariate	ROC-AUC	C/O	Sensitivity	Specificity
VDBP	0.465	-	0.568	1.36	0.25	1.00
RBP-4	0.176	-	0.625	1.13	0.75	0.55
FINC	0.176	-	0.625	2.20	1.00	0.45
Combined	-	-	0.743	0.62	0.55	0.90

Abbreviations: *Vitamin D-Binding Protein (**VDBP**); Retinol-Binding Protein 4 (**RBP-4**); Fibronectin (**FINC**); M-LR combined marker consisting of VDBP, RBP-4, and FINC (**Combined**); Receiver Operator Characteristics Area Under the curve (**ROC-AUC**); Cut-off for the reported sensitivity and specificity (**C/O**)*

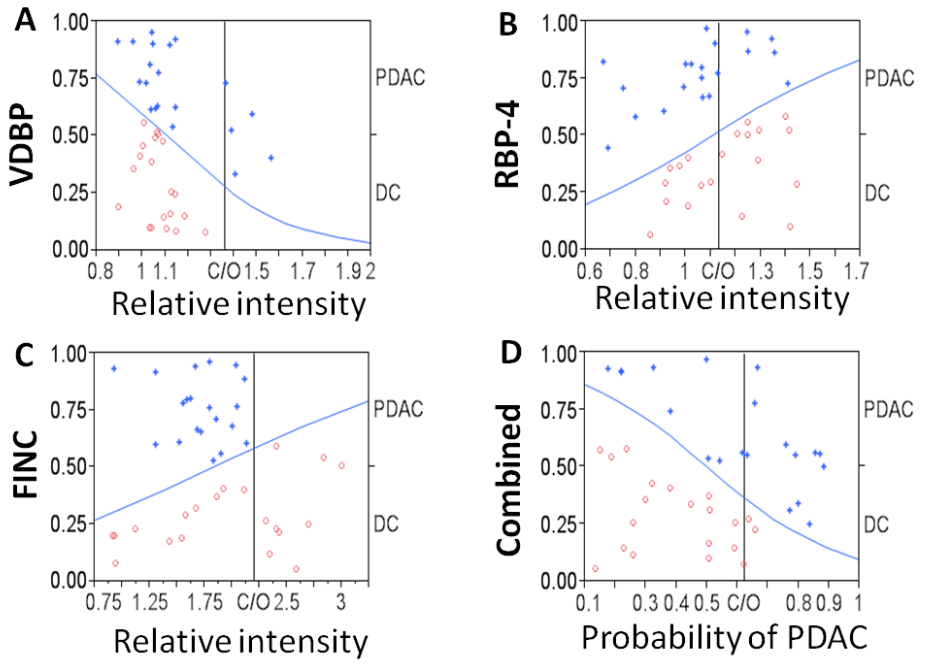


Figure 2.14 *Logistic Plots showing the diagnostic accuracy of **[A]** VDBP ,**[B]** RBP-4, and **[C]** FINC as individual markers of PDAC against disease control. **[D]** Logistic plot showing the diagnostic accuracy when all three proteins were combined. Individuals with PDAC are coloured **Blue** and disease controls are in **red**. The optimal cut-off value (**C/O**) was determined by ROC and is shown on the logistic plots as a solid black line.*

2.4.2.9 Diagnostic accuracy of PDAC against all controls

Results from the diagnostic accuracy analyses for PDAC against all controls are summarised in ***Table 2.14***. On univariate analysis, only RBP-4, and FINC were significant discriminators of pancreatic cancer against chronic pancreatitis (Wilcoxon's test, $p<0.05$). Furthermore, results from multivariate analysis indicate that only RBP-4 was an independent discriminator of pancreatic cancer against control subjects. Results from ROC analysis showed that VDBP, RBP-4, and FINC were insufficiently accurate at discriminating between PDAC against controls subjects with ROC-AUCs of 0.626, 0.696, and 0683, respectively. Similar to previous results, the diagnostic accuracy was increased when all three markers were combined by M-LR modelling with ROC-AUC of 0.756 and relative sensitivity and specificity of 0.70 and 0.80 at a cut-off of 0.36 (***Figure 2.15***), although this result would required verification with an independent sample set.

Table 2.14- The diagnostic accuracy of VDBP, RBP-4, and FINC for PDAC against all controls

Biomarker	**Pancreatic Cancer versus all controls**					
	Univariate	Multivariate	ROC-AUC	C/O	Sensitivity	Specificity
VDBP	0.113	-	0.626	1.069	0.55	0.75
RBP-4	**0.014**	**0.006**	0.696	1.130	0.75	0.67
FINC	**0.021**	0.067	0.683	1.853	0.70	0.65
Combined	-	-	0.756	0.36	0.70	0.80

Abbreviations: *Vitamin D-Binding Protein (**VDBP**); Retinol-Binding Protein 4 (**RBP-4**); Fibronectin (**FINC**); M-LR combined marker consisting of VDBP, RBP-4, and FINC (**Combined**); Receiver Operator Characteristics Area Under the curve (**ROC-AUC**); Cut-off for the reported sensitivity and specificity (**C/O**)*

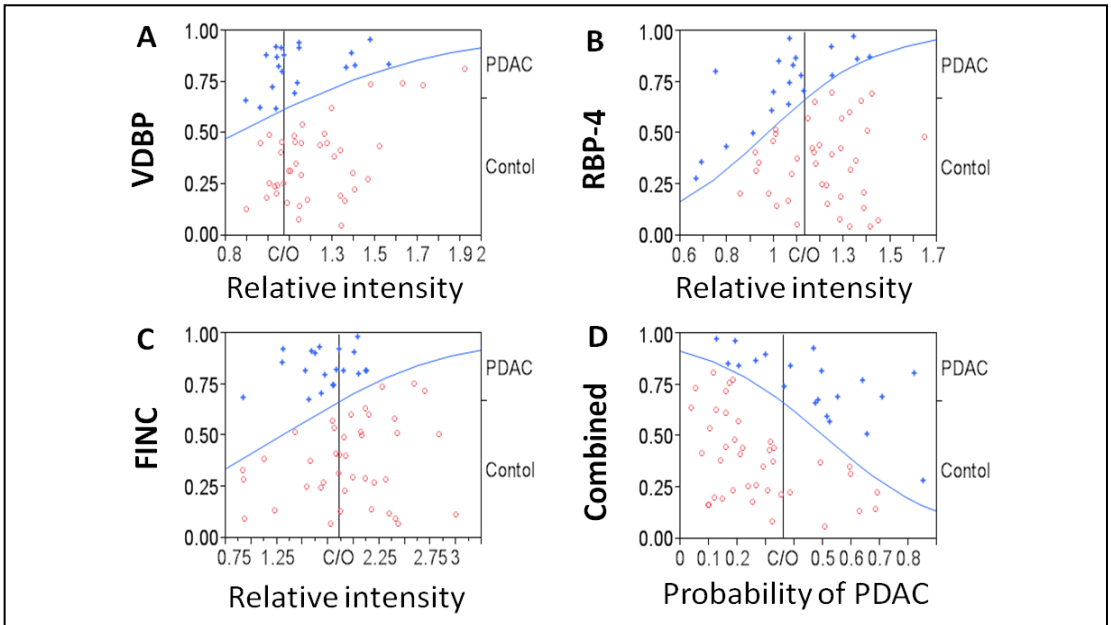


Figure 2.15 *Logistic Plots showing the diagnostic accuracy of **[A]** VDBP ,**[B]** RBP-4, and **[C]** FINC as individual markers of PDAC against all control subjects. **[D]** Logistic plot showing the diagnostic accuracy when all three proteins were combined. Individuals with PDAC are coloured **Blue** and controls are in **red**. The optimal cut-off value (**C/O**) was determined by ROC and is shown on the logistic plots as a solid black line.*

2.4.2.10 Accuracy of candidate markers for the diagnosis of PDAC against HC and CP combined

In view of the fact that existing literature on novel biomarkers invariably compared PDAC against HC and CP control subgroups only and to facilitate the comparison between the current and previous biomarker studies, a separate analysis was performed to determine the overall performance of the three markers in absence of the DC group (see ***Table 2.15***). Results from this analysis indicated that VDBP, RBP-4, and FINC were all independent markers of PDAC against the combination of HC and CP (***Figure 2.17*** and ***Figure*** 2.16). Furthermore, ROC analysis demonstrated that the markers have a relatively high accuracy for discriminating between PDAC against HC and CP with ROC-AUCs of 0.82, 0.77, and 0.74. Finally, results from the combination of the markers by M-LR were encouraging with ROC-AUC of 0.91 and relative sensitivity and specificity of 0.85 and 0.95, respectively, at the optimal cut-off of 0.59 (***Figure 2.17***). However, similar to previous analyses, results from this analysis would require verification with an independent sample set

Table 2.15- The diagnostic accuracy of VDBP, RBP-4, and FINC for PDAC against HC and CP

Biomarker	**Pancreatic Cancer versus healthy control and chronic pancreatitis**					
	Univariate	Multivariate	ROC-AUC	C/O	Sensitivity	Specificity
VDBP	**<0.001**	**0.002**	0.820	1.15	0.75	0.95
RBP-4	**0.004**	**0.015**	0.768	1.13	0.75	0.80
FINC	**0.009**	**0.028**	0.741	1.70	0.45	0.95
Combined	**-**	**-**	0.908	0.59	0.85	0.95

Abbreviations: *Vitamin D-Binding Protein (**VDBP**); Retinol-Binding Protein 4 (**RBP-4**); Fibronectin (**FINC**); M-LR combined marker consisting of VDBP, RBP-4, and FINC (**Combined**); Receiver Operator Characteristics Area Under the curve (**ROC-AUC**); Cut-off for the reported sensitivity and specificity (**C/O**)*

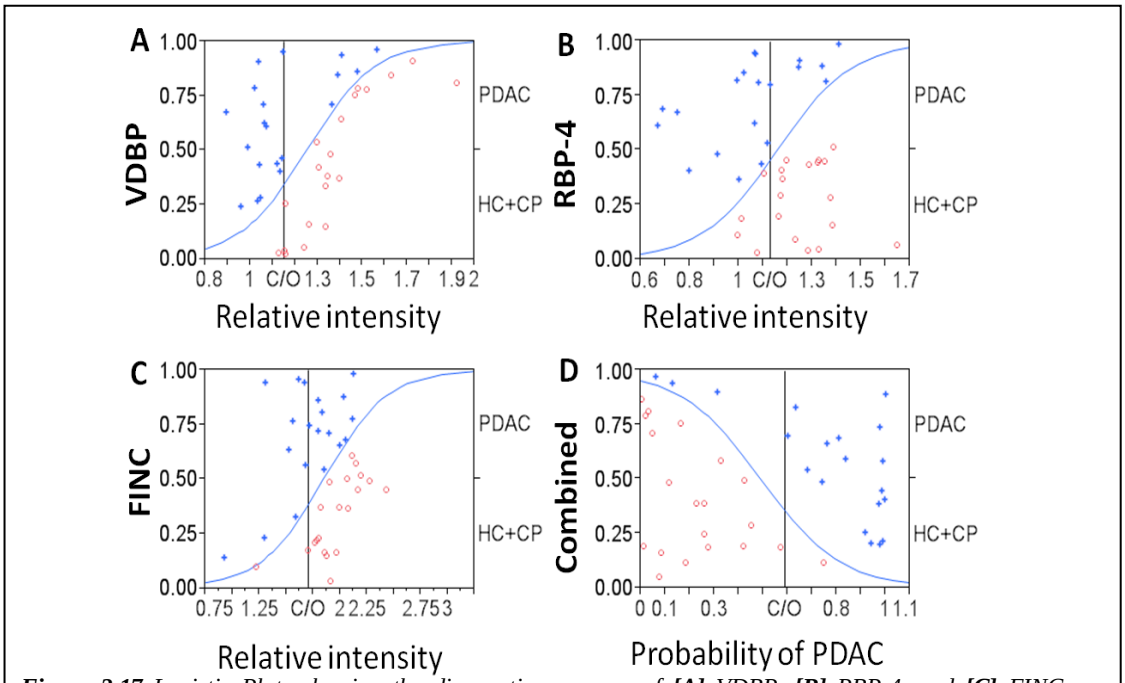


Figure 2.17 *Logistic Plots showing the diagnostic accuracy of* ***[A]*** *VDBP ,****[B]*** *RBP-4, and* ***[C]*** *FINC as individual markers of PDAC against the combined HC and CP group.* ***[D]*** *Logistic plot showing the diagnostic accuracy when all three proteins were combined. Individuals with PDAC are coloured* ***Blue*** *and HC/CP are in* ***red****. The optimal cut-off value (****C/O****) was determined by ROC and is shown on the logistic plots as a solid black line.*

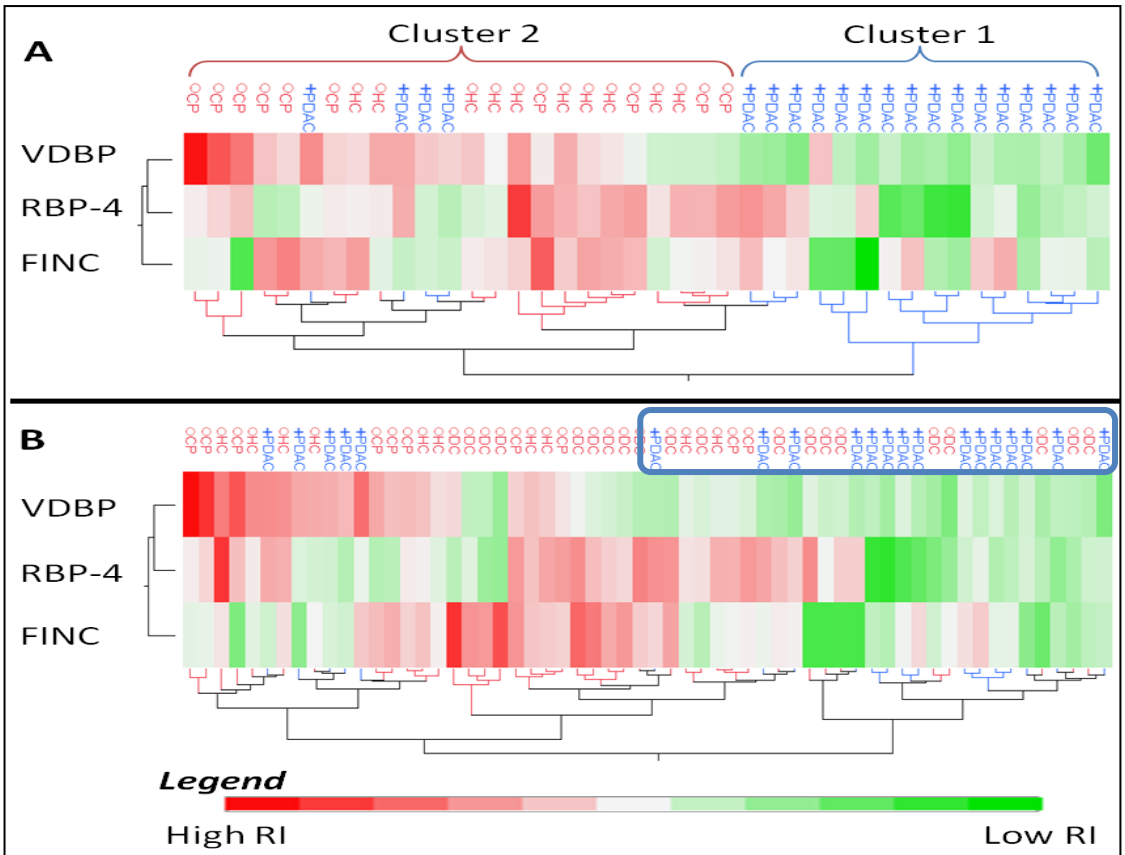


Figure 2.16- *Heat maps showing the clustering of patients according to the cancer status.* ***[A]*** *Clustering in absence of the DC subgroup resulted in two defined clusters. Cluster 1* ***(Blue)*** *consists of PDAC samples only whereas cluster 2 (Red) consists of mainly HC and CP samples.*
[B] *Clustering of all PDAC and control samples. Although the heat map pattern is similar to [A], Clustering of patients became much less define when the disease control group is added to the analysis (****marked****).*

2.4.2.11 Correlation between age and candidate protein markers

Univariate analysis of demographical data in ***section 2.4.2.1*** showed that individuals in the PDAC group were older compared to individuals in the control subgroups. This is supported by the pattern shown in ***Figure 2.18***. Therefore, the correlations between age and the relative serum concentrations of the VDBP, RBP-4, and FINC in PDAC patients and controls were independently assessed using the Kendall Tau non-parametric correlation analysis. Results indicated that there was minimal correlation (Kendall τ<0.5) between age, individual protein markers, and the M-LR disease-predicting model (***Table 2.16; see section 2.3.6.2 for a description of the Kendall Tau test and the interpretation of its results)***.

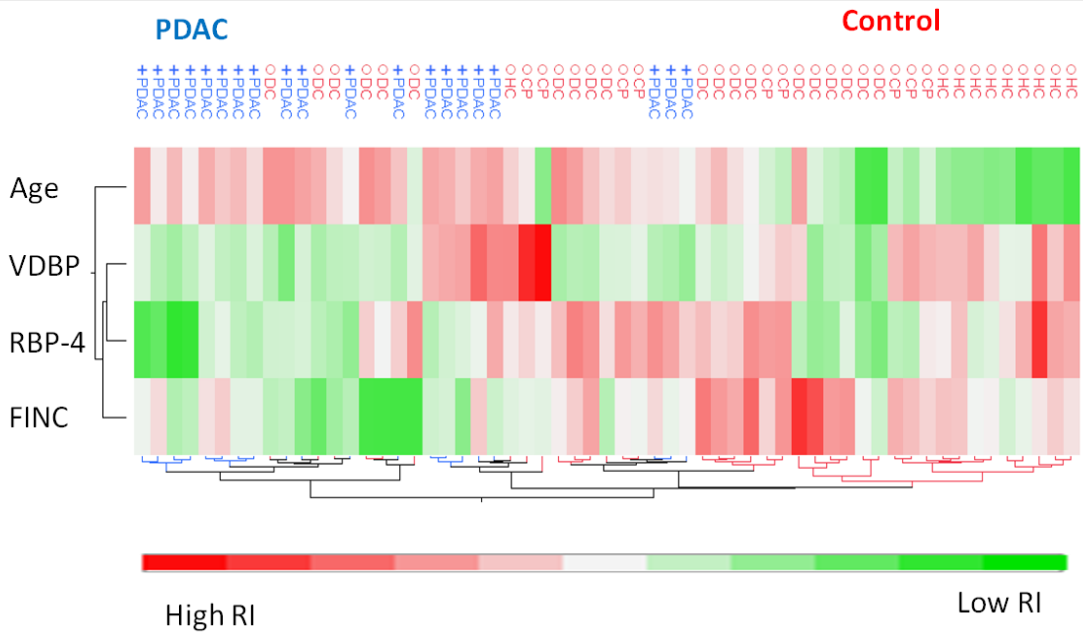


Figure 2.18- *Two-way clustering of Age, vitamin D-binding protein (**VDBP**), Retinol-Binding Protein 4 (**RBP-4**), and Fibronectin (**FINC**). Whilst the heat map pattern suggests that VDBP, RBP-4, and FINC may be decreased in older individuals, this is likely to be attributable to the fact that individuals in the PDAC groups are older than individuals in the control group. Standardised relative intensity on western blot densitometry analysis (**RI**)*

Table 2.16- *Correlation between Age, individual protein markers, and disease-predicting formulae*

	Controls		PDAC	
Parameters	Kendall τ	Prob>|τ|	Kendall τ	Prob>|τ|
Vitamin D-Binding Protein	-0.32	<0.01	0.27	0.10
Retinol-Binding Protein	0.06	0.58	0.00	1.00
Fibronectin	<-0.01	0.94	-0.11	0.52
M-LR algorithm	0.03	0.77	0.00	1.00

Abbreviations*: Multinomial Logistic Regression (**M-LR**); Artificial neural network (**NN**)*

Note: *Kendall Tau (τ) coefficients ranges from -1 (perfect negative correlation) to 0 (no correlation) to +1 (perfect positive correlation); Prob>$|\tau|$ represents the probability that the actual correlation coefficient (τ) would be greater than the calculated coefficient.*

2.4.3Phase III- Validation with pre-diagnostic samples (UKCTOCS)

2.4.3.1Characteristics of the UKCTOCS samples

Pre-Pancreatic cancer (PPC) samples were stratified into six time-categories according to the duration between sample collection and diagnosis: 0-0.5, 0.5-1, 1-2, 2-3, 3-4, and over 4 years). Each time group therefore, consists of 10 PPC samples and 10 matched HC samples. There were no significant difference in age, gender (all females), or delay in sample processing between the PPC and HC group.

Table 2.17*- UCKTOCKS samples characteristics*

Time category (years)	**Age/ years**		**Collection to processing Time (hours)**		**Collection to diagnosis Time (Months)**
	Median (range)	p	Median (range)	p	Median (range)
0 yr - 0.5yr					
HC, n=10	64 (52-75)	N.S.	22 (19-25)	N.S.	-
PPC, n=10	65 (57-78)		22 (20-26)		2.1 (1.3-5.5)
0.5yr – 1yr					
HC, n=10	62 (53-72)	N.S.	20 (19-24)	N.S.	-
PPC, n=10	70 (57-73)		21 (19-24)		9.5 (6.3-11.0)
1yr – 2yr					
HC, n=10	60 (53-72)	N.S.	23 (19-46)	N.S.	-
PPC, n=10	61 (53-72)		23 (20-46)		17.4 (12.7-22.7)
2yr – 3yr					
HC, n=10	61 (51-73)	N.S.	24 (18-25)	N.S.	-
PPC, n=10	61 (52-71)		24 (18-25)		28.2 (24.6-34.3)
3yr – 4yr					
HC, n=10	58 (52-71)	N.S.	23 (21-43)	N.S.	-
PPC, n=10	61 (57-75)		23 (21-44)		39.3 (36.6-47.9)
4+ years					
HC, n=10	61 (51-72)	N.S.	20 (4-26)	N.S.	-
PPC, n=10	65 (52-73)		20 (2-22)		60.6 (52.1-80.6)

*Abbreviations: **N.S.** Not significant (Wilcoxon's Test p>0.05); Healthy Control (**HC**); pre-diagnosis pancreatic cancer samples (**PPC**)*

2.4.3.2 Example of Western Blot images for VDBP and RBP-4

Examples from the Western Blot analyses of UKCTOCS samples for VDBP and RBP-4 are shown in ***Figure 2.19***.

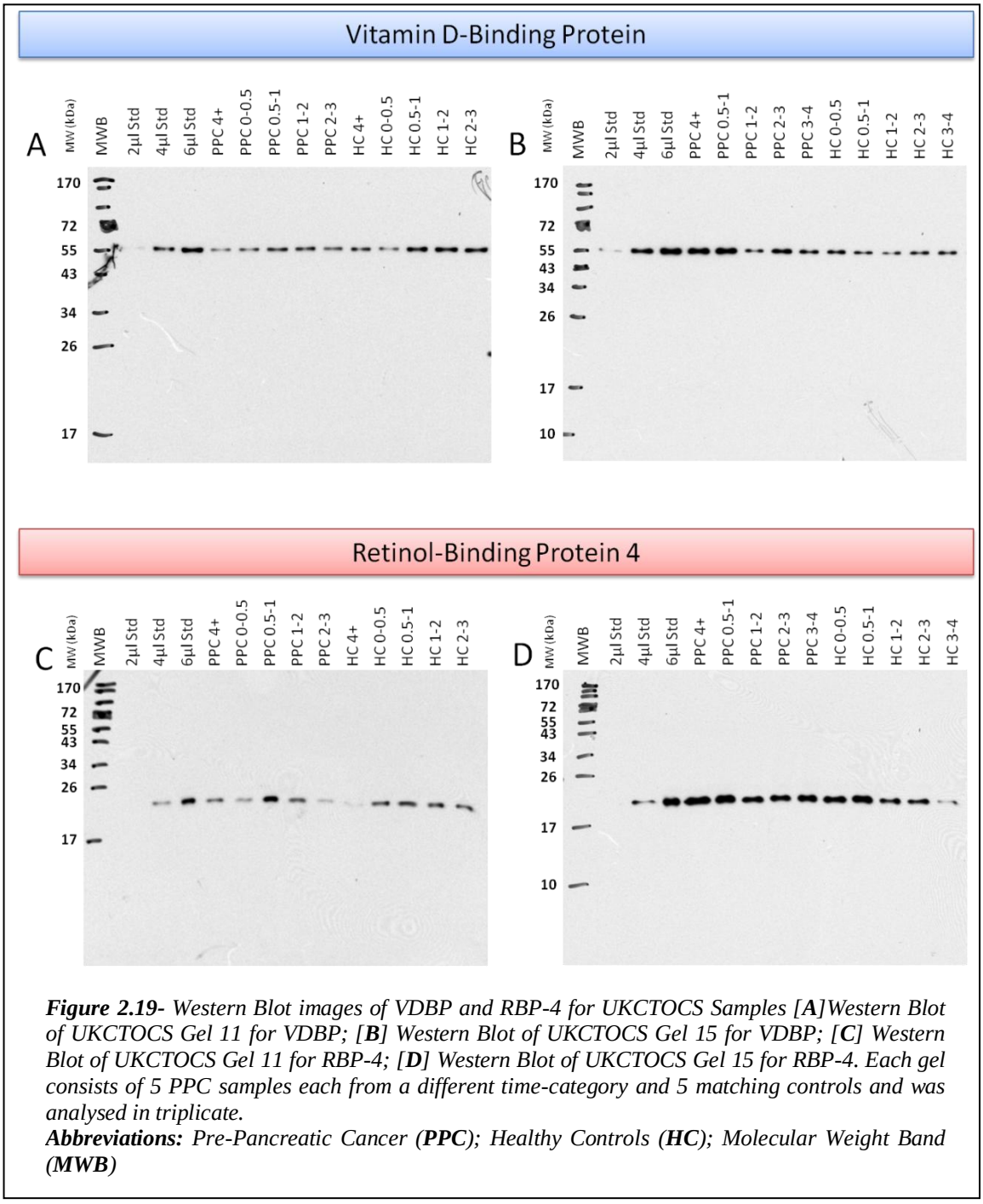


Figure 2.19- *Western Blot images of VDBP and RBP-4 for UKCTOCS Samples [**A**]Western Blot of UKCTOCS Gel 11 for VDBP; [**B**] Western Blot of UKCTOCS Gel 15 for VDBP; [**C**] Western Blot of UKCTOCS Gel 11 for RBP-4; [**D**] Western Blot of UKCTOCS Gel 15 for RBP-4. Each gel consists of 5 PPC samples each from a different time-category and 5 matching controls and was analysed in triplicate.*
Abbreviations: *Pre-Pancreatic Cancer (**PPC**); Healthy Controls (**HC**); Molecular Weight Band (**MWB**)*

2.4.3.3 Selection of markers for validation in Phase III

Due to a limitation in the availability of the UKCTOCS samples, it was decided that only RBP-4 and VDBP would be validated in Phase III. This is because RBP-4 was the only significant biomarker on multivariate analysis when comparing PDAC against control samples (HC, CP, and DC) in Phase-II of the current study. Furthermore, VDBP was also analysed since it was the best discriminator of PDAC against HC and CP. In addition, the protein requirement for the relative quantification of VDBP was a tenth compared to either RBP-4 or FINC.

2.4.3.4 The level of VDBP in pre-diagnosis samples compared to Controls

Univariate analysis by Wilcoxon's Test indicated that there is no significant difference in the serum level of VDBP in pre-diagnostic pancreatic cancer samples compared to the matched controls across most time categories (***Figure 2.20*** and ***Table 2.18***). However, in the 2-3 years category, there is a significant decrease in the serum level of VDBP in control subjects compared to PPC. The reason for this difference is unclear but likely to be co-incidental.

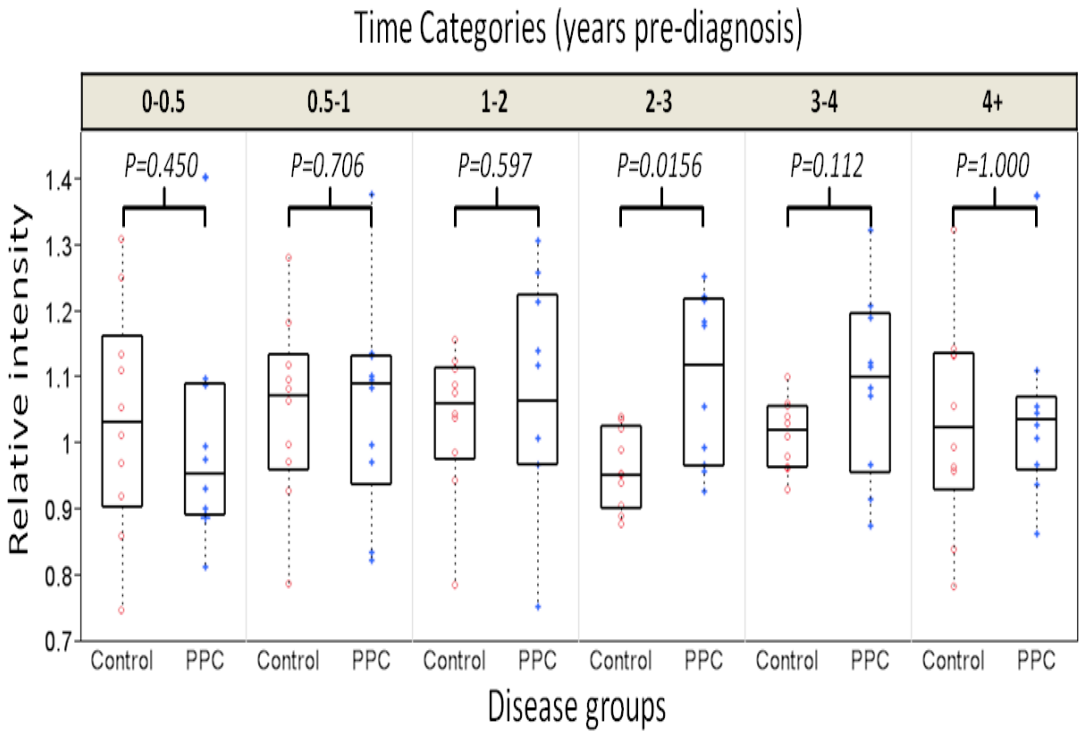


Figure 2.20- *Box Plots showing the level of VDBP between pre-diagnosis pancreatic cancer (**PPC**) and healthy controls (**HC**) across the six time categories. Univariate analysis by Wilcoxon's Test showed no significance expect for the 2-3 yr category where VDBP in HC is lower than PPC.*

Table 2.18- Relative quantification of serum VDBP in UKCTOCS samples by densitometry analysis

VDBP	Time Categories (years before diagnosis)/ median (range)					
	0-0.5	0.5-1	1-2	2-3	3-4	4+
Control	1.03 (0.75-1.31)	1.07 (0.79-1.28)	1.06 (0.79-1.16)	0.95 (0.88-1.04)	1.02 (0.93-1.10)	1.03 (0.78-1.32)
PPC	0.95 (0.81-1.40)	1.09 (0.82-1.38)	1.06 (0.75-1.31)	1.12 (0.93-1.25)	1.10 (0.88-1.32)	1.04 (0.86-1.38)

*Note: *relative intensity calculated from densitometry results: sample mean region of interest (ROI) intensity divided by the standard reference mean ROI Intensity*

2.4.3.5 Change in the relative serum level of VDBP through time

Subsequent analyses aimed to identify any trends in the serum level of VDBP through time. There were three individuals in the PPC group with data from at least four different time categories: PPC patients 10, 12, and 20. Interestingly, PPC patient 10 showed a progressive decrease in relative serum level of VDBP closer to diagnosis; however, this pattern was not maintained in the other two individuals with PPC (***Figure 2.21***). Therefore, whilst the relative serum level of VDBP may not a good predictor of pancreatic cancer development in all patients, a gradual decline in VDBP over time may be observed in certain individuals and may be an indication of pancreatic malignancy.

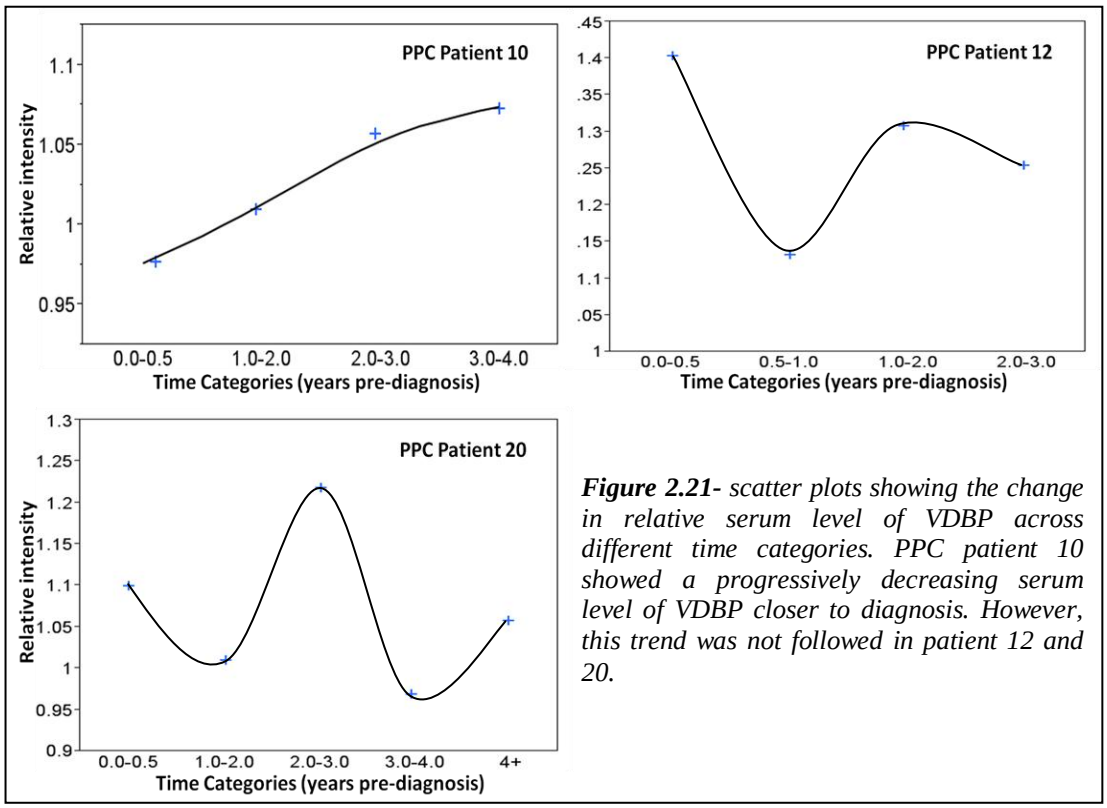


Figure 2.21- *scatter plots showing the change in relative serum level of VDBP across different time categories. PPC patient 10 showed a progressively decreasing serum level of VDBP closer to diagnosis. However, this trend was not followed in patient 12 and 20.*

2.4.3.6 The level of RBP-4 in pre-diagnosis samples compared to Controls

There appears to be a trend towards lower serum expression of RBP-4 in the PPDC group. However, univariate analysis by Wilcoxon's Test yielded insignificant results. Furthermore,the serum level of RBP-4 between the PPC and HC groups across the other time categories also yielded insignificant results (Wilcoxon's Test) p>0.05; ***Figure 2.22*** and ***Table 2.19***).

Table 2.19- Relative quantification of serum RBP-4 in UKCTOCS samples by densitometry analysis

RBP-4	Time Categories (years before diagnosis)/ median (range)					
	0-0.5	0.5-1	1-2	2-3	3-4	4+
Control	1.18 (0.94-2.00)	1.27 (0.99-1.76)	1.21 (1.05-1.47)	1.18 (0.98-1.33)	1.25 (1.11-1.64)	1.17 (0.93-1.91)
PPC	1.07 (0.82-1.84)	1.18 (0.86-1.56)	1.22 (0.93-1.49)	0.86 (1.24-1.43)	1.33 (1.03-1.62)	1.23 (0.93-1.58)

*Note: *relative intensity calculated from densitometry results: sample mean region of interest (ROI) intensity divided by the standard reference mean ROI Intensity*

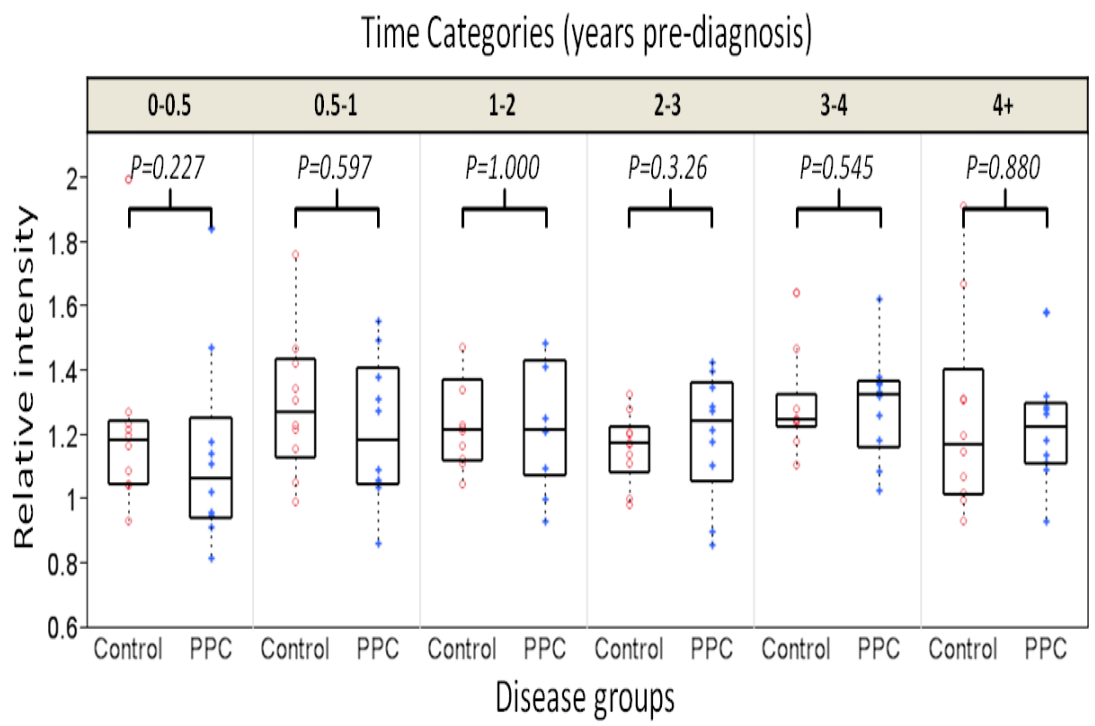


Figure 2.22- *Box Plots comparing the serum level of RBP-4 between healthy controls (**HC**) and pre-diagnosis pancreatic cancer (**PPC**) patients. Univariate analyses indicated that there is no significant difference between the two groups in any of the time groups.*

2.4.3.7 Change in the relative serum level of RBP-4 through time

Subsequent analyses aimed to identify any trends in the serum level of RBP-4 through time. There were three individuals in the PPC group with data from at least four different time categories: PPC patients 10, 12, and 20. Similar to VDBP, the serum level of RBP-4 in PPC patient 10 seemed to be decreasing closer to diagnosis. However, there was a sudden drop in RBP-4 level at 2-3 year pre-diagnosis. Once again, this pattern was not maintained in the other three individuals with PPC. Clearly, the relative serum level of RBP-4 is also not a good predictor of the development of pancreatic cancer (***Figure 2.23***).

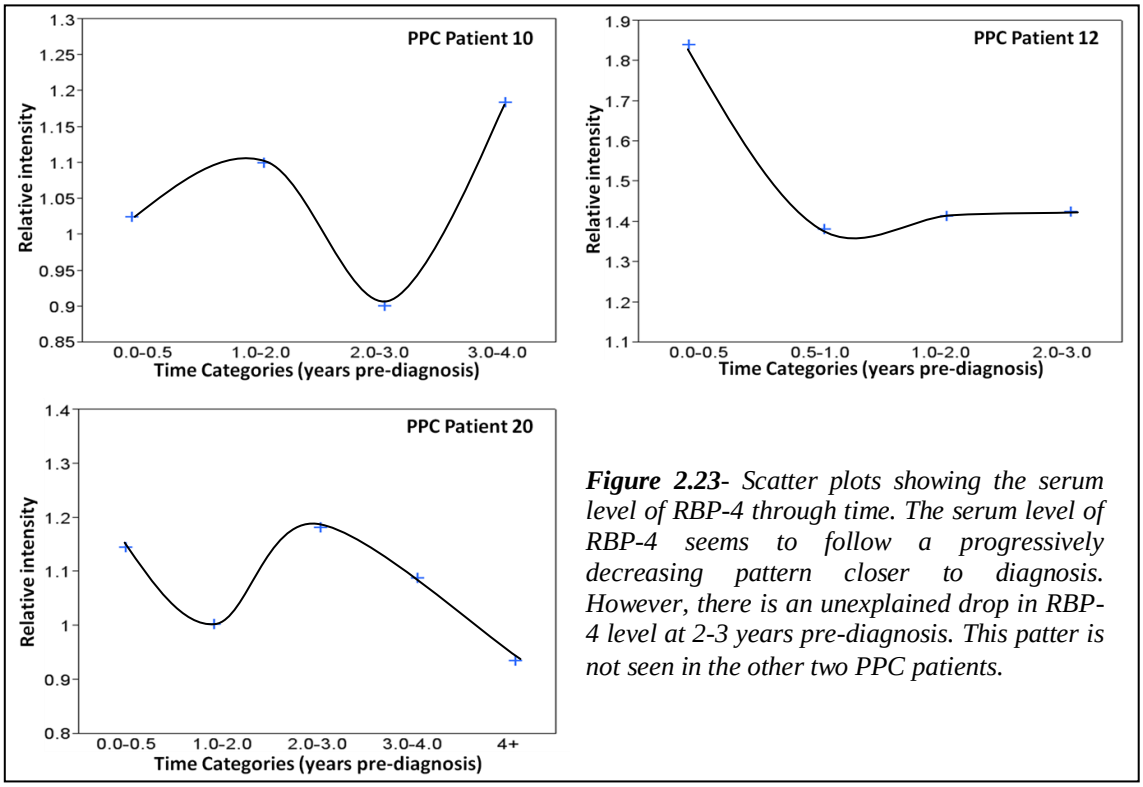


Figure 2.23- *Scatter plots showing the serum level of RBP-4 through time. The serum level of RBP-4 seems to follow a progressively decreasing pattern closer to diagnosis. However, there is an unexplained drop in RBP-4 level at 2-3 years pre-diagnosis. This patter is not seen in the other two PPC patients.*

2.4.4 Expression of VDBP, RBP-4, and FINC in cell lines

2.4.4.1 Quantification of cell lysate

The protein concentrations of cell lysates were quantified by Bradford Assay. Linear modelling of the reference standards generated a linear equation with a R^2 value of 0.998 therefore indicating that the linear equation fits the data almost perfectly. The concentration of each lysate was calculated using this linear equation and the volume of lysate required for 10µg of protein was calculated for each lysate.

Table 2.20- *Protein quantification of cell lysate by Bradford Assay*

Cell-line	Av. Bradford	Concentration µg/µl	Volume (µl)/ 10 µg
HEPG-2	1.08	6.74	1.48
HEK	0.65	2.01	4.97
BXPC-3	0.67	2.20	4.54
ASPC	1.20	8.08	1.24
PANC-1	1.02	6.09	1.64
CFPAK	0.94	5.18	1.93
FAMPAC	0.76	3.26	3.06

Note: *Protein Concentration= 0.5*(Bradford assay reading-0.4642)/ 0.0457;*
Reference standard curve R^2= 0.9976

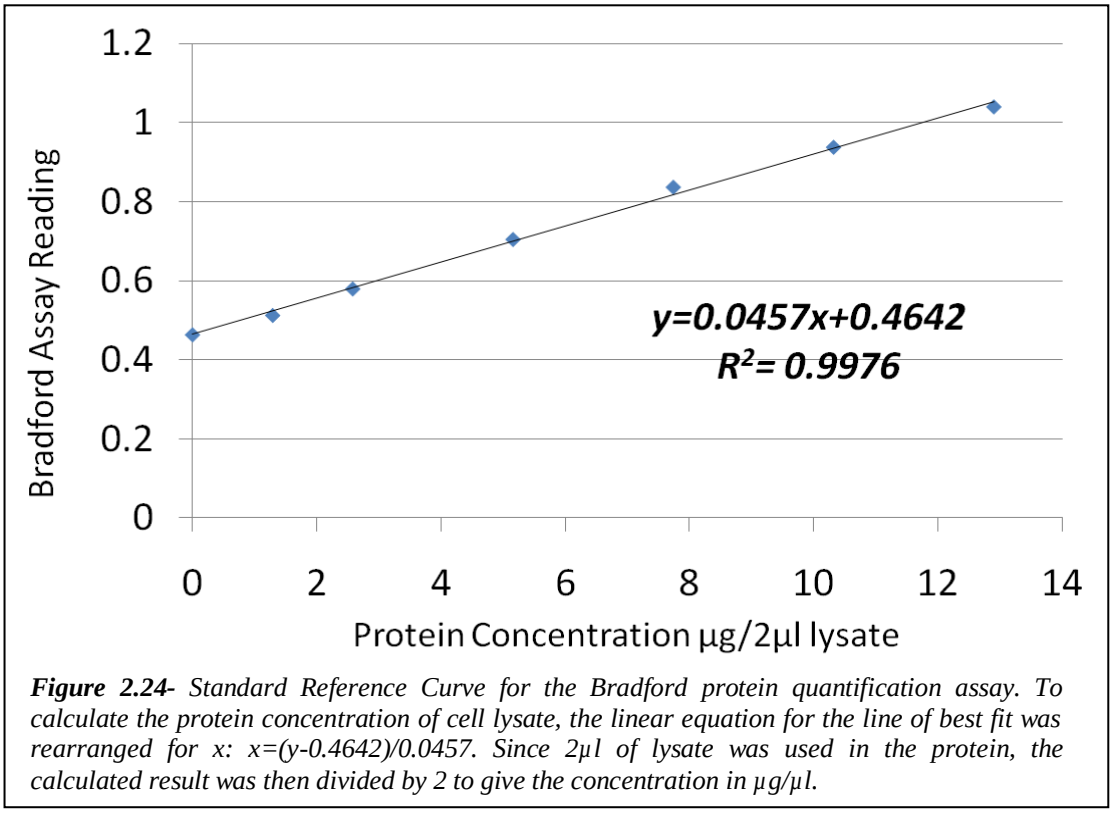


Figure 2.24- *Standard Reference Curve for the Bradford protein quantification assay. To calculate the protein concentration of cell lysate, the linear equation for the line of best fit was rearranged for x: x=(y-0.4642)/0.0457. Since 2µl of lysate was used in the protein, the calculated result was then divided by 2 to give the concentration in µg/µl.*

2.4.4.2 Expression of VDBP, RBP-4 and FINC in cancer cell lines

Western blot analysis of the two non-cancer cell lines HEPG-2 (hepatocellular cancer) and HEK (human embryonic kidney) and five pancreatic cancer cell lines BXPC-3, ASPC, PANC-1, CFPAK, and FAMPAC showed that VDBP, RBP-4, and FINC were present in an undetectable quantity in these cell lines (***Figure 2.25)***.

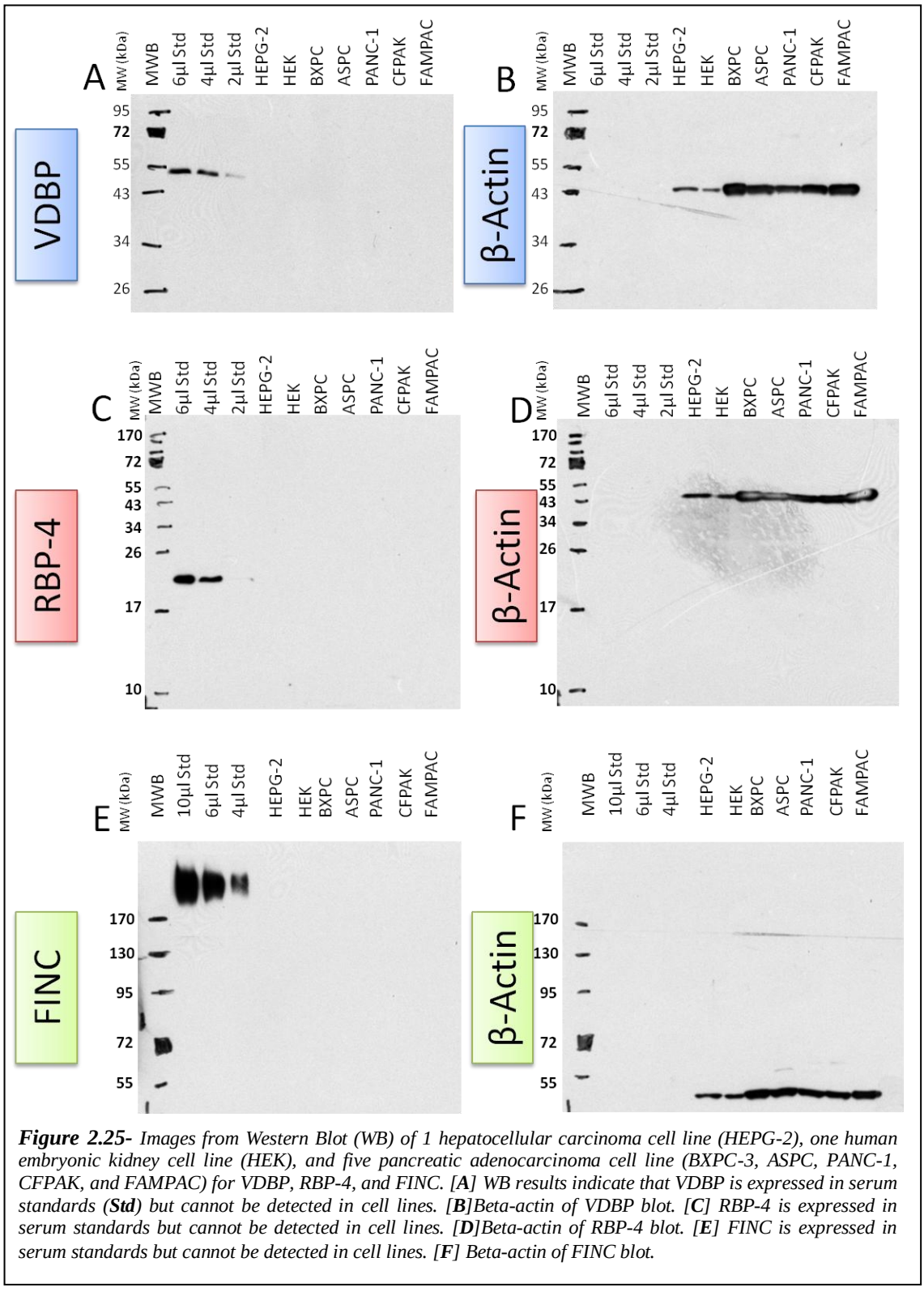


Figure 2.25- *Images from Western Blot (WB) of 1 hepatocellular carcinoma cell line (HEPG-2), one human embryonic kidney cell line (HEK), and five pancreatic adenocarcinoma cell line (BXPC-3, ASPC, PANC-1, CFPAK, and FAMPAC) for VDBP, RBP-4, and FINC. [**A**] WB results indicate that VDBP is expressed in serum standards (**Std**) but cannot be detected in cell lines. [**B**]Beta-actin of VDBP blot. [**C**] RBP-4 is expressed in serum standards but cannot be detected in cell lines. [**D**]Beta-actin of RBP-4 blot. [**E**] FINC is expressed in serum standards but cannot be detected in cell lines. [**F**] Beta-actin of FINC blot.*

2.5 Discussion

Recent advances in protein identification techniques have enabled researchers to mine the serum proteome in greater depth. Indeed, the advent of protein detection techniques such as iTRAQ/MS and protein microarray has opened up a new dimension in biomarker research. These new techniques have proven to be as reliable, accurate, and sensitive for the discovery of less abundant and smaller proteins compared to traditional methods such as spot identification for 2D polyacrylamide gel electrophoresis[128]. In particular, the mass spectrometry analysis performed by Murray *et al.* from this department has identified over 300 proteins with over 95% identification confidence including many novel proteins, which have never been associated with pancreatic cancer by previous studies.

2.5.1 VDBP, RBP-4, and FINC for the diagnosis of pancreatic cancer

The current study sought to verify the serum concentrations of three protein markers (VDBP, RBP-4, and FINC) identified by the iTRAQ experiment mentioned above. I observed that in concordance with the original iTRAQ results, western blot analyses of these proteins demonstrated a general decrease in relative serum concentrations of all three proteins in individuals with pancreatic cancer compared to controls (Phase I, ***section 2.4.1***). Furthermore, I found that in agreement with existing literature[225], validation with the LPCD independent sample set indicated that the presence of biliary obstruction has a negative impact on the diagnostic accuracies of all three markers (Phase II, ***sections 2.4.2.6 to 2.4.2.10***). This finding once again emphasised the importance of recognising the potential confounding effects of benign pancreaticobiliary diseases (e.g. biliary obstruction) in the discovery of serum proteomic biomarker for pancreatic cancer. Nonetheless, preliminary results from ***sections 2.4.2.6*****Error! Reference source not found.**, ***2.4.2.7, and 2.4.2.10*** suggest that VDBP, RBP-4, and FINC are accurate diagnostic markers of PDAC against individuals with CP or HC. In addition, we report that the use of the M-LR model to combine all three markers dramatically improved the diagnostic accuracy for PDAC in the absence of DC. However, further validation with an independent validation sample set would be necessary to assess the integrity of the individual and combined markers.

2.5.2 VDBP and RBP-4 as screening modalities for pancreatic cancer

Pancreatic cancer is a disease of late presentation where the majority of individuals with PDAC have metastatic, inoperable disease at diagnosis[54-55]. As such, much effort has been put into the search for a marker, which can detect the presence of early-stage PDAC with the ultimate aim of implementing a nation-wide screening programme for pancreatic cancer. Such implementation would facilitate the early detection of PDAC in otherwise asymptomatic individuals thereby increasing the chance of resectable disease, which in turn, would improve the prognosis for these patients.

Results from the current study thus far were based on the use of serum samples from patients with confirmed resectable (therefore by mostly early-stage) PDAC. Whilst biomarkers discovered and verified in this way can be concluded to be good indicators of early-stage PDAC, their potential as a screening modality remain untested. Phase-III of the current study sought to address this issue by using samples from the UKCTOCS trial, which consisted of serum samples taken from up to 6 years before the confirmed diagnosis of PDAC. Results from this analysis indicated that neither VDBP nor RBP-4 was sufficiently accurate as single markers of PDAC in the pre-diagnosis setting.

There were several limitations to this study. In particular, a relatively small number of samples were analysed in the first validation set meaning that the current study should be regarded as a pilot screen of candidate markers aimed at assessing their potential utility as markers of pancreatic cancer and further validation with a larger independent sample set would be required to confirm our findings. Furthermore, the use of Western Blotting for the quantitative protein analysis has its own inherent limitations, such as the need to compare values across different blots.

2.5.3 The roles of VDBP, RPB-4, and FINC in pancreatic cancer

The exact roles of VDBP, RPB-4, and FINC in pancreatic cancer and the reason behind their apparent decrease in the serum of individuals with PDAC are not clear. However, there are a number of hypotheses, which may explain the inverse relationship between the serum concentrations of the three proteins and the presence of PDAC. One explanation could simply be that patients with PDAC are generally under-nourished (e.g. due to anorexia or impaired absorption), which subsequently result in a decreased production of these proteins[21]. Furthermore, the fact that the 5-year survival rate of resectable pancreatic cancer is <25% due to cancer recurrence and metastasis[88] suggest that most patients with resectable PDAC may have an underlying micro-metastasis undetectable by conventional diagnostic procedures at the time of surgery. Therefore, specifically for the two hepatocyte-secreted proteins, VDBP and RBP-4, another possible explanation for the decrease in serum VDBP and RBP-4 could be that the production of both proteins is impaired in the presence of pancreatic liver metastasis.

2.5.3.1Vitamin D-Binding protein and cancer

There is a current lack of evidence associating VDBP specifically with pancreatic cancer. However, increasing evidence has suggested that VDBP may be associated with several types of cancers through its metabolite, DBP-MAF[226-229]. In particular, an *in vivo* study by Kisker *et al.* demonstrated that BxPC-3, a pancreatic cancer cell line, is capable of converting VDBP to DBP-MAF and in addition, the systemic administration of DBP-MAF (4ng/kg) has significant antiproliferative and antiangiogenic effects on immune-compromised mice implanted with BxPC-3[227]. The authors however, did not discuss the apparent contradiction in the ability of a cancer cell line to convert DBP into an anti-tumour factor. One possible explanation is that pancreatic cancer is characterised by an unusually hypoxic condition and as such, very limited amount of systemic DBP is converted into DBP-MAF and subsequently does not have a noticeable effect on the tumour. Another possible explanation is that VDBP may be readily converted to DBP-MAF and that this stimulates tumour-associated macrophages in the tumour microenvironment to promote growth and proliferation.

2.5.3.2 Retinol-Binding protein and cancer

The association between retinol-binding proteins (RBP) and pancreatic cancer was first described in 1984 by Fabris *et al.*[212]. The authors observed that the serum concentration of RBP is significantly lower in patients with chronic pancreatitis and pancreatic cancer compared to healthy subjects[212]. Furthermore, this association was also observed by later studies in a number of other epithelial cancers and benign diseases such as malnutrition and liver cirrhosis [220, 230-232]. Although the findings originally reported by Fabis *et al.*[212] was not fully recapitulated by this study, as I found that the levels of RBP-4 were down regulated in PDAC compared to HC but not in CP compared to HC, this difference may be accountable by the fact that the current study analysed a specific member of the RBP family (RBP-4) rather than RBP in general.

Interestingly, recent studies have focused on the role of RBP in type II diabetes and insulin resistance where the authors described an inverse association between the serum concentration of RBP-4 and insulin sensitivity [233-237]. This association may be of biological relevance to the poor glycaemic control in pancreatic cancer. Particularly in patients with tumour-induced dysfunction of pancreatic endocrine cells where there is an increased insulin resistance, a higher sensitivity to insulin would be desirable, which would correlate to decrease in circulating RBP. Therefore, combining the two observations given above, we propose that as a part of a complex compensatory mechanism for the decreased insulin secretion in PDAC, hepatocyte may inhibit the production of RBP-4 to increase the sensitivity of the liver to insulin.

2.5.3.3 Fibronectin and cancer

Fibronectin is an essential component of the extracellular matrix (ECM). It has been associated with tissue fibrosis and several types of cancer through its role in controlling cellular proliferation, differentiation, and organization of tissue architecture[147, 154, 185, 238-239]. In particular, several studies have demonstrated that there is extensive remodelling of the ECM in pancreatic cancer by fibroblasts and pancreatic stellate cells (PSCs), which may induce a permissive microenvironment for the transition of cancer cells from dormancy to growth[147, 152-154, 185, 238-240]. Furthermore, the release of angiogenic factors from the remodelled ECM may trigger the angiogenic switch thereby promoting tumour growth and metastasis[198]. Recently,

evidence indicated that pancreatic cancer-associated fibrosis might be partially attributable to elevated levels of TGF-β, which alters the ECM composition by promoting collagen and fibronectin gene transcription, production, and secretion in PSCs[185, 238, 241-242]. Therefore, it is conceivable that the decrease in the serum concentration of FINC may be due to an excess consumption of circulating FINC as the result of extensive remodelling of the tumour microenvironment by tumour-associated cells such as PSCs.

Finally, the finding that all three proteins were down regulated in individuals with benign biliary obstruction may provide some indication as to why they are observed to be down regulated in PDAC. Over 70% of patients with PDAC have biliary obstruction at presentation[21]. Such obstructions are frequently associated with the impairment in liver function, which may account for the decrease in the three proteins observed in PDAC.

Chapter 3

The Diagnostic Potential of CCGFs for Pancreatic Cancer

3.1 Introduction

The role of cytokines, chemokines, and growth factors (CCGFs) in pancreatic inflammatory diseases (e.g. acute and chronic pancreatitis) has been reported by many studies[138-144]. Although there is a current lack of evidence directly associating CCGFs with pancreatic cancer, recently, evidence from clinical and experimental studies have demonstrated that CCGFs play a pivotal role in mediating cancer-related inflammation[144-150], which in turn is an essential component for the formation, propagation, survival, and metastasis of several other types of cancers[143, 151-154]. In 2000, Hanahan *et al.*[45] proposed in their review that there are six hallmarks of cancer, of these, five hallmarks were mediated by CCGFs including self-sufficiency in growth signals, insensitivity to anti-growth signals, limitless replicative potential, sustained angiogenesis, and evading apoptosis. Recent studies have now associated the sixth hallmark (tissue invasion and metastasis) with CCGFs by demonstrating their role in stimulating tumour-associated macrophages to promote the remodelling of the extracellular matrix in the tumour microenvironment[187, 198]. More importantly, results from a recent review by Colotta *et al.*[150] would suggest that cytokine mediated cancer-related inflammation should be regarded as the seventh hallmark of cancer.

Therefore, based on the evidence above, we postulate that CCGFs should also be heavily involved in pancreatic cancer and that certain CCGFs would be differentially expressed only in cancer-related inflammation and not benign inflammatory conditions. To test this hypothesis, we utilise a relatively new multiplex assay, the LUMINEX cytokines assay, to quantify and compare the presence of CCGFs in serum samples from four distinct disease groups: pancreatic cancer, chronic pancreatitis, biliary obstruction, and healthy controls. The multiplex cytokines assay combines several existing technologies including microsphere, flow cytometry, conventional ELISA, digital, traditional chemistry, digital signal processing, and dual laser detection to quantify the serum concentrations of 27 different CCGFs simultaneously in a single sample. Furthermore, we will use the resulting data was used to assess the diagnostic potential of CCGFs as individual markers of resectable PDAC. In addition, we will present a statistical method for the selection of independent biomarkers amongst a large number of potential biomarkers and two independent methods, which can be employed to combine the selected markers in to a single diagnostic marker panel for PDAC.

3.2 Study Aims

The aims of the current chapter include:

- To determine the serum expression of 27 CCGFs in resectable PDAC and Control subjects.
- To determine the diagnostic potential of CCGFs for resectable pancreatic cancer
- To compare the diagnostic potential of novel CCGF markers against CA19-9
- To determine the effects of combining candidate CCGF markers on the diagnostic accuracy for pancreatic cancer
- To validate the disease-predicting algorithms generated in aim (4) using an independent validation sample set.

3.3 Materials and Methods

3.3.1 Patients and Samples- The Liverpool Pancreatic Cancer Database

One-hundred and eighty pre-operative serum samples were prospectively collected at the Royal Liverpool University Hospital from patients with resectable PDAC (n=90), CP (n=30), obstructive jaundice (DC, n=30), and healthy volunteers (HC, n=30) between 1996 and 2010; ***Table 3.1***). Serum samples were collected in Sarstedt Monovette tubes (Sarstedt Ltd, Leicester, UK) and allowed to clot at room temperature for 15 minutes. The serum fraction was acquired by centrifugation at 800x g for 10 minutes, which was then aliquoted into cryotubes (Nunc GmbH & Co KG., Thermo Fisher Scientific, Langenselbold, Germany). All samples were stored at minus 80℃ until analysis.

Of the 90 individuals in the control group, 44 were female and 46 were male and of the 90 individuals with PDAC, 42 were female and 48 were male. Although there was a significant difference in age between the PDAC group and the individual control subgroups groups (CP, DC, and HC), there was no significant difference when the PDAC group was analysed against the overall Control group (Wilcoxon's Test, P=0.073; ***Table 3.1***). Furthermore, analysis of available demographical data showed that there was no significant difference in gender, histories of smoking, diabetes, or pancreatitis in the PDAC group compared to the controls (Fisher's Exact,

p>0.05). In addition, comparison of the serum bilirubin levels indicated that individuals in the PDAC and DC groups have a significantly higher serum concentration of bilirubin compared to the CP group. Furthermore, the serum CA19-9 levels in the PDAC group is significantly higher compared to all control subgroups (Wilcoxon's Test, p<0.001; ***Table 3.1***). It should be noted that some demographical data could not be obtained due to privacy and confidentiality reasons.

In subsequent analyses, samples within each disease group were randomised into the *Discovery Phase* and the *Validation Phase* in a 2 to 1 ratio (see ***sections 3.3.1.1***, ***3.3.1.2***, and ***3.3.2*** for details).

Table 3.1- *Demographical and clinical characteristics of individuals involved in this study*

	Disease groups					Cancer status		
Characteristics	**CP**	**DC**	**HC**	**PDAC**	**p (K-W)**	**Control**	**PDAC**	**p (W)**
Count, n	30	30	30	90	-	90	90	-
Age								
Median	60	72.5	56.4	66.5	**<0.001**	62	66.5	N.S.
Range	50-78	24-86	37-77	36-81		24-86	36-81	
Gender, n								
Female/male	14/16	15/15	15/15	42/48	N.S.	44/46	42/48	N.S.
Smoking								
Yes	19/6	20/7	-/-	51/22	N.S.	39/13	51/22	N.S.
Unknown	5	3	30	17		38	17	
Diabetes								
Yes	7/18	4/23	-/-	14/59	N.S.	11/41	14/59	N.S.
Unknown	5	3	30	17		38	17	
Alcohol consumption								
Yes	16/8	17/10	-/-	55/18	N.S.	33/18	55/18	N.S.
Unknown	6	3	30	17		39	17	
Pancreatitis								
Yes	30/0	3/0	-/-	4/43	-/-	33/0	4/43	-/-
Unknown	0	27	30	43		57	43	
Bilirubin								
Median	6	34	-/-	46.5	**<0.001**	15	46.5	**<0.001**
Range	1-261	6-643	-/-	5-448		1-643	5-448	
CA19-9								
Median	21.5	24.5	4.5	656.5	**<0.001**	13.4	188.1	**<0.001**
Range	0-331	0-1402	0-18	0-42094		0-1402	0-42094	

Abbreviations*: chronic pancreatitis (**CP**); disease controls (**DC**); healthy controls (**HC**); pancreatic cancer (**PDAC**); Kruskal Wallis's test [**p (K-W)**]; Wilcoxon's Test [**p (W)**]; p-value <0.05 (**N.S.**); data unavailable (**-/-**)*

3.3.1.1 Discovery Phase patient demographics and sample characteristics

One-hundred and twenty samples including sixty individuals with PDAC and twenty individuals from each of the three control subgroups (CP, DC, and HC) were randomly chosen and analysed in *Discovery Phase (*Table 3.2*)*. Although a significant difference in age was observed between HC and other disease subgroups, there were no significant difference in age, gender, histories of smoking, diabetes and alcohol consumption between the PDAC group and the combined Control group. As expected, the serum levels of bilirubin and CA19-9 were significantly higher in the PDAC group compared to the overall control group.

Table 3.2- *Demographical and clinical characteristics of individuals in the Discovery Phase*

Characteristics	**Disease groups**					**Cancer status**		
	CP	**DC**	**HC**	**PDAC**	**p (K-W)**	**Control**	**PDAC**	**p (W)**
Count, n	20	20	20	60	-	60	60	-
Age								
Median	60	74	56	67	**0.002**	62	67	N.S.
Range	50-78	45-86	37-77	36-81		37-86	36-81	
Gender, n								
Female/male	8/12	11/9	10/10	31/29	N.S.	29/31	31/29	N.S.
Smoking								
Yes/No	12/4	13/5	-/-	34/15	N.S.	25/9	34/15	N.S.
Unknown	4	2	20	11		26	11	
Diabetes								
Yes/ No	5/11	4/14	-/-	12/28	N.S.	9/25	12/28	N.S.
Unknown	4	2	20	10		26	10	
Alcohol consumption								
Yes/ No	10/5	12/6	-/-	34/15	N.S.	22/11	34/15	N.S.
Unknown	5	2	20	11		27	11	
Pancreatitis								
Yes/ No	20/0	1/0	-/-	3/24	-/-	21/0	3/24	-/-
Unknown	0	19	20	23		39	23	
Bilirubin								
Median	8	30	-/-	38	**<0.001**	14	38	**<0.0028**
Range	4-58	6-585	-/-	5-379		4-585	5-379	
CA19-9								
Median	21	44	5	293	**<0.001**	15	293	**<0.001**
Range	0-188	4-1402	0-18	0-42094		0-1402	0-42094	

Abbreviations*: chronic pancreatitis (**CP**); disease controls (**DC**); healthy controls (**HC**); pancreatic cancer (**PDAC**); Kruskal Wallis's test [p (**K-W**)]; Wilcoxon's Test [**p (W)**]; p-value <0.05 (**N.S.**); data unavailable (**-/-**)*

3.3.1.2 Validation Phase patient demographics and sample characteristics

The *Validation Phase* utilized sixty independent samples for the validation of the two disease-predicting algorithms generated in the *Discovery Phase* (Table 3.3). The sample set includes thirty individuals with PDAC and ten individuals from each of the control sub-groups (CP, DC, and HC). Again, whilst a significant difference in age was observed between the HC and the other disease groups, there were no statistical difference in age, gender, histories of smoking, diabetes, and alcohol consumption between individuals with PDAC and the combined Control group. Similar to the *Discovery Phase*, there was a significant increase in serum concentration of bilirubin and CA19-9 in individuals with PDAC compared to the Controls group.

Table 3.3- *Demographical and clinical characteristics of individuals in the Discovery Phase*

	Disease groups					Cancer status		
Characteristics	**CP**	**DC**	**HC**	**PDAC**	**p (K-W)**	**Control**	**PDAC**	**p (W)**
Count, n	10	10	10	30	-	30	30	-
Age								
Median	60	73	57	66	**0.029**	61	66	N.S.
Range	52-77	24-84	44-71	45-77		24-84	45-77	
Gender, n								
Female/male	6/4	4/6	5/5	11/19	N.S.	15/15	11/19	N.S.
Smoking								
Yes/No	7/2	7/2	-/-	17/7	N.S.	14/4	17/7	N.S.
Unknown	1	1	10	6		2	6	
Diabetes								
Yes/No	2/7	0/9	-/-	2/1	N.S.	2/16	2/1	N.S.
Unknown	1	1	10	7		12	7	
Alcohol consumption								
Yes/No	6/3	5/4	-/-	21/3	N.S.	11/7	21/3	N.S.
Unknown	1	1	10	6		12	6	
Pancreatitis								
Yes/No	10/0	1/0	-/-	1/17	-/-	11/0	1/17	-/-
Unknown	0	9	10	12		19	12	
Bilirubin								
Median	6	51	-/-	52	**0.005**	22	52	**<0.0028**
Range	1-261	6-643	-/-	11-448		1-643	11-448	
CA19-9								
Median	24	22	4	125	**<0.001**	12	125	**<0.001**
Range	5-331	0-265	0-15	5-2749		0-331	5-2749	

Abbreviations*: chronic pancreatitis (**CP**); disease controls (**DC**); healthy controls (**HC**); pancreatic cancer (**PDAC**); Kruskal Wallis's test [p (**K-W**)]; Wilcoxon's Test [**p (W)**]; p-value <0.05 (**N.S.**); data unavailable (**-/-**)*

3.3.2 Study design

Approval for the current study was obtained from the relevant research ethics committee including the multicentre research ethics committee (MREC) for the use of the serum samples from the Liverpool Pancreatic Cancer Database (LPCD). Informed consent was obtained from all individuals involved.

The serum concentrations of CCGFs in one hundred and eighty serum samples were quantified using the BioRad Pro 27-Plex Human Cytokines Assay. Basic analyses (including median, univariate, multivariate, and ROC analyses) were performed to determine the serum expression and diagnostic accuracies of each CCGF for PDAC against the following disease groups: HC, CP, DC, and all controls. Subsequent analyses were designed to evaluate the impact of combining several CCGF markers into a single combined marker on the diagnostic accuracy for PDAC. This was achieved by dividing the study into a *Discovery Phase* and a *Validation Phase* (***Figure 3.1***).

In the *Discovery Phase*, CCGF data from 120 randomly selected serum samples were analysed by stepwise regression model to select candidate markers for combination by two independent modelling methods (multinomial logistic regression model [**M-LR**] and artificial neural network model [**NN**]) to generate two independent disease-predicting algorithms. The accuracies of the disease-predicting algorithms for DPAC were represented by ROC-AUC. In the *Validation Phase*, the two disease-predicting algorithms were directly applied to the CCGF data from the sixty remaining samples and their diagnostic accuracies assessed by ROC.

A separate analysis was performed to compare the diagnostic accuracy of the current biomarker standard, CA19-9, against those of the individual and combined CCGF markers. This was achieved by completing the quantification of serum CA19-9 using a CA19-9 ELISA assay and by statistical comparison using ROC-AUCs, sensitivities, and specificities. Furthermore, the effects of including CA19-9 as the fifth marker in the generation of the disease-predicting algorithms on the diagnostic accuracy in the Discovery and *Validation Phase*s will be determined. Finally, the impact of each candidate biomarker on the M-LR CCGF algorithm and the M-LR CCGF-CA19-9 algorithm will be evaluated by comparing the diagnostic accuracies of the algorithms when each of the four CCGFs was independently removed (***section 3.4.5***).

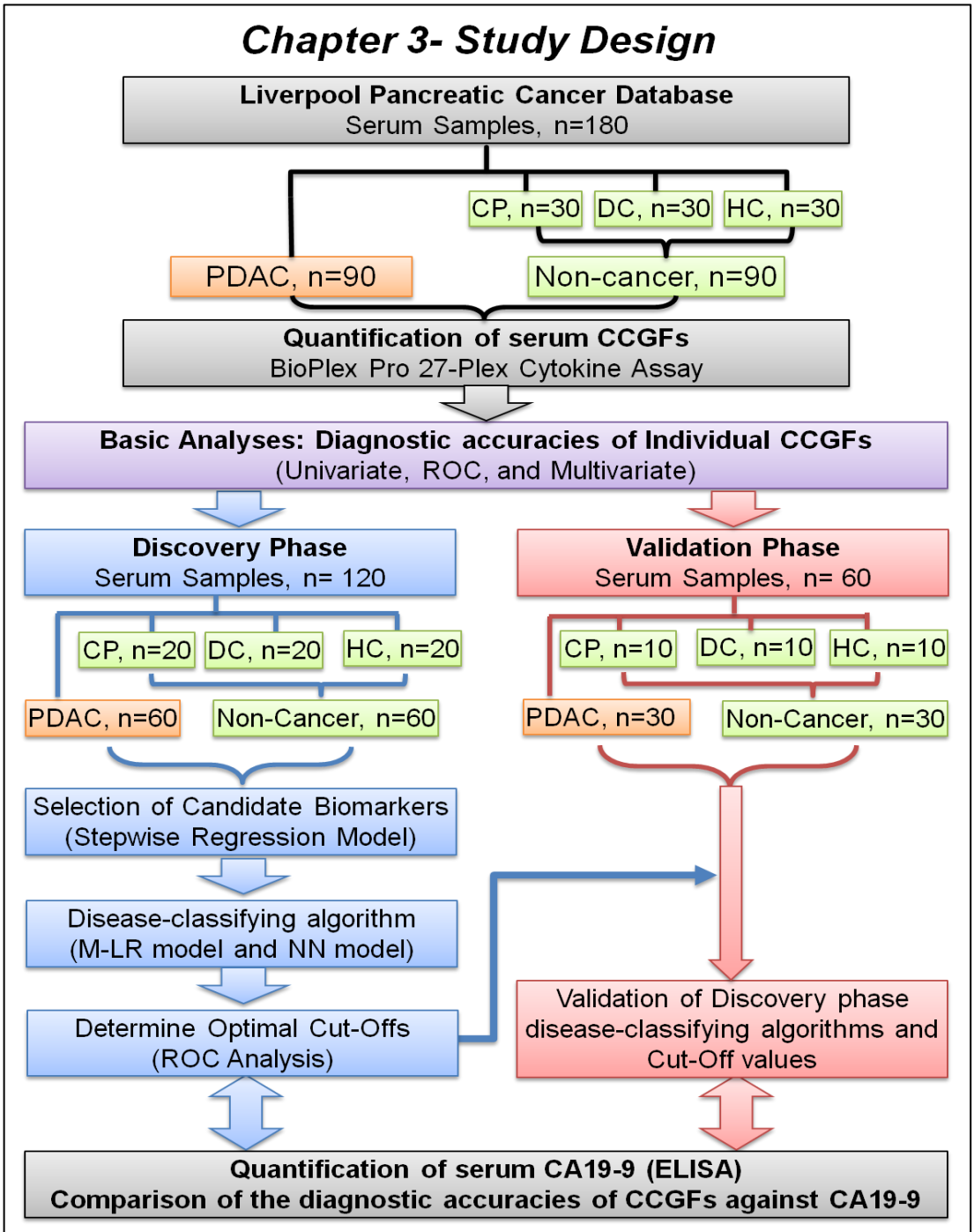


Figure 3.1- Study design. *180 serum samples from the Liverpool Pancreatic Cancer database were quantified for 27 serum CCGFs in this study, of these, 90 were PDAC, 90 were non-cancer: 30 CP, 30 DC, and 30 HC. A basic analysis was performed to determine the serum expression and diagnostic accuracy of CCGFs for pancreatic cancer against all control groups using univariate, multivariate, and ROC analyses. The study is then divided into 2 Phases. In the* ***Discovery Phase,*** *CCGF data from 120 serum samples was used to select potential CCGF markers for combination by* ***stepwise regression****. A disease predicting algorithm was then generated using two independent modelling methods (multinomial logistic regression [****M-L****] and artificial neural network [****NN****] models). Finally, the diagnostic accuracies of the two models were determined using ROC. The* ***Validation Phase*** *aimed to validate the diagnostic accuracies of CCGFs as individual markers of PDAC and to validate the Discovery Phase disease-predicting algorithms using an independent sample set consisting of 60 serum samples. Results were then compared to CA19-9 in a separate section.*

3.3.3 Quantification of Cytokines, Chemokines, and Growth Factors

A panel of 27 CCGFs were simultaneously quantified by the BioPlex Pro 27-Plex Human Cytokines Assay kit (BioRad Laboratories Inc, California, USA). The CCGFs analysed were Platelet-Derived Growth Factor (PDGF), Interleukin (IL)1β, IL-1 receptor alpha (IL-1Rα), IL-2, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-12, IL-13, IL-15, IL-17, Eotaxin, Basic Fibroblast Growth Factor (FGF-b), Granulocyte Colony-Stimulating Factor (G-CSF), Granulocyte-Macrophage Colony-Stimulating Factor (GM-CSF), Interferon Gamma (IFN-γ), Monocyte Chemotactic Protein-1 (MCP-1), Macrophage Inflammatory Protein 1-alpha (MIP-1α), Macrophage Inflammatory Protein 1-beta (MIP-1β), Interferon Gamma Inducible Protein-10 (IP-10), Vascular Endothelial Growth Factor (VEGF), Tumour Necrosis Factor-alpha (TNF-α), and Regulated Upon Activation Normal T-Cell Expressed And Secreted Protein (RANTES). This BioPlex Pro 27-Plex Human Cytokines Assay kit provides a mixture of 27 magnetic micro-bead populations pre-coated with antibodies specific to a corresponding CCGF. Each bead population is internally dyed with a mixture of red and infrared dye, which produces a unique signature that can be identified by the dual laser technology in the Luminex-200 system. Serum samples were analysed in duplicate using a standardised assay protocol:

3.3.3.1 Sample and Standard Preparation

All samples and standards were thawed and kept on ice until ready to use. 30µl of each serum sample was diluted by a factor of four using 90µl of sample diluent. Premixed lyophilized standard was reconstituted in 500µL of sample diluent, vortexed for 3 seconds then incubated on ice for 30 minutes. 128µl of reconstituted standard was added to 72µl of standard diluent in a 1.5ml Eppendorf microfuge tube followed by seven 1/4 serial dilutions to generate eight standard references. 200µl of standard diluent was aliquoted for use as a "Blank" sample (***Figure 3.2***).

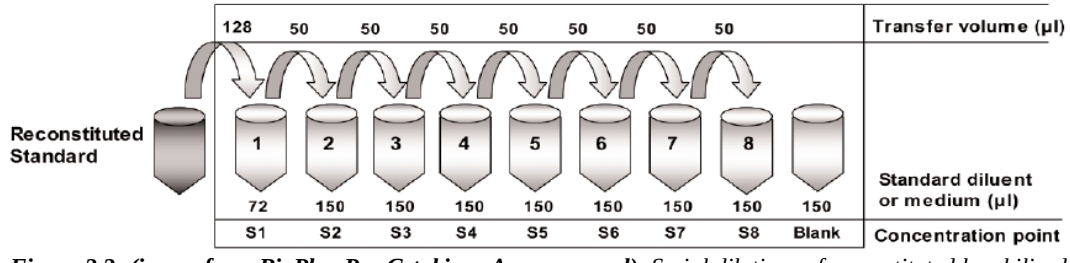


Figure 3.2- (image from BioPlex Pro Cytokines Assay manual). *Serial dilutions of reconstituted lyophilized standards. 128 µl of Reconstituted standard was added to 72µl of standard diluent followed by seven 1/4 serial dilutions to generate 8 standard references. 150µl of standard diluent was used for blank sample.*

3.3.3.2 Coupled magnetic beads, detection antibody, and Streptavidin-PE preparation

Coupled magnetic beads (10x) were centrifuged for 30 seconds and then re-suspended using a pipette. 575µl of coupled beads (10x) were diluted to 1x concentration using 5175µl of assay buffer. Reconstituted coupled beads were kept on ice and protected from light until ready for use.

The detection antibody was prepared when the magnetic coupled beads were incubated with standard/sample. Detection antibody (10x) was centrifuged for 30 seconds before 300µl of detection antibody (10x) was diluted with 2700µl of detection antibody diluent into 1x concentration. The diluted serum detection antibody was protected from light until ready for use.

The Streptavidin-PE was prepared when the magnetic coupled beads were incubated with detection antibody. Streptavidin-PE (100x) was centrifuged for 30 seconds before 60µl of Streptavidin-PE (100x) was diluted with 5940µl of assay buffer into 1x concentration. The diluted Streptavidin-PE (1x) was protected from light until ready for use.

3.3.3.3 Assay Procedure

The 96-well multiplex plate was pre-wet with 100µl of assay buffer. The buffer was removed by vacuum manifold and the bottom of the plate was dried using a clean paper towel. The coupled beads (1x) were vortexed for 30 seconds and 50µl was aliquoted into each of well. The plate was washed (and vacuum emptied) twice with 100µl of wash buffer. 40µl of prepared standards, blank, or serum samples were vortexed for 3 seconds before being added to each well, and incubated at room temperature on a shaker (900rpm) for 30 minutes. The plate was washed (and vacuum emptied) three-times with 100µl of wash buffer. 25µl of 1x detection antibody (for preparation see ***section 3.3.3.2***) was vortexed for 3 seconds before being added to each well and incubated on a shaker (900rpm) at room temperature for 30 minutes. The plate was washed (and vacuum emptied) three-times with 100µl of wash buffer. 50µl of Streptavidin-PE (1x) was vortexed for 3 seconds before being added to each well and incubated on a shaker (900rpm) at room temperature for 10 minutes. The plate was washed (and vacuum emptied) three-times with 100µl of wash buffer. 125µl of assay buffer was added to each well and shaken (900rpm) for

30 seconds at room temperature before detection by the LUMINEX-200 systems (Luminex Co, Austin, USA).

The LUMINEX-200 was calibrated per experiment and the Bio-Plex Manager Software Version 5.0 was used covert assay data to serum concentration in pg/ml.

3.3.4 Quantification of serum CA19-9 by ELISA

The serum concentration of CA19-9 in 141 serum samples has been previously measured, either at diagnosis or in the department by Dr Sarah Tonack. The current study quantified the remaining 39 samples using a CA19-9 ELISA kit (Alpha Diagnostics).

A 96-well streptavidin-coated ELISA plate was pre-wet with 200µl of wash buffer. Then 25µl of standards (provided by the kit) and sample were aliquoted into the appropriate wells in duplicate followed by 100µl of biotinylated anti-CA19-9 antibody. After the mixture was allowed to incubate at room temperature for 60 minutes, the content was discarded and the wells were washed five times manually using 200µl of wash buffer per well. 100µl of anti-CA19-9-HRP conjugate was then added to each well. The content was mixed and allowed to incubate at room temperature for a further 60 minutes. Once completed, the wells were washed manually five times using 200µl of wash buffer and incubated with 100µl of chromogenic HRP substrate solution for 30 minutes at room temperature. This will cause a varying degree of colour shift from clear to blue, which is directly proportional to the concentration of CA19-9. After incubation, the reaction was stopped by adding 50µl of stop solution to all wells and mixed gently for 3-10 seconds. The stop solution changes the colour of the solution from blue to yellow and the intensity of yellow is again proportional to the serum concentration of CA19-9. Finally, an ELISA reader was used to quantify the colour intensity at a wavelength of 450nm.

The ELISA readings were exported to Microsoft Excel for further analysis. A linear equation was generated using the ELISA readings of the standard references with known concentrations of CA19-9. This equation was then used to convert ELISA readings from the samples into estimated serum CA19-9 concentrations.

3.3.5 Statistical Analysis

3.3.5.1 General statistics, univariate, and multivariate analyses

All continuous data (e.g. Age, serum analyte levels) were classified as non-parametric data and were summarised using median and inter-quartile range (IQR). The Kruskal-Wallis test was used to determine whether there is a significant difference in non-parametric distribution between multiple (>2) groups (e.g. age versus PDAC, CP, HC, and DC). For univariate analysis, the Wilcoxon's test was used to assess any difference in the distribution of a non-parametric variable (e.g. age) between two groups and the Fisher's Exact test was used to assess the difference between two categorical variables (e.g. gender versus PDAC/Control). Variables that were significant on univariate analysis were further tested using multivariate analysis to identify any independent variables. For univariate and multivariate analyses, a p-value of <0.05 was considered statistically significant.

In Hierarchical Cluster Analysis, data from each variable was first standardised by mean and standard deviation to give a value between -1 and 1. The standardised values were then analysed by Ward's minimal variance method and represented graphically in a heat map.

Data from the current study were also graphically represented using box plots, mosaic plots, logistic plots, and receiver-operator characteristic (ROC) curves.

3.3.5.2 Correlation analysis

The correlation between age and the relative serum concentration of CCGFs were assessed using Kendall Tau (τ) multivariate correlation analysis for non-parametric tied data. The Kendall τ coefficient ranges from -1 to 0 to 1 where a τ coefficient of -1 or 1 represents a perfect negative or positive correlation, respectively, and a τ coefficient of 0 indicates that there is no correlation. The associated p-value represents how confident the test is that the actual τ value would be more negative or more positive than then calculated τ value. In the current study, a Kendall Tau coefficient of >0.5 was regarded as significant.

3.3.5.3 Diagnostic accuracies of CCGFs for PDAC

The Receiver-Operator Characteristic (ROC) Area Under Curve (AUC) was used to assess the diagnostic accuracies of the markers. A ROC-AUC of >0.70 was considered statistically significant. An optimal cut-off value was obtained from the ROC analysis and the relative sensitivity and specificity at the optimal cut-off was reported.

3.3.5.4 Selection of candidate markers

Cytokine markers, which showed significance on univariate analyses were analysed in the *Discovery Phase* using the Stepwise Regression model with parameters of <0.05 for entering and >0.05 for leaving the model. Briefly, at each step, the cytokine marker with the lowest Wald/Score p-value that is also <0.05 is entered into the model. The Wald/Score p-values for each cytokine marker were then recalculated and a marker will leave the model if it has a p-value of >0.05. This process is iterated until no further cytokine marker can enter or leave the model.

3.3.5.5 Generating disease-predicting mathematical algorithms: M-LR and NN

The four most significant cytokine markers selected by the stepwise regression model were used to generate two independent disease-predicting algorithms using two separate modelling methods: Multinomial Logistic Regression (M-LR) model and Artificial Neural Network (NN) model. Both models are designed to estimate the probability of PDAC using the serum concentration of selected CCGF markers. The resulting probability value ranges from zero to one where an estimated probability value close to zero indicates that the sample is likely to be a non-cancer sample whereas a probability value closer to one suggests a pancreatic cancer sample.

Briefly, the M-LR model combines the markers using a simple logistic function with the appropriate coefficients. This function was then transformed mathematically into a value ranging from zero to one. The artificial neural network model utilizes multiple complex logistic functions. The input data (i.e. serum concentrations from selected CCGFs) were randomly entered into a pre-defined number of ‘hidden nodes’. At each hidden node, a logistic function was generated. The logistic functions from all the nodes were then combined into a single logistic function and transformed into a probability value, which ranges from zero to one.

3.3.5.6 Diagnostic accuracies of the models

The optimal cut-off values for the estimated probability of pancreatic cancer generated by the models were determined by ROC. A sample with an estimated probability above the cut-off value was classified as pancreatic cancer and a probability value below the cut-off was classified as control. The sensitivity and specificity of the predicted classifications were the then calculated and the results were graphically represented as mosaic plots.

$$\boldsymbol{Sensitivity} = \frac{Number\ of\ True\ Positives}{Number\ of\ True\ Positives + Number\ of\ False\ negatives}$$

$$\boldsymbol{Specificity} = \frac{Number\ of\ True\ Negatives}{Number\ of\ True\ Negatives + Number\ of\ False\ Positives}$$

3.3.5.7 Software for statistical analyses

All statistical analyses were performed using JMP version 8.02 (SAS, Buckinghamshire, UK). In addition, Microsoft Excel 2007 (Microsoft Limited, Berkshire, UK) was used to graphically represent the data.

3.4 Results

3.4.1 Basic analysis- CCGFs as individual markers of PDAC

3.4.1.1 Diagnostic accuracy of individual CCGFs for PDAC against HC

Results from the basic analysis of the twenty-seven CCGFs and CA19-9 serum concentration data obtained using the 180 serum samples are summarised in Table 3.4. Fourteen CCGFs and CA19-9 were observed to have a significantly different serum concentration between the PDAC group and the HC control subgroup, (Univariate by Wilcoxon's Test, $p<0.05$). In addition to CA19-9, eleven CCGFs were identified as up regulated in PDAC compared to HC including PDGF, IL-4, IL-6, IL-8, IL-9, IL-10, IL-17, IFN-γ, IP-10, MIP-1β and VEGF. Furthermore, three CCGFs were down regulated in PDAC compared to HC including IL-1β, IL-5, and IL-13.

Subsequent analyses assessed the diagnostic accuracies of the twenty-seven CCGFs and CA19-9 for differentiating between PDAC and HC. Results from ROC analysis showed that seven CCGFs and CA19-9 had a ROC-AUC of >0.7 and are therefore good discriminators between the two disease groups. In particular, IP-10, IFN-γ, and CA19-9 were highly accurate markers of PDAC with ROC-AUCs of 0.90, 0.91, and 0.95, respectively (***Table 3.4***).

We report that whilst IFN-γ achieved a sensitivity of 0.89 and a specificity of 0.87 at the optimal cut off value of 65.44 pg/ml, CA19-9 remained the most accurate single marker of PDAC against HC with a sensitivity of 0.92 and a specificity of 0.97 at a cut-off of 16 U/ml (***Figure 3.3***)

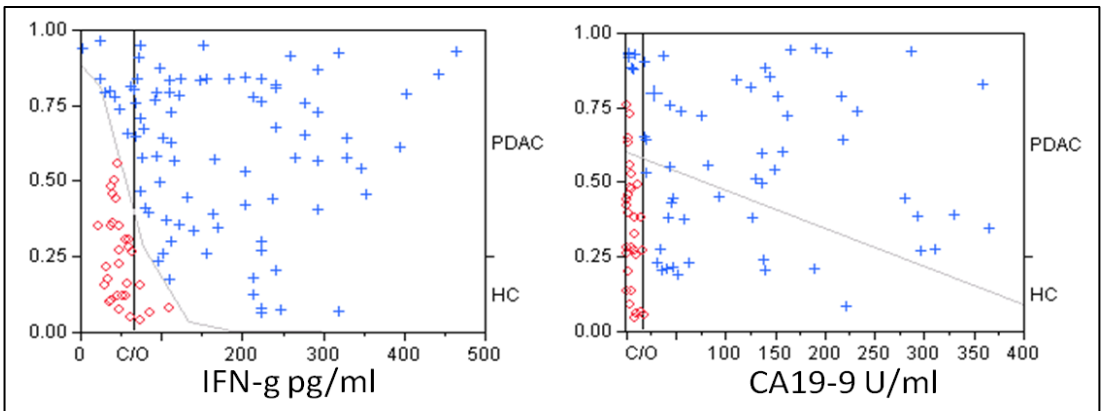


Figure 3.3- *Logistic Plot showing the diagnostic accuracies of IFN-g (left) and CA19-9 (right) for PDAC against HC. Optimal cut-off as determined by ROC (C/O). Individuals with PDAC are coloured* ***Blue*** *and individuals in the HC group are in* ***Red****.*

Table 3.4- *Diagnostic potential of CCGFs and CA19-9 for PDAC against HC*

Analyte	Quartile (pg/ml)						Univariate (p)	Reg.	Diagnostic accuracy			
	Healthy Control			PDAC								
	25th	Median	75th	25th	Median	75th			AUC	C/O	Sens.	Spec.
PDGF	9037.4	12048.2	16887.7	12415.5	15557.1	20338.0	**0.013**	Up	0.65	12463.50	0.76	0.53
IL-1β	1.1	1.5	1.9	0.0	0.0	0.0	**<0.001**	Down	0.19	2.41	0.14	0.97
IL-1Rα	40.2	76.5	108.1	32.1	70.8	160.3	0.687	-	0.53	112.74	0.36	0.80
IL-2	0.0	0.0	0.0	0.0	0.0	0.0	0.313	-	0.52	2.78	0.03	1.00
IL-4	1.3	1.7	2.0	1.5	3.4	6.2	**<0.001**	Up	**0.74**	**2.40**	**0.62**	**0.97**
IL-5	1.3	1.5	1.9	0.0	0.0	0.0	**<0.001**	Down	0.19	3.91	0.13	0.97
IL-6	1.9	3.3	5.4	2.6	9.5	21.0	**<0.001**	Up	0.72	**7.72**	**0.71**	**0.93**
IL-7	2.7	4.4	5.7	0.0	9.5	24.6	0.094	-	0.60	8.62	0.54	0.97
IL-8	6.8	8.3	10.7	10.7	23.6	40.3	**<0.001**	Up	**0.77**	**16.18**	**0.64**	**1.00**
IL-9	0.0	0.0	0.0	0.0	0.0	11.2	**0.007**	Up	0.65	1.34	0.46	0.87
IL-10	0.0	0.0	0.0	0.0	2.4	18.7	**<0.001**	Up	**0.77**	**0.86**	**0.59**	**0.97**
IL-12	5.0	11.1	18.5	0.0	13.2	38.4	0.946	-	0.50	22.92	0.39	0.87
IL-13	0.0	0.2	2.7	0.0	0.0	0.0	**0.004**	Down	0.35	5.66	0.12	0.97
IL-15	0.0	0.0	0.0	0.0	0.0	0.0	0.091	-	0.54	0.40	0.09	1.00
IL-17	0.0	0.0	0.0	0.0	0.0	136.9	**<0.001**	Up	**0.72**	**32.27**	**0.44**	**1.00**
Eotaxin	74.1	109.0	166.8	0.0	106.8	187.6	0.394	-	0.55	0.00	0.38	1.00
FGF Basic	0.0	0.0	0.0	0.0	0.0	0.0	0.237	-	0.53	11.08	0.09	1.00
G-CSF	0.0	0.0	3.4	0.0	0.0	1.7	0.415	-	0.46	22.82	0.07	1.00
GM-CSF	0.0	0.0	0.0	0.0	0.0	1.8	0.328	-	0.55	0.53	0.30	0.80
IFN-γ	38.2	47.4	59.6	90.3	153.1	247.5	**<0.001**	Up	**0.91**	**65.44**	**0.89**	**0.87**
IP-10	320.6	448.1	700.6	880.2	1368.9	2298.6	**<0.001**	Up	**0.90**	**749.31**	**0.84**	**0.80**
MCP-1	18.4	31.6	47.1	8.5	40.3	64.7	0.799	-	0.52	45.81	0.40	0.77
MIP-1α	0.0	0.0	0.0	0.0	0.0	0.0	0.313	-	0.52	17.15	0.03	1.00
MIP-1β	59.8	83.4	127.9	83.0	111.4	161.3	**0.002**	Up	0.69	68.57	0.91	0.40
RANTES	6638.6	7709.4	9163.9	6059.4	7421.9	9591.6	0.451	-	0.55	6539.75	0.40	0.79
TNF-α	0.0	0.0	0.0	0.0	0.0	0.0	0.301	-	0.53	30.45	0.08	1.00
VEGF	16.8	52.5	108.1	35.2	167.2	307.0	**0.002**	Up	0.69	153.80	0.54	0.90
CA19-9	2.0	4.5	9.3	49.0	188.1	656.5	**<0.001**	Up	**0.95**	**16.40**	**0.92**	**0.97**

Note: Sensitivity (Sens.); Specificity (Spec.); Wilcoxon's Test p-value (Univariate, (P)); Inter-quartile range= 25th to 75th quartile; Receiver-Operator Characteristic Area Under Curve (AUC); Optimal Cut-off as determined by ROC (C/O); Regulation in PDAC (Reg.)

3.4.1.2 Diagnostic accuracy of individual CCGFs for PDAC against CP

Results from the basic analysis of the twenty-seven CCGFs and CA19-9 serum concentration data obtained using the 180 serum samples are summarised in ***Table 3.5***. Thirteen CCGFs and CA19-9 were observed to have a significantly different serum concentration between the PDAC group and the CP control subgroup, (Univariate by Wilcoxon's Test, $p<0.05$). In addition to CA19-9, seven CCGFs were identified as up regulated in PDAC compared to CP including IL-4, IL-8, IL-10, IL-17, IFN-γ, IP-10, and VEGF. Furthermore, six CCGFs were down regulated in PDAC compared to CP including IL-1β, IL-1Rα, IL-5, IL-13, G-CSF, and Eotaxin.

Subsequent analyses assessed the diagnostic accuracies of the twenty-seven CCGFs and CA19-9 for differentiating between PDAC and CP. Results from ROC analysis showed that four CCGFs and CA19-9 had a ROC-AUC of >0.7 and are therefore good discriminators between the two disease groups. In particular, IL-1β, IFN-γ, and CA19-9 achieved a relatively high accuracy for diagnosing PDAC with ROC-AUCs of 0.83, 0.83, and 0.84, respectively. Furthermore, we report that IL-1β and IFN-γ achieved similar sensitivity (0.77 and 0.76, respectively) and specificities (1.00 and 0.90) at optimal cut-off values of 0.00 pg/ml and 90.73 pg/ml compared to CA19-9 (sensitivity, 0.78; specificity; 0.83 at cut-off of 45.8 U/ml; ***Figure 3.4***).

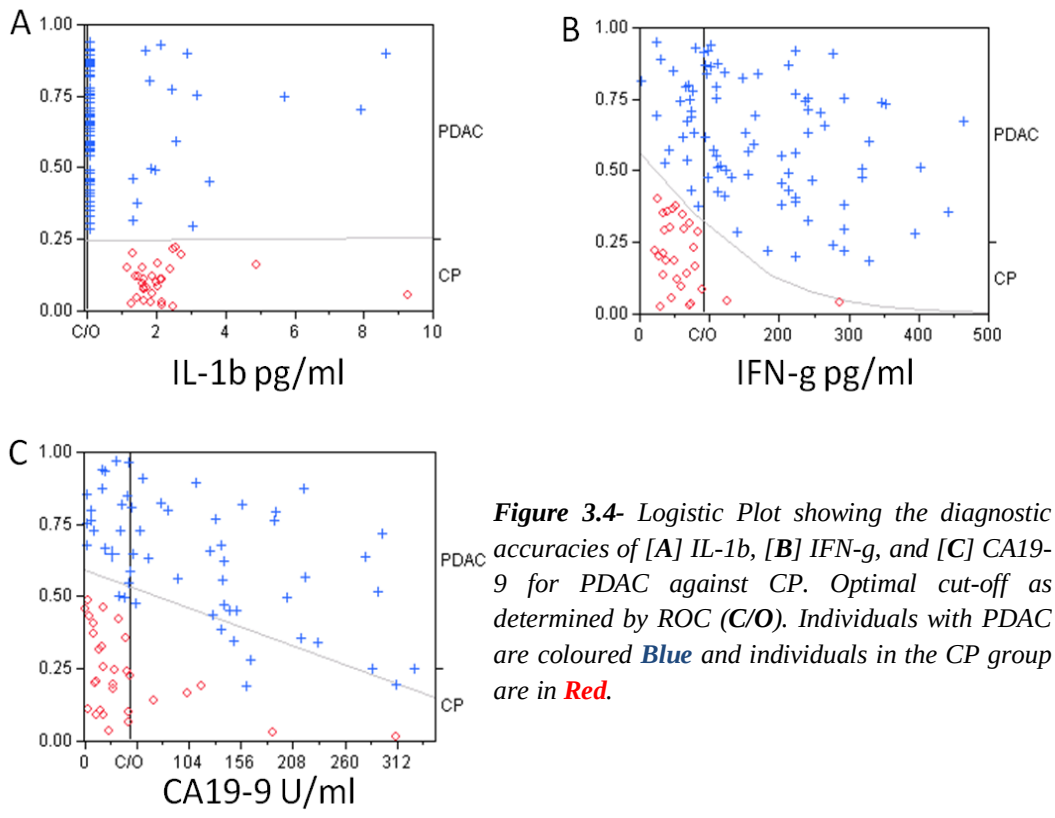


Figure 3.4- *Logistic Plot showing the diagnostic accuracies of [**A**] IL-1b, [**B**] IFN-g, and [**C**] CA19-9 for PDAC against CP. Optimal cut-off as determined by ROC (**C/O**). Individuals with PDAC are coloured **Blue** and individuals in the CP group are in **Red**.*

Table 3.5- *Diagnostic potential of CCGFs and CA19-9 for PDAC against CP*

Analyte	Quartile (pg/ml)						Univariate (p)	Reg.	Diagnostic accuracy			
	Chronic Pancreatitis			PDAC								
	25th	Median	75th	25th	Median	75th			AUC	C/O	Sens.	Spec.
PDGF	7996.0	13657.4	22684.7	12415.5	15557.1	20338.0	0.163	-	0.59	8989.14	0.93	0.33
IL-1β	1.6	1.9	2.2	0.0	0.0	0.0	**<0.001**	Down	**0.83**	**0.00**	**0.77**	**1.00**
IL-1Rα	87.3	111.7	204.6	32.1	70.8	160.3	**0.026**	Down	0.36	221.72	0.18	0.87
IL-2	0.0	0.0	0.0	0.0	0.0	0.0	0.313	-	0.52	2.78	0.03	1.00
IL-4	1.4	1.8	2.2	1.5	3.4	6.2	**0.003**	Up	0.68	2.40	0.62	0.87
IL-5	1.7	2.1	2.9	0.0	0.0	0.0	**<0.001**	Down	0.17	6.11	0.08	1.00
IL-6	5.1	8.5	17.8	2.6	9.5	21.0	0.879	-	0.51	7.72	0.71	0.47
IL-7	4.4	5.8	8.2	0.0	9.5	24.6	0.309	-	0.56	14.07	0.47	0.97
IL-8	12.3	15.6	21.1	10.7	23.6	40.3	**0.039**	Up	0.63	29.46	0.40	0.97
IL-9	0.0	0.0	18.9	0.0	0.0	11.2	0.717	-	0.49	16.86	0.80	0.27
IL-10	0.0	0.0	1.7	0.0	2.4	18.7	**0.004**	Up	0.67	1.93	0.54	0.87
IL-12	3.1	12.0	24.2	0.0	13.2	38.4	0.729	-	0.48	20.97	0.42	0.73
IL-13	0.2	1.9	3.9	0.0	0.0	0.0	**<0.001**	Down	0.26	9.59	0.09	1.00
IL-15	0.0	0.0	0.0	0.0	0.0	0.0	0.091	-	0.54	0.40	0.09	1.00
IL-17	0.0	0.0	0.0	0.0	0.0	136.9	**<0.001**	Up	**0.70**	**32.27**	**0.44**	**1.00**
Eotaxin	106.2	132.6	205.7	0.0	106.8	187.6	**0.045**	Down	0.62	0.00	0.38	0.97
FGF Basic	0.0	0.0	0.0	0.0	0.0	0.0	0.637	-	0.48	123.88	0.03	1.00
G-CSF	0.0	1.0	10.1	0.0	0.0	1.7	**0.010**	Down	0.65	7.28	0.88	0.43
GM-CSF	0.0	0.0	2.9	0.0	0.0	1.8	0.994	-	0.50	21.32	0.08	0.97
IFN-γ	35.8	52.3	73.9	90.3	153.1	247.5	**<0.001**	Up	**0.83**	**90.73**	**0.76**	**0.90**
IP-10	401.5	588.0	1053.8	880.2	1368.9	2298.6	**<0.001**	Up	**0.81**	**797.21**	**0.83**	**0.67**
MCP-1	27.5	50.3	101.2	8.5	40.3	64.7	0.061	-	0.61	9.99	0.27	1.00
MIP-1α	0.0	0.0	0.0	0.0	0.0	0.0	0.977	-	0.50	17.15	0.03	1.00
MIP-1β	65.0	136.0	188.2	83.0	111.4	161.3	0.585	-	0.53	161.65	0.77	0.47
RANTES	6159.9	8188.9	8883.9	6059.4	7421.9	9591.6	0.752	-	0.54	10996.50	0.18	0.90
TNF- α	0.0	0.0	0.0	0.0	0.0	0.0	0.530	-	0.48	326.57	0.04	1.00
VEGF	17.3	58.2	145.8	35.2	167.2	307.0	**0.007**	Up	0.66	164.89	0.51	0.83
CA19-9	10.3	21.5	42.5	54.8	225.0	721.5	**<0.001**	Up	**0.84**	**45.80**	**0.78**	**0.83**

Note*: Sensitivity (**Sens.**); Specificity (**Spec.**); Wilcoxon's Test p-value (**Univariate, (P)**); Inter-quartile range= 25th to 75th quartile; Receiver-Operator Characteristic Area Under Curve (**AUC**); Optimal Cut-off as determined by ROC (**C/O**); Regulation in PDAC (**Reg**.)*

3.4.1.3 Diagnostic accuracy of individual CCGFs for PDAC against biliary obstruction (DC)

Results from the basic analysis of the twenty-seven CCGFs and CA19-9 serum concentration data obtained using the 180 serum samples are summarised in Table 3.6. Thirteen CCGFs and CA19-9 were observed to have a significantly different serum concentration between the PDAC group and the DC (biliary obstruction) control subgroup, (Univariate by Wilcoxon's Test, $p<0.05$). In addition to CA19-9, seven CCGFs were identified as up regulated in PDAC compared to DC including PDGF, IL-4, IL-10, IL-17, IFN-γ, IP-10, and VEGF. Furthermore, six CCGFs were down regulated in PDAC compared to DC including IL-1β, IL-1Rα, IL-5, IL-13, G-CSF, MCP-1, and IL-13.

Subsequent analyses assessed the diagnostic accuracies of the twenty-seven CCGFs and CA19-9 for differentiating between PDAC and DC. In additional to CA19-9m results from ROC analysis showed that four CCGFs had a ROC-AUC of >0.7 including IL-17, PDGF, IFN-γ, and IP-10, which indicated that they are good discriminators of PDAC against biliary obstruction. In particular, IFN-γ achieved a much higher accuracy for diagnosing PDAC against DC (ROC-AUC, 0.86 and 0.75, respectively). Furthermore, we report that IFN-γ is the most accurate diagnostic marker for PDAC against DC with a sensitivity of 0.74 and a specificity of 0.87 at a cut-off of 92.51 pg/ml whereas CA19-9 only achieved a sensitivity of 0.87 and a specificity of 0.57 at an optimal cut-off of 33U/ml (***See Figure 3.5***).

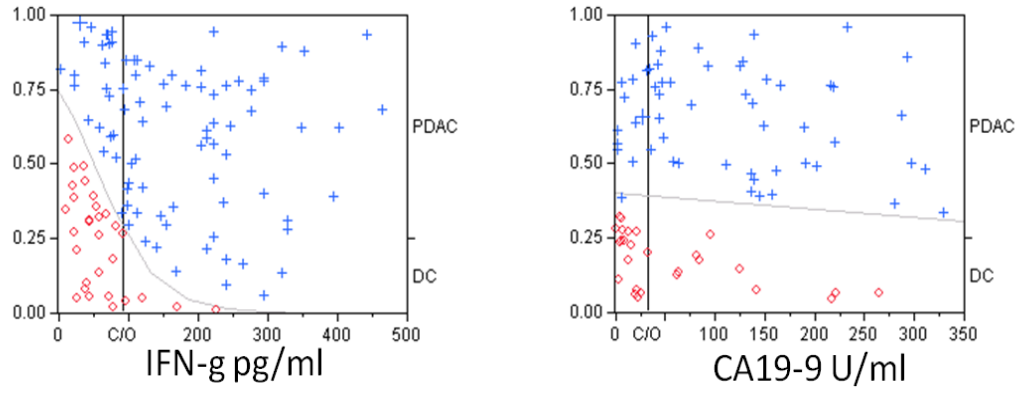


Figure 3.5- *Logistic Plot showing the diagnostic accuracies of IFN-g (left) and CA19-9 (right) for PDAC against DC. Optimal cut-off as determined by ROC (C/O). Individuals with PDAC are coloured* ***Blue*** *and individuals in the HC group are in* ***Red****.*

***Table 3.6**- Diagnostic potential of CCGFs and CA19-9 for PDAC against DC*

Analyte	Quartile (pg/ml)						Univariate (p)	Reg.	Diagnostic accuracy			
	Disease Control			PDAC								
	25th	Median	75th	25th	Median	75th			AUC	C/O	Sens.	Spec.
PDGF	2397.6	7893.3	16658.8	12415.5	15557.1	20338.0	**<0.001**	Up	**0.76**	**9198.56**	**0.91**	**0.60**
IL-1β	1.2	1.8	2.2	0.0	0.0	0.0	**<0.001**	Down	0.18	5.67	0.08	1.00
IL-1Rα	80.7	147.5	242.4	32.1	70.8	160.3	**0.007**	Down	0.66	104.02	0.63	0.70
IL-2	0.0	0.0	0.0	0.0	0.0	0.0	0.992	-	0.50	10.55	0.99	0.03
IL-4	0.5	1.5	2.1	1.5	3.4	6.2	**<0.001**	Up	0.70	2.40	0.62	0.90
IL-5	0.7	1.8	2.6	0.0	0.0	0.0	**<0.001**	Down	0.19	4.97	0.10	1.00
IL-6	6.8	11.4	17.2	2.6	9.5	21.0	0.543	-	0.46	56.59	0.13	0.97
IL-7	1.1	5.7	9.8	0.0	9.5	24.6	0.228	-	0.57	13.56	0.48	0.90
IL-8	13.0	19.4	29.2	10.7	23.6	40.3	0.606	-	0.53	24.64	0.48	0.70
IL-9	0.0	11.6	25.4	0.0	0.0	11.2	0.051	-	0.62	11.10	0.76	0.53
IL-10	0.0	0.3	1.9	0.0	2.4	18.7	**0.008**	Up	0.66	6.22	0.41	0.97
IL-12	3.4	13.4	27.2	0.0	13.2	38.4	0.975	-	0.50	37.86	0.27	0.97
IL-13	0.2	2.5	4.0	0.0	0.0	0.0	**<0.001**	Down	0.26	15.83	0.07	1.00
IL-15	0.0	0.0	0.0	0.0	0.0	0.0	0.091	-	0.54	0.40	0.09	1.00
IL-17	0.0	0.0	0.0	0.0	0.0	136.9	**<0.001**	Up	**0.72**	**4.59**	**0.47**	**0.97**
Eotaxin	64.1	132.3	158.6	0.0	106.8	187.6	0.350	-	0.45	169.60	0.30	0.83
FGF Basic	0.0	0.0	0.0	0.0	0.0	0.0	0.237	-	0.53	11.08	0.09	1.00
G-CSF	0.0	3.3	9.1	0.0	0.0	1.7	**0.006**	Down	0.64	2.42	0.80	0.50
GM-CSF	0.0	0.0	7.1	0.0	0.0	1.8	0.111	-	0.42	116.61	0.03	1.00
IFN-γ	26.9	47.6	79.4	90.3	153.1	247.5	**<0.001**	Up	**0.86**	**92.51**	**0.74**	**0.87**
IP-10	461.1	571.1	951.1	880.2	1368.9	2298.6	**<0.001**	Up	**0.79**	**733.62**	**0.86**	**0.63**
MCP-1	28.7	56.0	99.7	8.5	40.3	64.7	**0.014**	Down	0.65	12.80	0.28	1.00
MIP-1α	0.0	0.0	0.0	0.0	0.0	0.0	0.313	-	0.52	17.15	0.03	1.00
MIP-1β	73.6	107.4	158.9	83.0	111.4	161.3	0.544	-	0.54	80.88	0.81	0.37
RANTES	4648.6	6196.6	8911.7	6059.4	7421.9	9591.6	0.063	-	0.39	17980.18	0.96	0.13
TNF-α	0.0	0.0	0.0	0.0	0.0	0.0	0.093	-	0.54	2.68	0.09	1.00
VEGF	14.0	49.3	111.6	35.2	167.2	307.0	**<0.001**	Up	0.70	197.16	0.48	0.93
CA19-9	8.5	24.5	128.6	49.0	188.1	656.5	**<0.001**	Up	**0.75**	**33.00**	**0.87**	**0.57**

***Note**: Sensitivity (**Sens.**); Specificity (**Spec.**); Wilcoxon's Test p-value (**Univariate, (P)**); Inter-quartile range= 25th to 75th quartile; Receiver-Operator Characteristic Area Under Curve (**AUC**); Optimal Cut-off as determined by ROC (**C/O**); Regulation in PDAC (**Reg**.)*

3.4.1.4 Diagnostic accuracy of individual CCGFs for PDAC against all Controls

Results from the basic analysis of the twenty-seven CCGFs and CA19-9 serum concentration data obtained using the 180 serum samples are summarised in Table 3.7. Fourteen CCGFs and CA19-9 were observed to have a significantly different serum concentration between the PDAC group and the overall control group, (Univariate by Wilcoxon's Test, $p<0.05$). In addition to CA19-9, nine CCGFs were identified as up regulated in PDAC compared to CP including PDGF, IL-4, IL-8, IL-10, IL-17, IL-15, IFN-γ, IP-10, and VEGF. Furthermore, four CCGFs were down regulated in PDAC compared to Controls including IL-1β, IL-5, IL-13, and G-CSF.

Subsequent analyses assessed the diagnostic accuracies of the twenty-seven CCGFs and CA19-9 for differentiating between PDAC and Controls. Results from ROC analysis showed that four CCGFs and CA19-9 had a ROC-AUC of >0.7 and are therefore good discriminators between the two disease groups. In particular, IFN-γ, and CA19-9 achieved a relatively high accuracy for diagnosing PDAC with ROC-AUCs of 0.87 and 0.85, respectively. Furthermore, we report that IFN-γ and CA19-9 have comparable diagnostic accuracy at their optimal cut-off values for PDAC against Control subjects with sensitivities of 0.77 and 0.87 and specificities of 0.89 and 0.73, respectively (***Figure 3.6***).

A multivariate analysis was performed on CCGFs, which were significant on univariate analysis to identify independent discriminators of PDAC amongst control subjects. Results indicate that IL-4, IL-17, G-CSF, and IL-10 were independently significant as diagnostic markers of PDAC (***Table 3.7***).

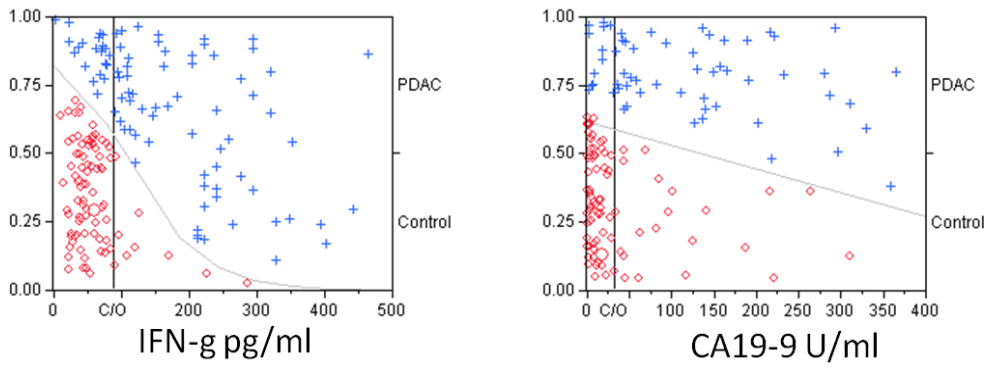


Figure 3.6- *Logistic Plot showing the diagnostic accuracies of IFN-g (left) and CA19-9 (right) for PDAC against all controls. Optimal cut-off as determined by ROC (C/O). Individuals with PDAC are coloured* ***Blue*** *and individuals in the HC group are in* ***Red****.*

Table 3.7- *Diagnostic potential of CCGFs and CA19-9 for PDAC against all Controls*

Analyte	Quartile (pg/ml)						Uni (p)	Multi (p)	Diagnostic accuracy			
	Disease Control			PDAC								
	25th	Median	75th	25th	Median	75th			Reg.	AUC	C/O	Sens.	Spec.
PDGF	6767.2	11958.9	17565.9	12415.5	15557.1	20338.0	**<0.001**	0.294	Up	0.67	9198.56	0.91	0.40
IL-1β	1.3	1.7	2.2	0.0	0.0	0.0	**<0.001**	0.902	Down	0.18	5.67	0.08	0.98
IL-1Rα	60.7	106.0	190.6	32.1	70.8	160.3	**0.034**	0.862	Down	0.41	316.37	0.12	0.91
IL-2	0.0	0.0	0.0	0.0	0.0	0.0	0.316	-	-	0.51	2.78	0.03	0.99
IL-4	1.2	1.7	2.0	1.5	3.4	6.2	**<0.001**	**<0.001**	Up	**0.73**	**2.40**	**0.62**	**0.91**
IL-5	1.3	1.8	2.6	0.0	0.0	0.0	**<0.001**	0.338	Down	0.18	4.69	0.12	0.96
IL-6	3.8	6.3	13.6	2.6	9.5	21.0	0.143	-	-	0.56	7.72	0.71	0.58
IL-7	3.3	5.2	8.0	0.0	9.5	24.6	0.068	-	-	0.58	13.56	0.48	0.93
IL-8	8.8	13.6	20.2	10.7	23.6	40.3	**0.001**	0.756	Up	0.64	23.30	0.52	0.82
IL-9	0.0	0.0	17.0	0.0	0.0	11.2	0.639	-	-	0.52	2.47	0.44	0.62
IL-10	0.0	0.0	0.7	0.0	2.4	18.7	**<0.001**	**0.033**	Up	0.70	1.93	0.54	0.87
IL-12	3.5	11.7	21.8	0.0	13.2	38.4	0.885	-	-	0.49	20.71	0.43	0.74
IL-13	0.0	1.2	3.8	0.0	0.0	0.0	**<0.001**	0.481	Down	0.29	5.66	0.12	0.94
IL-15	0.0	0.0	0.0	0.0	0.0	0.0	**0.004**	1.00	Up	0.54	0.40	0.09	1.00
IL-17	0.0	0.0	0.0	0.0	0.0	136.9	**<0.001**	**<0.001**	Up	**0.71**	**36.76**	**0.43**	**1.00**
Eotaxin	82.2	123.2	180.0	0.0	106.8	187.6	0.076	-	-	0.58	0.00	0.38	0.99
FGF Basic	0.0	0.0	0.0	0.0	0.0	0.0	0.387	-	-	0.52	11.08	0.09	0.97
G-CSF	0.0	0.9	8.5	0.0	0.0	1.7	**0.004**	**<0.001**	Down	0.38	103.72	0.03	1.00
GM-CSF	0.0	0.0	3.0	0.0	0.0	1.8	0.736	-	-	0.49	99.65	0.04	0.99
IFN-γ	36.8	48.1	71.4	90.3	153.1	247.5	**<0.001**	0.605	Up	**0.87**	**88.95**	**0.77**	**0.89**
IP-10	399.8	533.4	890.8	880.2	1368.9	2298.6	**<0.001**	0.059	Up	**0.83**	**797.21**	**0.83**	**0.70**
MCP-1	24.5	44.5	93.6	8.5	40.3	64.7	0.054	-	-	0.58	9.99	0.27	0.99
MIP-1α	0.0	0.0	0.0	0.0	0.0	0.0	0.305	-	-	0.51	17.15	0.03	1.00
MIP-1β	68.4	104.1	164.4	83.0	111.4	161.3	0.128	-	-	0.57	81.26	0.80	0.39
RANTES	5911.6	7498.2	9013.1	6059.4	7421.9	9591.6	0.688	-	-	0.48	7272.12	0.49	0.55
TNF-α	0.0	0.0	0.0	0.0	0.0	0.0	0.359	-	-	0.52	30.45	0.08	0.98
VEGF	17.3	51.0	113.3	35.2	167.2	307.0	**<0.001**	0.120	Up	0.68	156.68	0.53	0.86
CA19-9	4.8	13.4	41.5	49.0	188.1	656.5	**<0.001**	-	Up	**0.85**	**33.00**	**0.87**	**0.73**

Note*: Sensitivity (**Sens.**); Specificity (**Spec.**); Univariate analysis by Wilcoxon's Test p-value (**Uni (P)**);Multivariate by Logistic Regression p-value (**Multi (P)**) Inter-quartile range= 25th to 75th quartile; Receiver-Operator Characteristic Area Under Curve (**AUC**); Optimal Cut-off as determined by ROC (**C/O**); Regulation in PDAC (**Reg**.)*

3.4.1.5 Correlation studies to determine the relationship between CCGFs

The fourteen CCGFs, which were significant on univariate analysis, were analysed for non-parametric correlation. The correlation between CCGFs were analysed independently for the Control group (HC, CP, and DC) and the PDAC group.

Five pairs of CCGFs were found to have a significantly positive correlation (Kendall τ Test, $\tau>0.5$ and $p<0.05$) in both the control group and the PDAC group. In particular, results from ***Table 3.8*** and ***Table 3.9*** suggest that regardless of the disease status, there seemed to be a positive correlation between the serum concentrations of IL-1β and IL-5 as well as between IL-4 and IFN-γ. It is also interesting to note that in general, the strength of the correlations appeared to be stronger in PDAC patients (Kendall Tau coefficient >0.7) compared to controls (Kendall Tau coefficient <0.65).

In addition to non-parametric correlation studies, the relationships between individual CCGFs were analysed by hierarchy clustering. Results from this analysis confirmed the findings from the non-parametric correlation studies. Indeed, the heat map in ***Figure 3.7*** identified separate three groups of CCGFs with similar patterns of serum expression (especially in the PDAC patients): (a) IL-10 and IL-17; (b) IL-4 and IFN-γ; (c) IL-1β, IL-5, and IL-13.

Table 3.8*- The correlation between individual CCGFs in the Control Group*

CCGF 1	CCGF 2	Kendall τ	Prob>|τ|	Plot
IL-5	IL-1β	0.6323	<0.0001	++++++++
IFN-γ	IL-4	0.6089	<0.0001	++++++++
IL-4	IL-1β	0.5386	<0.0001	+++++++
IL-4	PDGF	0.5133	<0.0001	++++++
G-CSF	IL-1β	0.5049	<0.0001	++++++

Note: *Kendall Tau (t) coefficient ranges from -1 to 1 where -1 and 1 represents a perfect negative (-) and positive (+) correlation , respectively and a Tau coefficient of 0 represents no correlation. + under Plot represents positive correlation*

Table 3.9*- The correlation between individual CCGFs in the PDAC group*

CCGF 1	CCGF 2	Kendall τ	Prob>|τ|	Plot
IL-13	IL-5	0.973	<0.0001	++++++++++++
IL-13	IL-1β	0.968	<0.0001	++++++++++++
IL-5	IL-1β	0.968	<0.0001	++++++++++++
IFN-γ	IL-4	0.721	<0.0001	+++++++++
IL-17	IL-10	0.613	<0.0001	++++++++

Note: *Kendall Tau (t) coefficient ranges from -1 to 1 where -1 and 1 represents a perfect negative (-) and positive (+) correlation , respectively and a Tau coefficient of 0 represents no correlation. + under Plot represents positive correlation*

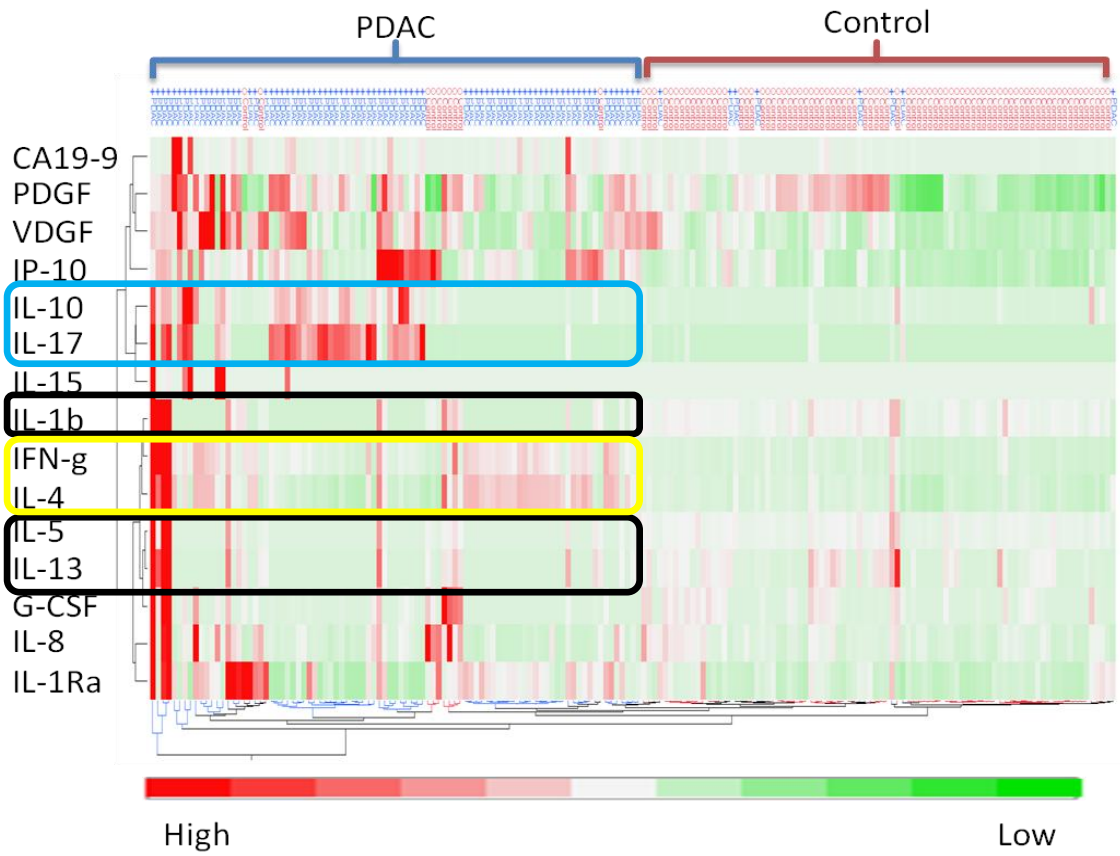


Figure 3.7- *Two-Way Hierarchical clustering of the fourteen significant CCGFs and CA19-9. Individuals with resectable pancreatic cancer (PDAC) are displayed in* ***blue*** *and control patients in* ***red****. It can be seen from the heat map that the serum concentrations of IL-10 and IL-17 were similar, especially in individuals with PDAC. Similarly, IL-4 and IFN-g appeared to be increased in the same individuals with PDAC. In addition, it is also worth noting that the patterns of IL-1b, IL-5, and IL-13 expression in the serum were almost identical to each other in individuals with PDAC.*

In summary, results from initial analysis indicated that individual CCGFs were insufficiently accurate as standalone markers for resectable pancreatic cancer. However, IFN-γ has been shown to have a similar accuracy compared to the standard biomarker, CA19-9. Nevertheless, the aim of the *Discovery Phase* will be to determine the value of combining a panel of CCGF to improve the diagnostic accuracy of CCGFs as markers for pancreatic cancer.

3.4.2 Discovery Phase- Diagnostic potential of CCGFs in combination

One hundred and twenty randomly selected samples were analysed in the *Discovery Phase* to identify candidate biomarkers, which can be combined as a single marker for the diagnosis of PDAC.

3.4.2.1 Selection of CCGFs for combination

The fourteen CCGFs, which were significant in ***section 3.4.1.4*** were shortlisted for further selection by stepwise regression (***see section 3.3.5.4***). The selection process was completed in four steps, after which no further CCGF satisfied the criteria for entering or leaving the model (***Table 3.10***). Amongst the fourteen CCGF markers, four were identified as independently significant combined markers for discriminating PDAC against control subjects: IL-4, IL-17, G-CSF, and IP-10 (***Table 3.10***). In view of the results from the previous section, which suggested that some CCGFs might share a common serum expression pattern, the stepwise regression model was repeated four times, each with one marker (IL-4, IL-17, G-CSF, or IP-10) excluded from the analysis. However, none of the alternative models was better than the original (***Table 3.11***).

Cluster analysis was performed on the four CCGFs selected by the stepwise regression model (***Candidate CCGFs)*** and a heat map was created to illustrate graphically the regulation of CCGFs in the serum of cancer and control subjects (***Figure 3.8***). It appeared from the heat map pattern that PDAC is associated with the up-regulation of IL-4, IL-17, or IP-10 whereas the up-regulation of G-CSF appeared to identify a group of control patients, some of whom have a high serum IL-4 concentration.

Table 3.10- CCGF markers selected by the stepwise regression model for combination

Step	Parameter	Action	Wald/Score p-value	R^2
1	**Il-17**	Entered	< 0.0001	0.23
2	**IL-4**	Entered	< 0.0001	0.51
3	**G-CSF**	Entered	< 0.0001	0.78
4	**IP-10**	Entered	0.0003	0.86

Note: *The R^2 value reflects how closely the data fits with the model*

Table 3.11 *-Alternative markers for combination*

Step	Parameter	Action	Wald/Score p-value	R^2
IL-17 excluded				
1	**IP-10**	Entered	<0.001	0.18
2	**PDGF**	Entered	<0.001	0.28
3	**VEGF**	Entered	0.012	0.32
4	**IL-4**	Entered	<0.001	0.46
5	**PDGF**	Removed	0.066	0.42
6	**IL-1β**	Entered	<0.001	0.49
7	**G-CSF**	Entered	<0.001	0.69
8	**IFN-γ**	Entered	<0.001	0.76
9	**VEGF**	Removed	0.072	0.74
10	**IL-5**	Entered	<0.001	0.83
11	**IL-4**	Removed	0.184	0.82
IL-4 removed				
1	**IL-17**	Entered	<0.001	0.23
2	**IFN-γ**	Entered	<0.001	0.45
3	**G-CSF**	Entered	<0.001	0.70
4	**IL-17**	Removed	<0.001	0.53
G-CSF removed				
1	**IL-17**	Entered	<0.001	0.23
2	**IL-4**	Entered	<0.001	0.51
3	**IL-1β**	Entered	0.018	0.55
4	**IP-10**	Entered	<0.001	0.62
5	**IFN-γ**	Entered	0.017	0.66
6	**IL-17**	Removed	<0.001	0.57
IP-10 removed				
1	**IL-17**	Entered	<0.001	0.23
2	**IL-4**	Entered	<0.001	0.51
3	**G-CSF**	Entered	<0.001	0.78
4	**IL-8**	Entered	0.004	0.83
5	**IL-5**	Entered	0.011	0.87
6	**IL-1β**	Entered	0.139	0.88

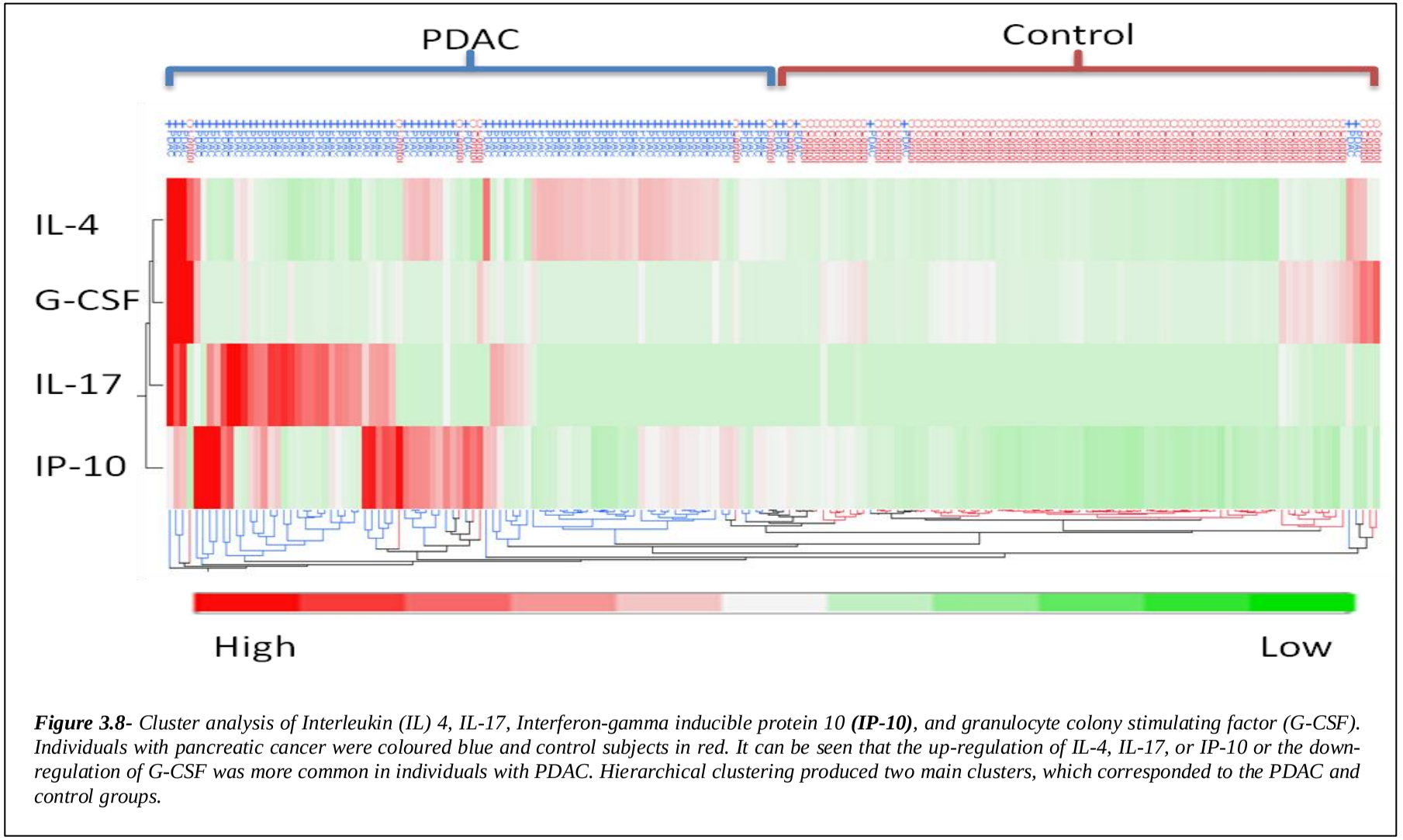


Figure 3.8- *Cluster analysis of Interleukin (IL) 4, IL-17, Interferon-gamma inducible protein 10* ***(IP-10)****, and granulocyte colony stimulating factor (G-CSF). Individuals with pancreatic cancer were coloured blue and control subjects in red. It can be seen that the up-regulation of IL-4, IL-17, or IP-10 or the down-regulation of G-CSF was more common in individuals with PDAC. Hierarchical clustering produced two main clusters, which corresponded to the PDAC and control groups.*

3.4.2.2 Diagnostic accuracy of the combined CCGF marker

Subsequent analyses combined the four candidate markers using two independent modelling methods: Multinomial logistic regression (M-LR) and Artificial Neural Network (NN). Each model generated a mathematical algorithm, which predicted the probability of PDAC based on the serum concentrations of IL-4, IL-17, G-CSF, and IP-10. The resulting probability values range from zero (likely control) to one (likely PDAC). The diagnostic accuracy of each model are summarised in ***Table 3.12*** and graphically represented in ***Figure 3.9*** and ***Figure 3.10***.

Results from the analysis of the diagnostic accuracy of the M-LR model were promising. In the *Discovery Phase*, the M-LR model achieved a very high diagnostic accuracy with ROC-AUCs of >0.99 for discriminating PDAC from the CP, DC, and HC control subgroups individually and as a combined Control group (***Figure 3.9***). Two PDAC patients were misdiagnosed as Control at the optimal cut-off value of 0.4 (as determined by ROC) and one Control subject (from the CP subgroup) was misclassified as PDAC (***Table 3.12***).

Results from the analysis of the diagnostic accuracy of the NN model were equally encouraging. In the *Discovery Phase*, the NN model again achieved a very high diagnostic accuracy with ROC-AUCs of 0.99 for discriminating PDAC from CP, DC, and HC control subgroups individually and as a combined Control group (***Figure 3.10***). Only one Control subject (CP subject) was misdiagnosed as PDAC at the optimal cut-off value of 0.365 (as determined by ROC) and two individuals with PDAC were misclassified as Control (***Table 3.12***).

Table 3.12- *Diagnostic accuracy of the combined CCGF marker in the Discovery Phase*

Parameter	Multinomial Logistic Regression				Artificial Neural Network			
	C/O	AUC	Sens.	Spec.	C/O	AUC	Sens.	Spec.
PDAC Vs CP	0.4	0.99	0.97	0.95	0.365	0.99	0.97	0.95
PDAC Vs DC	0.4	1.00	0.97	1.00	0.365	1.00	0.97	1.00
PDAC Vs HC	0.4	1.00	0.97	1.00	0.365	1.00	0.97	1.00
PDAC Vs Controls	0.4	0.99	0.97	0.98	0.365	1.00	0.97	0.98

Abbreviations*: Cut-off value (**C/O**) for the relative Sensitivity (**Sens.**)and specificity (**Spec.**); Receiver Operatory Characteristics Area Under Curve (**AUC**); resectable pancreatic cancer (**PDAC**); chronic pancreatitis (**CP**); disease controls (**DC**, mainly obstructive jaundice); Healthy volunteers (**HC**)*

Note*: The same algorithms (**M-LR** or **NN**) were used for each parameter.*

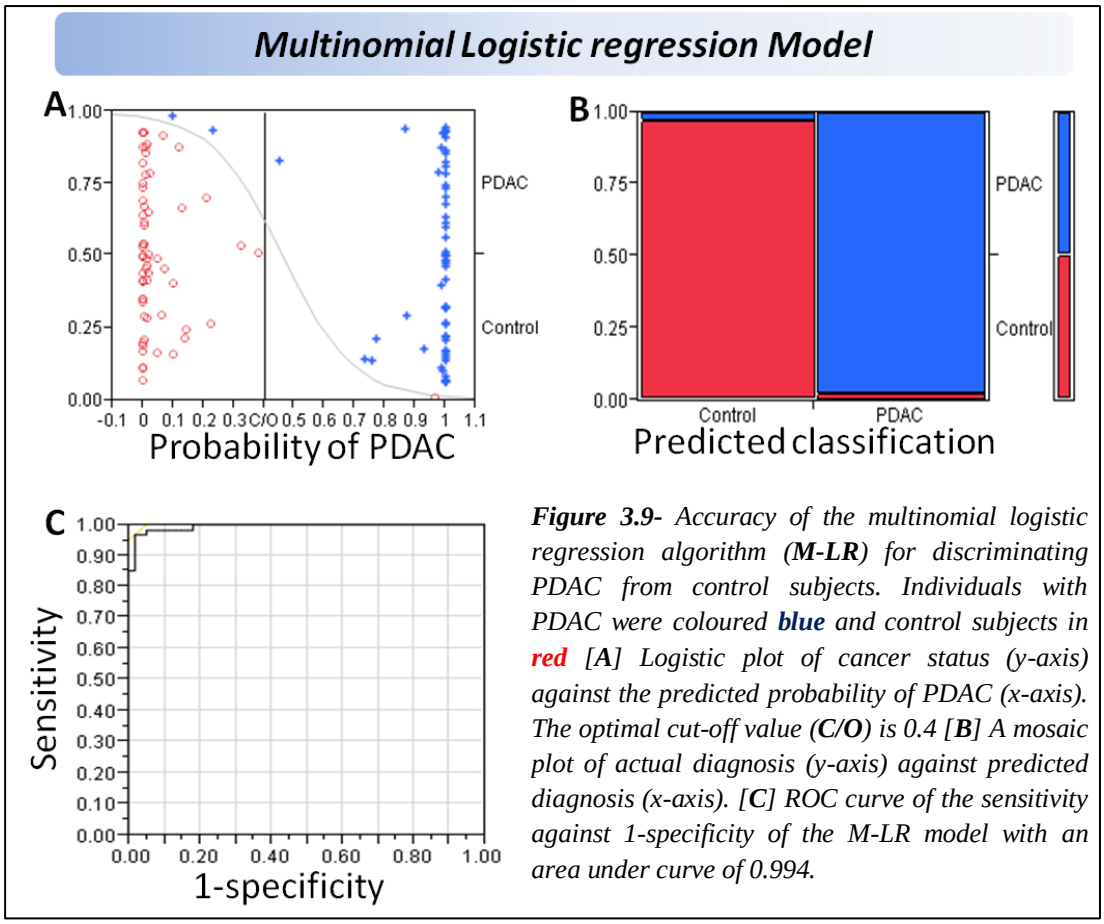


Figure 3.9- *Accuracy of the multinomial logistic regression algorithm (**M-LR**) for discriminating PDAC from control subjects. Individuals with PDAC were coloured **blue** and control subjects in **red** [**A**] Logistic plot of cancer status (y-axis) against the predicted probability of PDAC (x-axis). The optimal cut-off value (**C/O**) is 0.4 [**B**] A mosaic plot of actual diagnosis (y-axis) against predicted diagnosis (x-axis). [**C**] ROC curve of the sensitivity against 1-specificity of the M-LR model with an area under curve of 0.994.*

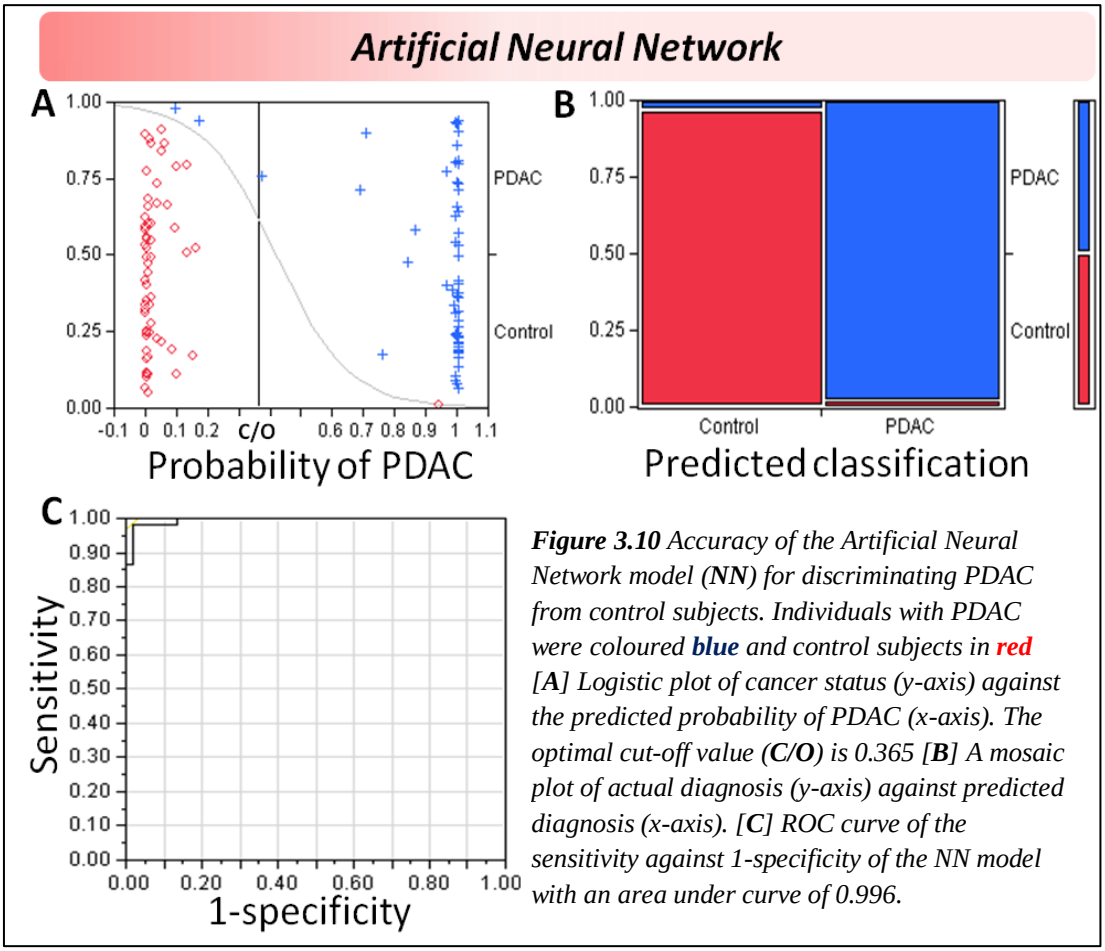


Figure 3.10 *Accuracy of the Artificial Neural Network model (**NN**) for discriminating PDAC from control subjects. Individuals with PDAC were coloured **blue** and control subjects in **red** [**A**] Logistic plot of cancer status (y-axis) against the predicted probability of PDAC (x-axis). The optimal cut-off value (**C/O**) is 0.365 [**B**] A mosaic plot of actual diagnosis (y-axis) against predicted diagnosis (x-axis). [**C**] ROC curve of the sensitivity against 1-specificity of the NN model with an area under curve of 0.996.*

3.4.3 Validation Phase- Validation of the disease-predicting algorithms

3.4.3.1 The diagnostic accuracy of the prediction models in the Validation Phase

Both algorithms generated in the *Discovery Phase* by the M-LR and the NN models were directly applied to the CCGF data from the sixty samples in the validation sample set. The resulting probability values were classified into PDAC if the probability was >0.4 and Control if the probability was ≤0.4. The results are summarized in ***Table 3.13***, **Error! Reference source not found.** and ***Figure 3.11***.

Although the diagnostic accuracy of the M-LR combined CCGF algorithm was slightly reduced in the *Validation Phase* compared to the *Discovery Phase,* the M-LR model still performed very well with ROC-AUCs of >95% for discriminating PDAC from CP, DC, and HC control subgroups individually and as a combined Control group. Furthermore, we report that the M-LR model only misclassified two individuals in the PDAC group and no Control subjects were misclassified.

Results from the validation of the NN combined CCGF algorithm were remarkable with ROC-AUCs of 1.00 discriminating PDAC from the CP, DC, and HC control subgroups as individually and as a combined Control group. Furthermore, ***all individuals in the validation sample set were correctly classified*** into their disease group at the *Discovery Phase* defined cut-off value of 0.4.

Table 3.13- *Diagnostic accuracy of the combined CCGF marker in the Validation Phase*

Parameter	Multinomial Logistic Regression				Artificial Neural Network			
	C/O	AUC	Sens.	Spec.	C/O	AUC	Sens.	Spec.
PDAC Vs CP	0.4	0.97	0.93	1.00	0.365	1.00	1.00	1.00
PDAC Vs DC	0.4	0.95	0.93	1.00	0.365	1.00	1.00	1.00
PDAC Vs HC	0.4	0.95	0.93	1.00	0.365	1.00	1.00	1.00
PDAC Vs Controls	0.4	0.96	0.93	1.00	0.365	1.00	1.00	1.00

Abbreviations*: Cut-off value (**C/O**) for the relative Sensitivity (**Sens.**)and specificity (**Spec.**); Receiver Operatory Characteristics Area Under Curve (**AUC**); resectable pancreatic cancer (**PDAC**); chronic pancreatitis (**CP**); disease controls (**DC**, mainly obstructive jaundice); Healthy volunteers (**HC**)*

Note*: The same algorithms (**M-LR** or **NN**) were used for each parameter.*

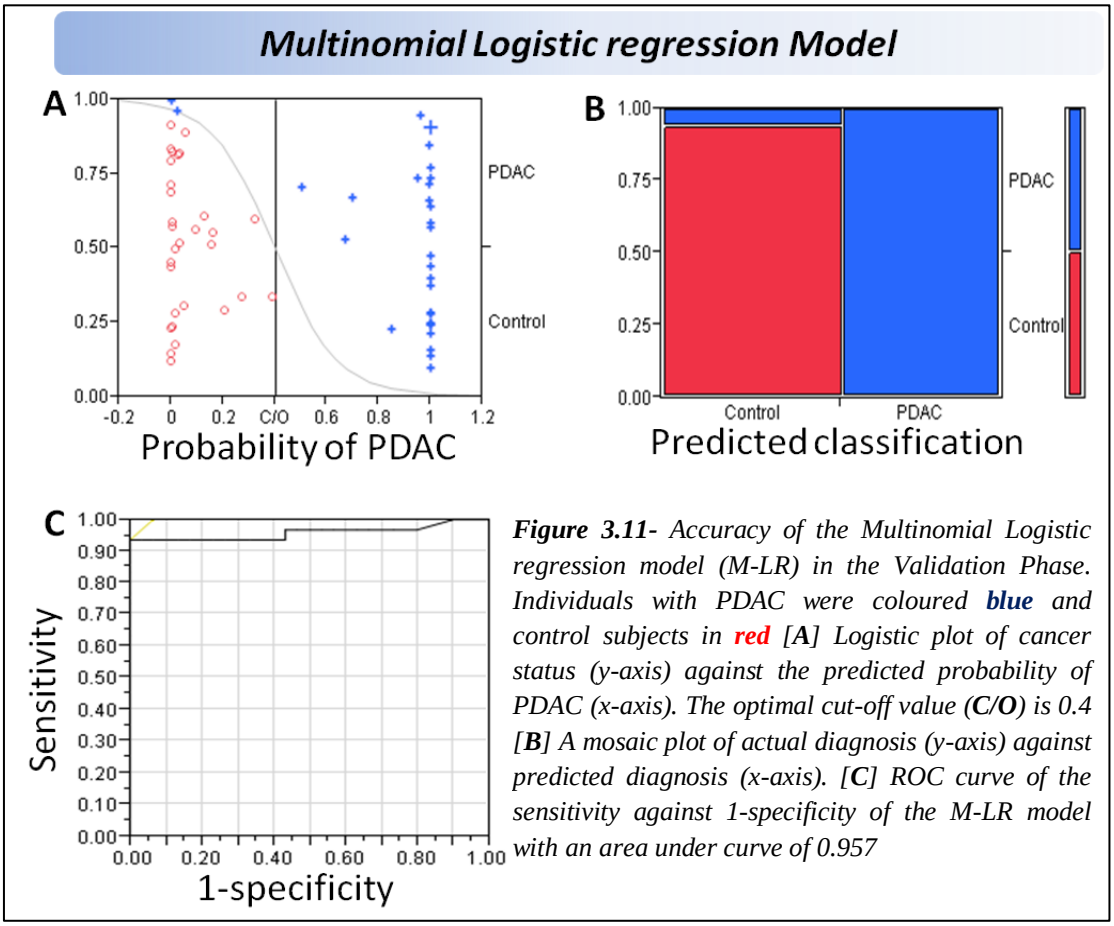


Figure 3.11- *Accuracy of the Multinomial Logistic regression model (M-LR) in the Validation Phase. Individuals with PDAC were coloured* ***blue*** *and control subjects in* ***red*** *[**A**] Logistic plot of cancer status (y-axis) against the predicted probability of PDAC (x-axis). The optimal cut-off value (**C/O**) is 0.4 [**B**] A mosaic plot of actual diagnosis (y-axis) against predicted diagnosis (x-axis). [**C**] ROC curve of the sensitivity against 1-specificity of the M-LR model with an area under curve of 0.957*

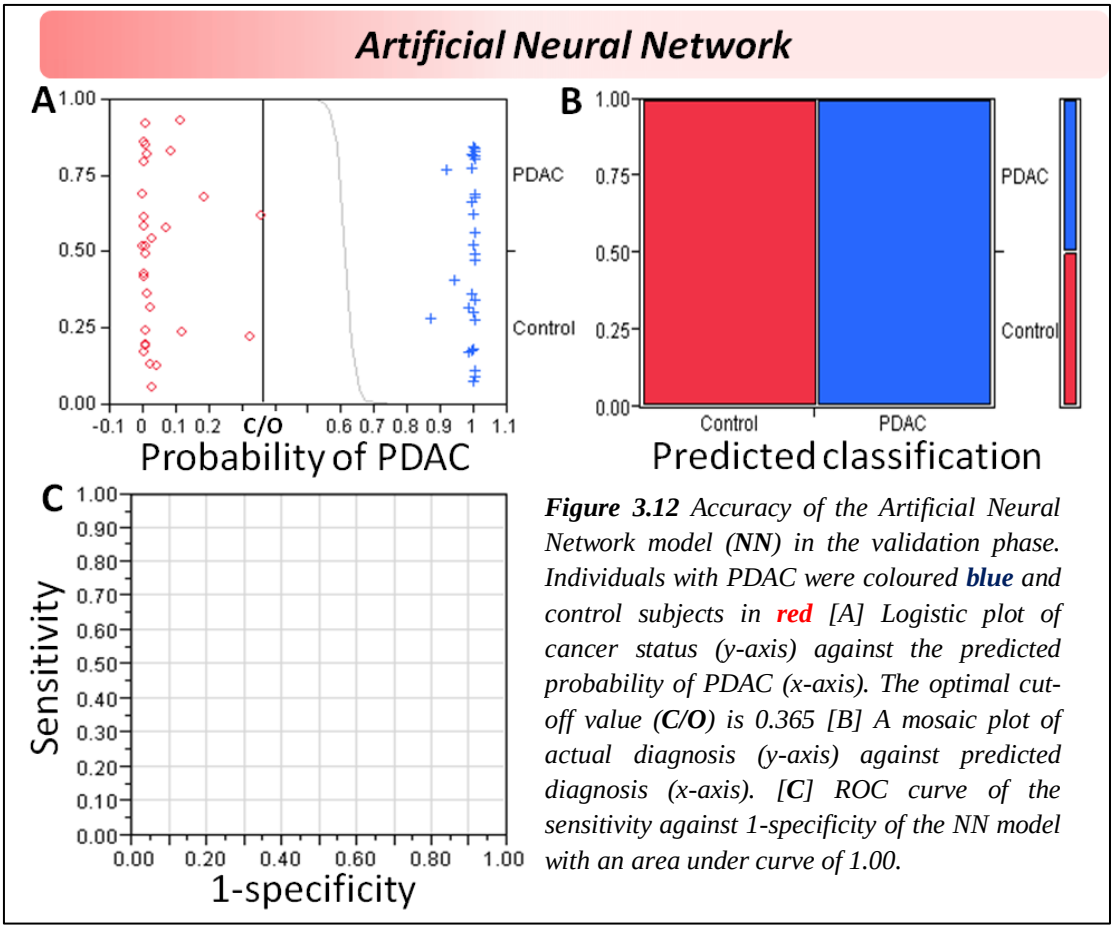


Figure 3.12 *Accuracy of the Artificial Neural Network model (**NN**) in the validation phase. Individuals with PDAC were coloured* ***blue*** *and control subjects in* ***red*** *[A] Logistic plot of cancer status (y-axis) against the predicted probability of PDAC (x-axis). The optimal cut-off value (**C/O**) is 0.365 [B] A mosaic plot of actual diagnosis (y-axis) against predicted diagnosis (x-axis). [**C**] ROC curve of the sensitivity against 1-specificity of the NN model with an area under curve of 1.00.*

3.4.4 The diagnostic accuracy of CA19-9 compared to CCGFs

3.4.4.1 Quantification of CA19-9

The serum concentration of CA19-9 in the 180 individuals involved in the current study was determined by ELISA. Of these, 141 samples were previously quantified (***section 3.3.4***) and 39 samples were quantified in the current study using a commercially available CA19-9 ELISA kit. An image of the ELISA plate at quantification is shown in ***Figure 3.13***. It can be seen in ***Figure 3.13*** that many PDAC samples are clearly more yellow in colour compared to healthy controls (virtually all colourless) and the intense yellow colour in some DC and CP samples suggest these patients are likely to have a high serum expression of CA19-9.

	1	2	3	4	5	6	7	8	9	10	11	12
A	A	E	X3	X7	X11	X15	X19	X23	X27	X31	X35	X39
B	A	E	M1479	M1004	P316	T265a	P222	T331	L934	R2212	R2218	M1911
C	B	F	X4	X8	X12	X16	X20	X24	X28	X32	X36	X40
D	B	F	M1421	M1429	T340	P295	P139	P184	L984 (1:1)	R2211	M1905	M1298
E	C	X1	X5	X9	X13	X17	X21	X25	X29	X33	X37	X41
F	C	L1412	L990	M1664	M1483	P311	P218	P215	P306	R2209	M1953	M1298 (1:1)
G	D	X2	X6	X10	X14	X18	X22	X26	X30	X34	X38	X42
H	D	M1804	L1346	P314	P279	P288	P263	L937	R2213	R2216	M1916	M1298 (1:10)
Key	PDAC	BIL	CP	HC								

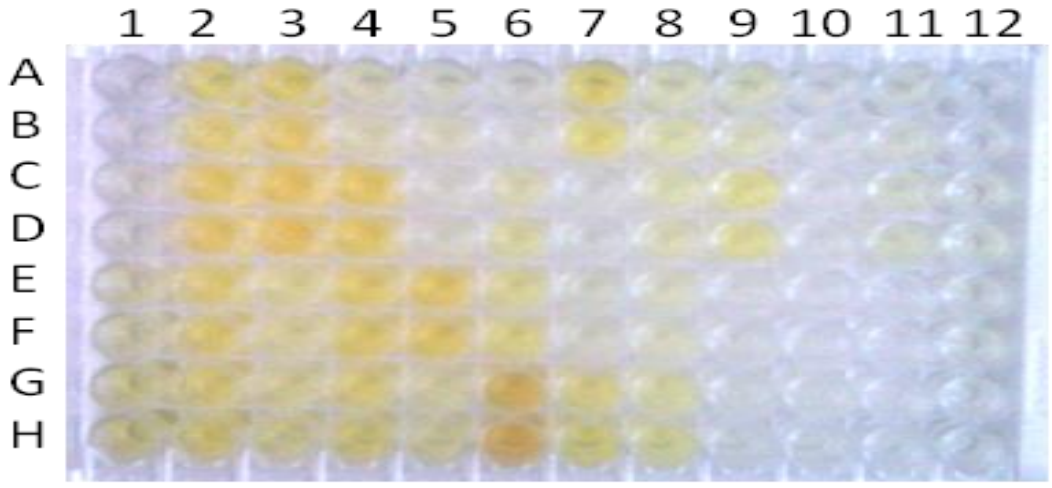


Figure 3.13- [Above] *CA19-9 ELISA plate layout. Note that sample were analysed in duplicates. Wells 1A to 2D were pre-mixed standard references. PDAC samples were coloured in* ***light blue****, disease controls (BIL) in* ***pink****, chronic pancreatitis (**CP**) in* ***grey*** *and healthy volunteers in* ***green****. Sample X40-X42 were from the sample patient but at different dilutions.*
[Below] *A photograph of the ELISA plate at plate reading. Note that the intensity of the yellow colour directly correlates to the concentration of CA19-9.*

3.4.4.2 The diagnostic accuracy of CA19-9 compared to the NN-CCGF algorithm

Univariate analysis showed that CA19-9 was significantly raised in pancreatic cancer compared all control groups *(Wilcoxon's Test, p<0.01;* ***sections 3.4.1.1 to 3.4.1.4*** *)*. Furthermore, it was the most accurate individual marker of PDAC against the HC and CP control subgroups. Therefore, subsequent analyses aimed to compare the diagnostic accuracy of CA19-9 with that of the M-LR and NN combined CCGF algorithms described in ***sections 3.4.2 and 3.4.3*** (for ease of reading, the results for the NN combined CCGF algorithm from ***sections 3.4.2*** is included in ***Table 3.14***).

In the *Discovery Phase*, CA19-9 was highly accurate for discriminating PDAC against HC with ROC-AUC of 0.94 with an optimal sensitivity and specificity of 0.92 and 0.95, respectively (***Table 3.14***). However, CA19-9 was less accurate for discriminating PDAC from the CP and DC control subgroups with ROC-AUCs of 0.86 and 0.73, respectively. Furthermore, on analysis of the diagnostic accuracy of CA19-9 for PDAC against all control subjects, ROC analysis yielded a ROC-AUC of 0.85 and an optimal sensitivity and specificity of 0.73 and 0.90, respectively (***Table 3.14***). Comparing the diagnostic accuracy of CA19-9 for the CP, DC, and HC control groups individually and as a combined Control group showed that CA19-9 alone was less accurate than either the M-LR or the NN combined CCGF algorithm.

Similarly in the *Validation Phase*, CA19-9 was highly accurate at discriminating between PDAC and HC (ROC-AUC= 0.98) but performed less well against CP and DC (ROC-AUC 0.78 and 0.78, respectively; ***Table 3.14***). Furthermore, the diagnostic accuracy of CA19-9 against all control groups yielded ROC-AUC of 0.85 and a relative sensitivity of 0.5 and a relative specificity of 0.90 at the diagnostic phase cut-off value of 125U/ml. Again, comparison between the accuracies of CA19-9 and the M-LR or NN combined CCGF algorithm in the *Validation Phase* showed that CA19-9 alone was a less accurate diagnostic marker of PDAC.

Table 3.14- *The diagnostic accuracies of CA19-9 and the CCGF Neural Network Algorithm*

		Carbohydrate Antigen 19-9							**CCGF Neural Network Algorithm**					
		Discovery			Validation				Discovery			Validation		
Parameters	C/O	AUC	Sens.	Spec.	AUC	Sens.	Spec.	C/O	AUC	Sens.	Spec.	AUC	Sens.	Spec.
PDAC Vs CP	45.0	0.86	0.78	0.90	0.78	0.73	0.70	0.365	0.99	0.97	0.95	1.00	1.00	1.00
PDAC Vs DC	125.0	0.73	0.73	0.75	0.78	0.50	0.80	0.365	1.00	0.97	1.00	1.00	1.00	1.00
PDAC Vs HC	16.0	0.94	0.92	0.95	0.98	0.93	1.00	0.365	1.00	0.97	1.00	1.00	1.00	1.00
PDAC Vs Controls	125.0	0.85	0.73	0.90	0.85	0.50	0.90	0.365	1.00	0.97	0.98	1.00	1.00	1.00

Abbreviations*: ROC area under curve (**AUC**); cut-off value (**C/O**); sensitivity (**Sens.**); Specificity (**Spec.**); Chronic pancreatitis (**CP**), Disease controls (**DC**), Healthy volunteers (**HC**), combined CP+DC+HC (**Controls**)*

3.4.5 CA19-9 in combination with CCGFs

The *Discovery Phase* was repeated with CA19-9 included in the selection process for candidate CCGFs. Interestingly, the same four CCGFs were selected in step 1-4 in the stepwise regression model. However, CA19-9 remained statistically significant after step 4 and therefore it was added to the combination in step 5 (***Table 3.15***). Two-way hierarchical clustering of CA19-9 with the four candidate markers showed a similar pattern compared to the same analysis in ***section 3.4.2.1, Figure 3.8***. However, it was unclear from the heat map as to the exact contribution of CA19-9 in the diagnostic algorithm.

Table 3.15- *CCGF markers selected by the stepwise regression model for combination*

Step	Parameter	Action	Wald/Score p-value	R^2
1	**IL-17**	Entered	< 0.0001	0.23
2	**IL-4**	Entered	< 0.0001	0.51
3	**G-CSF**	Entered	< 0.0001	0.78
4	**IP-10**	Entered	0.0003	0.86
5	**CA19-9**	Entered	0.0048	0.90

Abbreviations: *Interleukin (**IL**); granulocyte colony stimulating factor (**G-CSF**); Interferon-gamma inducible protein 10 (**IP-10**), Carbohydrate Antigen 19-9 (**CA19-9**)*

Note: *R^2 ranges from 0-1 and directly correlates how closely the model fits the data*

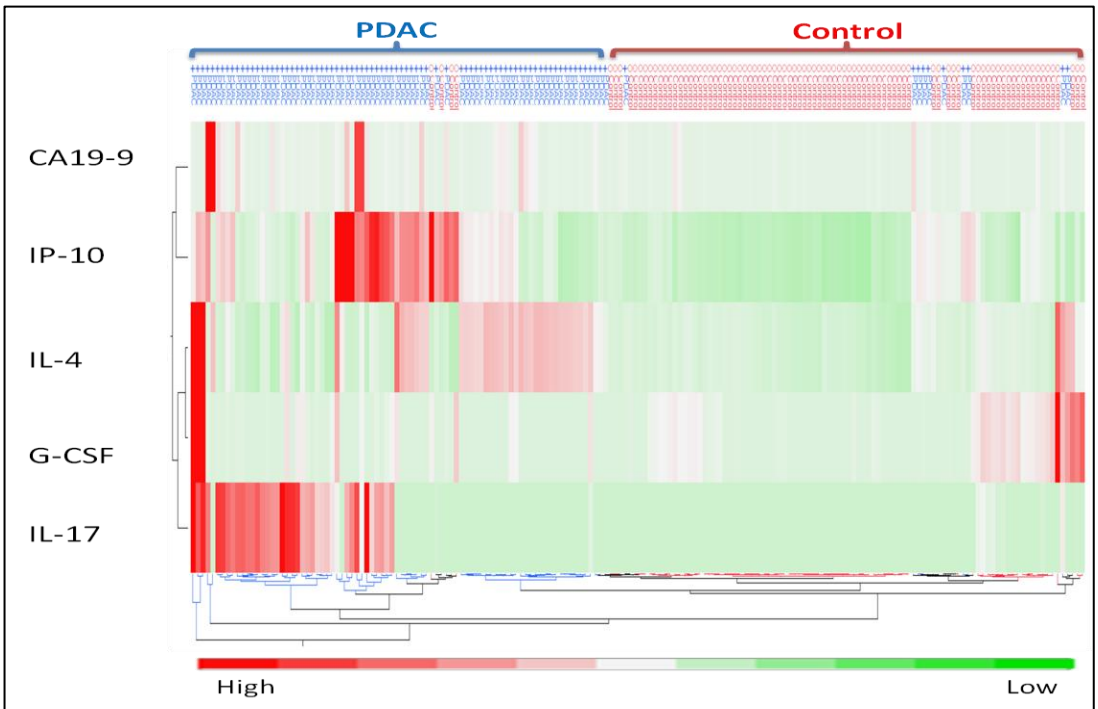


Figure 3.14- *Two-Way hierarchical analysis of CA19-9, IP-10, IL-4, G-CSF, and IL-17. Individuals with PDAC were coloured Blue and controls in Red. Results from the heat map showed a similar pattern to **Figure 3.8** and the samples were broadly clustered into a PDAC dominated group and a Control dominated group. However, the exact contribution of CA19-9 in the prediction process could not be elucidated from the heat map.*

3.4.5.1 Discovery Phase- accuracy of the combined CCGF-CA19-9 marker

Using the same sample randomisation as described in ***section 3.3.2***, the CCGF-CA19-9 combined algorithm achieved ROC-AUCs of 0.999 for both the M-LR and the NN models in the *Discovery Phase*. At the optimal cut-off of 0.38 (M-LR model) and 0.4 (NN model), the models yielded sensitivities of 1.00 and relative specificities of 0.98 (***Table 3.16***, ***Figure 3.15***, ***Figure 3.16***). The CCGF-CA19-9 model appeared to be slightly more robust than the CCGF combined algorithm as evident by the fact that it only misclassified one control subject as PDAC. Furthermore, the logistic plot of cancer status against the predicted probability of PDAC was better defined than the previous models i.e. most samples were either at the control end (probability= 0) or at the PDAC end (probability =1).

Table 3.16- *Diagnostic accuracy of the CCGF-CA19-9 marker in the Discovery Phase*

Parameter	Multinomial Logistic Regression				Artificial Neural Network			
	C/O	AUC	Sens.	Spec.	C/O	AUC	Sens.	Spec.
PDAC Vs CP	0.38	1.00	1.00	0.95	0.4	1.00	1.00	0.95
PDAC Vs DC	0.38	1.00	1.00	1.00	0.4	1.00	1.00	1.00
PDAC Vs HC	0.38	1.00	1.00	1.00	0.4	1.00	1.00	1.00
PDAC Vs Controls	0.38	1.00	1.00	0.98	0.4	1.00	1.00	0.98

Abbreviations*: Cut-off value (**C/O**) for the relative Sensitivity (**Sens.**)and specificity (**Spec.**); Receiver Operatory Characteristics Area Under Curve (**AUC**); resectable pancreatic cancer (**PDAC**); chronic pancreatitis (**CP**); disease controls (**DC**, mainly obstructive jaundice); Healthy volunteers (**HC**)*

Note*: The same algorithms (**M-LR** or **NN**) were used for each parameter. AUCs rounded to 2dp*

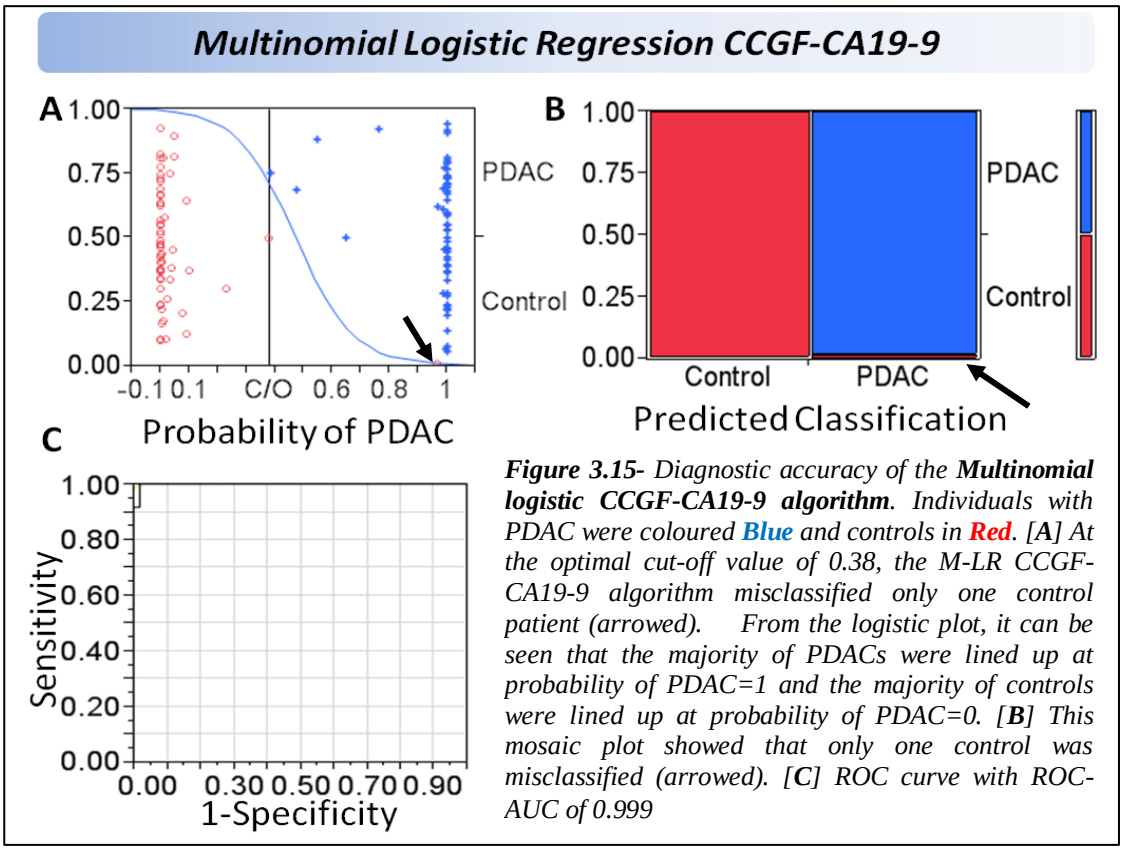


Figure 3.15- *Diagnostic accuracy of the **Multinomial logistic CCGF-CA19-9 algorithm**. Individuals with PDAC were coloured **Blue** and controls in **Red**. [**A**] At the optimal cut-off value of 0.38, the M-LR CCGF-CA19-9 algorithm misclassified only one control patient (arrowed). From the logistic plot, it can be seen that the majority of PDACs were lined up at probability of PDAC=1 and the majority of controls were lined up at probability of PDAC=0. [**B**] This mosaic plot showed that only one control was misclassified (arrowed). [**C**] ROC curve with ROC-AUC of 0.999*

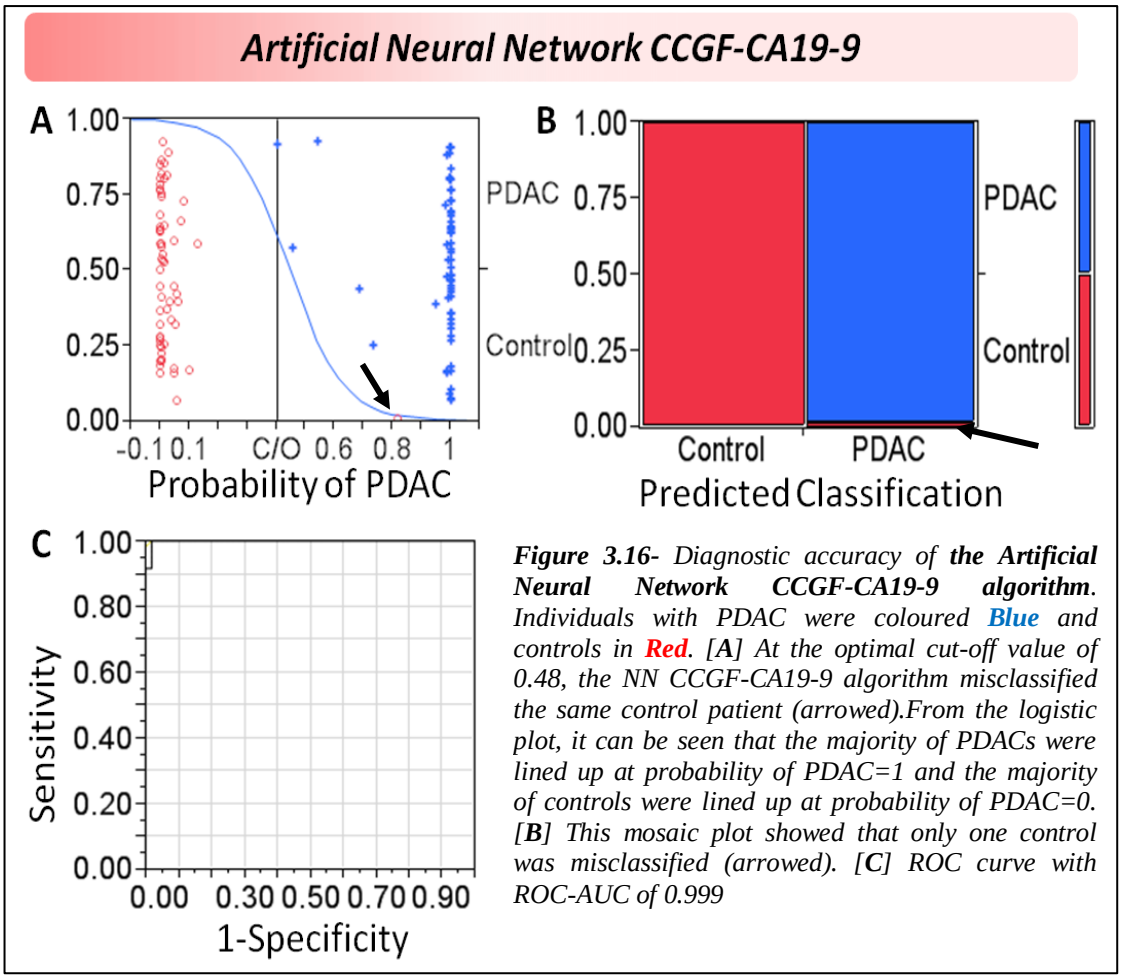


Figure 3.16- *Diagnostic accuracy of* ***the Artificial Neural Network CCGF-CA19-9 algorithm****. Individuals with PDAC were coloured* ***Blue*** *and controls in* ***Red****. [**A**] At the optimal cut-off value of 0.48, the NN CCGF-CA19-9 algorithm misclassified the same control patient (arrowed).From the logistic plot, it can be seen that the majority of PDACs were lined up at probability of PDAC=1 and the majority of controls were lined up at probability of PDAC=0. [**B**] This mosaic plot showed that only one control was misclassified (arrowed). [**C**] ROC curve with ROC-AUC of 0.999*

3.4.5.2 Validation Phase- validation of the CCGF-CA19-9 algorithms

The combined CCGF-CA19-9 algorithms and their relative cut-off values were directly applied to an independent validation sample set consisting of 30 PDAC, 10 CP, 10 DC, and 10 HC (***section 3.3.1.2*** for details regarding the samples).

With the *Discovery Phase* cut-off of 0.38, the multinomial logistic CCGF-CA19-9 algorithm achieved the same accuracy as the M-LR CCGFs algorithm described in ***section 3.4.2.2*** with a ROC-AUC of 0.96 and relative sensitivity and specificity of 0.93 and 1.00, respectively. Interestingly, the two misclassified individuals with pancreatic cancer in the CCGF-CA19-9 algorithm were the same as the two identified by the M-LR combined CCGF algorithm in ***section 3.4.3.1***.

Similarly, the NN CCGF-CA19-9 algorithm achieved the same accuracy as the NN combined CCGF algorithm described in ***section 3.4.3.1*** with a ROC-AUC of 1.00 and relative sensitivity and specificity of both 1.00.

In view of the fact that only one individual with chronic pancreatitis was misclassified in both the Discovery and *Validation Phase*, it would seem that the addition of CA19-9 to the NN algorithm might slightly improve the diagnostic accuracy of the model. Interestingly, review of the misclassified patient's medical records revealed that this patient was initially suspected of pancreatic cancer but there were no signs of the primary tumour of metastasis on contrast enhanced CT and biopsy was negative.

Table 3.17- *Diagnostic accuracy of the CCGF-CA19-9 markers in the Validation Phase*

Parameter	Multinomial Logistic Regression				Artificial Neural Network			
	C/O	AUC	Sens.	Spec.	C/O	AUC	Sens.	Spec.
PDAC Vs CP	0.38	0.96	0.93	1.00	0.4	1.00	1.00	1.00
PDAC Vs DC	0.38	0.95	0.93	1.00	0.4	1.00	1.00	1.00
PDAC Vs HC	0.38	0.95	0.93	1.00	0.4	1.00	1.00	1.00
PDAC Vs Controls	0.38	0.96	0.93	1.00	0.4	1.00	1.00	1.00

Abbreviations*: Cut-off value (**C/O**) for the relative Sensitivity (**Sens.**)and specificity (**Spec.**); Receiver Operatory Characteristics Area Under Curve (**AUC**). **Note**: The same algorithms (**M-LR** or **NN**) were used for each parameter. AUCs rounded to 2dp*

Validation: Multinomial Logistic Regression CCGF-CA19-9

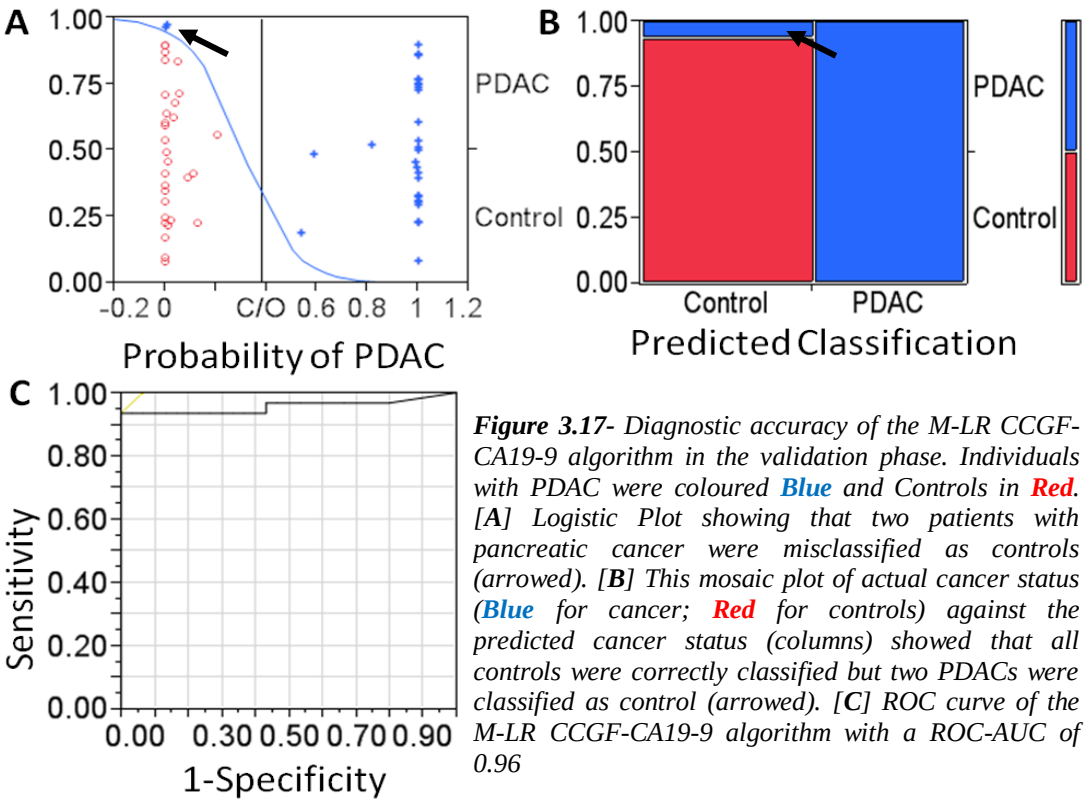


Figure 3.17- *Diagnostic accuracy of the M-LR CCGF-CA19-9 algorithm in the validation phase. Individuals with PDAC were coloured* ***Blue*** *and Controls in* ***Red****. [**A**] Logistic Plot showing that two patients with pancreatic cancer were misclassified as controls (arrowed). [**B**] This mosaic plot of actual cancer status (**Blue** for cancer; **Red** for controls) against the predicted cancer status (columns) showed that all controls were correctly classified but two PDACs were classified as control (arrowed). [**C**] ROC curve of the M-LR CCGF-CA19-9 algorithm with a ROC-AUC of 0.96*

Validation: Artificial Neural Network CCGF-CA19-9

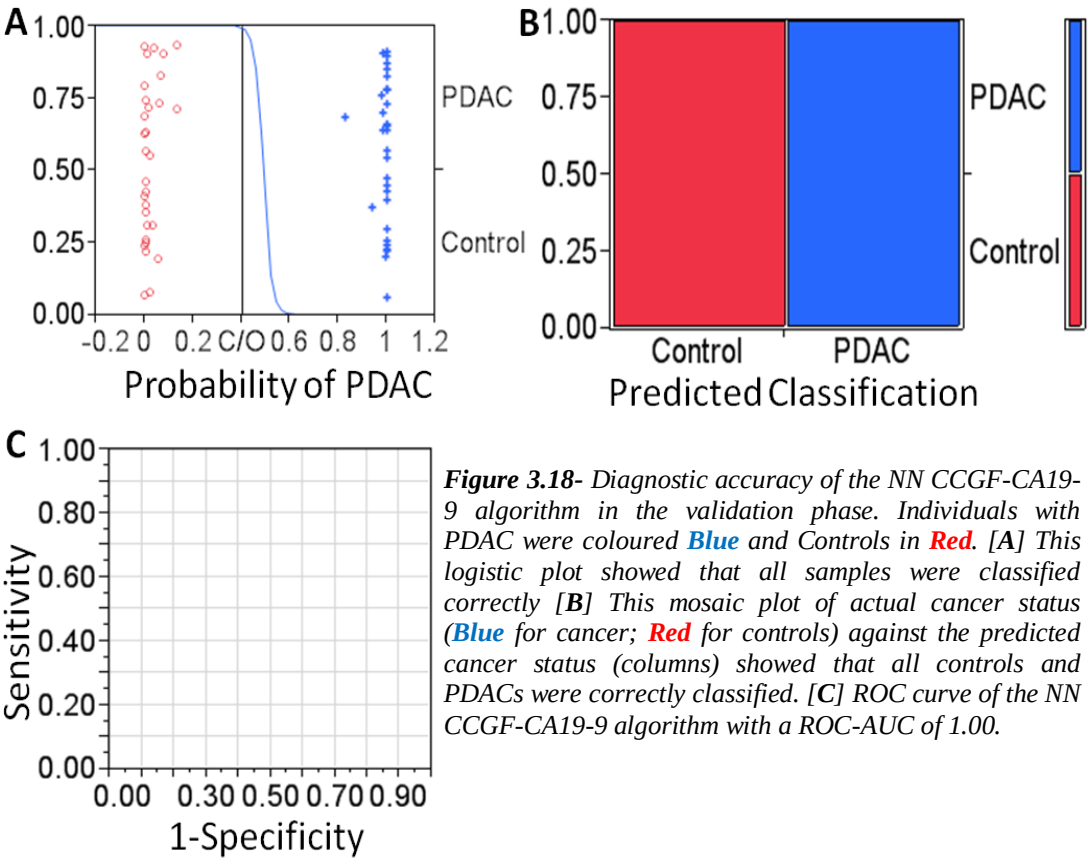


Figure 3.18- *Diagnostic accuracy of the NN CCGF-CA19-9 algorithm in the validation phase. Individuals with PDAC were coloured* ***Blue*** *and Controls in* ***Red****. [**A**] This logistic plot showed that all samples were classified correctly [**B**] This mosaic plot of actual cancer status (**Blue** for cancer; **Red** for controls) against the predicted cancer status (columns) showed that all controls and PDACs were correctly classified. [**C**] ROC curve of the NN CCGF-CA19-9 algorithm with a ROC-AUC of 1.00.*

3.4.6 The impact of each biomarker on the M-LR diagnostic algorithm

A separate analysis was performed to assess the contribution of each CCGF marker towards the diagnostic accuracy of the M-LR CCGF algorithm and M-LR CCGF-CA19-9 algorithm for PDAC against all controls. This was achieved by assessing the diagnostic accuracies of the M-LR algorithms when each CCGF was independently removed. It should be noted that a new M-LR algorithm was generated upon the removal of each CCGF in order to maximise the diagnostic accuracy of the remaining markers.

Results from the *Discovery Phase* suggest that the removal of IL-4, IL-17, G-CSF, IP-10, or CA19-9 from the M-LR algorithms will lead to a decrease in the diagnostic accuracy of the models (***Table 3.18***). However, the effect of removing one CCGF appeared to be well compensated by the remaining markers. In particular, the removal of IP-10 caused minimal decrease in the ROC-AUC in both the M-LR CCGF and the M-LR CCGF-CA19-9 algorithms (AUC, 0.994 versus 0.984 and 0.999 versus 0.995, respectively; ***Table 3.18***).

Results from the *Validation Phase* are displayed in ***Table 3.19***. When the M-LR CCGF algorithm generated in the *Discovery Phase* were directly applied to the Validation sample set, the impact of the removal of IL-17 was immediately obvious with a decrease in AUC from 0.940 to 0.814 on removal. Furthermore, there is a general decrease in the diagnostic accuracy in the *Validation Phase* compared to the *Discovery Phase* upon the removal of any CCGF.

Table 3.18- *Accuracy of the M-LR CCGF and CCGF-CA19-9 markers in the Discovery Phase*

Parameter	M-LR CCGF algorithm				M-LR CCGF-CA19-9 algorithm			
(PDAC versus all controls)	C/O	AUC	Sens.	Spec.	C/O	AUC	Sens.	Spec.
Original Algorithm	0.400	0.994	0.97	0.98	0.38	0.999	1.00	0.98
Removal of IL-4	0.369	0.928	0.90	0.87	0.38	0.931	0.88	0.88
Removal of IL-17	0.380	0.940	0.90	0.90	0.33	0.957	0.93	0.92
Removal of G-CSF	0.460	0.958	0.88	0.93	0.39	0.969	0.93	0.90
Removal of IP-10	0.350	0.984	0.97	0.97	0.40	0.995	0.98	0.97

Table 3.19- *Accuracy of the M-LR CCGF and CCGF-CA19-9 markers in the Validation Phase*

Parameter	M-LR CCGF algorithm				M-LR CCGF-CA19-9 algorithm			
(PDAC versus all controls)	C/O	AUC	Sens.	Spec.	C/O	AUC	Sens.	Spec.
Original Algorithm	0.400	0.957	0.93	1.00	0.38	0.956	0.93	1.00
Removal of IL-4	0.369	0.898	0.80	0.80	0.38	0.956	0.93	1.00
Removal of IL-17	0.380	0.814	0.77	0.83	0.33	0.929	0.80	0.83
Removal of G-CSF	0.460	0.989	0.87	0.97	0.39	0.993	0.93	0.97
Removal of IP-10	0.350	0.972	0.97	0.97	0.40	0.967	0.93	1.00

3.5 Impact of clinical-demographical factors on the accuracy of the disease-predicting models

3.5.1 Impact of Patient Age on the serum levels of candidate CCGFs

Analyses of demographical data in ***section 3.3.1*** suggested that, as expected individuals with PDAC were generally older compared to individuals in the chronic pancreatitis or the healthy control subgroups. Although results shown no significant difference in age between the PDAC and the DC control sub-group, it is still important to identify any correlations between age, individual CCGFs, and the disease-predicting formulae.

Correlation analyses of age against serum concentrations of IL-4, IL-17, G-CSF, and IP-10 yielded Kendall Tau correlation coefficients (τ) of <0.3 (both control and PDAC group) suggesting that age is independent to the serum concentrations of the four candidate CCGFs (***Table 3.20***). Furthermore, correlation analysis of age and estimated probabilities generated from the prediction formulae yielded Kendal τ coefficients of <0.2, which indicate that there is minimal correlation between the two variables. A hierarchical clustering analysis was performed for age against each CCGF individually and graphically represented in ***Figure 3.19***. Results showed that although individuals with PDAC tended to be older, the serum concentrations the four CCGFs were not influenced by age.

Table 3.20- *Correlation between Age, individual CCGFs, and disease-predicting formulae*

	Controls		PDAC	
Parameters	Kendall τ	Prob>|τ|	Kendall τ	Prob>|τ|
IL-4	-0.03	0.67	0.05	0.52
IL-17	0.01	0.87	0.02	0.81
G-CSF	0.13	0.11	-0.07	0.40
IP-10	0.28	<0.01	0.15	0.05
M-LR algorithm	-0.08	0.27	0.15	0.04
NN algorithm	-0.18	0.01	0.04	0.63

Abbreviations*: Interleukin (**IL**); Granulocyte colony stimulating factor (**G-CSF**); Interferon-gamma inducible protein-10 (**IP-10**); Multinomial Logistic Regression (**M-LR**); Artificial neural network (**NN**)*
Note: *Kendall Tau (τ) coefficients ranges from -1 (perfect negative correlation) to 0 (no correlation) to +1 (perfect positive correlation); Prob>|τ| represents the probability that the actual correlation coefficient (τ) would be greater than the calculated coefficient.*

Figure 3.19- *Two-way hierarchical clustering of Age against IP-10, IL-4, G-CSF, and IL-17. Results from the heat map suggested that although individuals with PDAC were generally older, the IL-4, IL-17, and IP-10 were not raised with to age and G-CSF was not decreased with age. Therefore it can be deduced that there is no correlation between age and the serum concentrations of the four CCGFs*

3.5.2 Impact of diabetes on the serum levels of candidate CCGFs

The correlation between the presence of diabetes (a frequently reported disease in PDAC), the serum levels of candidate CCGFs, and the estimated probability values generated by the disease-predicting formulae were investigated. Univariate analyses of the serum levels of candidate CCGF markers and the estimated probabilities generated by the disease-predicting formulae against available data on diabetes status yielded insignificant results (Wilcoxon's Test, p>0.05; ***Table 3.21***). Furthermore, a final analysis was performed to determine the accuracy of the two disease-predicting formulae for discriminating between diabetics and non-diabetic (***Figure 3.20***). Results from this analysis showed that neither algorithms is accurate at discriminate between diabetics and non-diabetics (ROC<0.7). It should be noted however, that there is a substantial number of missing data (especially for HC subgroup), which may decrease the integrity of these results.

Table 3.21- *Correlation between Diabetes, individual CCGFs, and disease-predicting formulae*

Parameters	Characteristic	Univariate Analysis, p-value		
		Controls	PDAC	All data
IL-4	Diabetes	0.43	0.54	0.81
IL-17		0.30	0.96	0.78
G-CSF		0.92	0.93	0.98
IP-10		0.17	0.46	0.23
M-LR algorithm		0.08	0.70	0.86
NN algorithm		0.85	0.65	0.72

Abbreviations*: Interleukin (**IL**); Granulocyte colony stimulating factor (**G-CSF**); Interferon-gamma inducible protein-10 (**IP-10**); Multinomial Logistic Regression (**M-LR**); Artificial neural network (**NN**)*

Note: *Wilcoxon's Test was used for the univariate analysis of non-parametric variables between two groups*

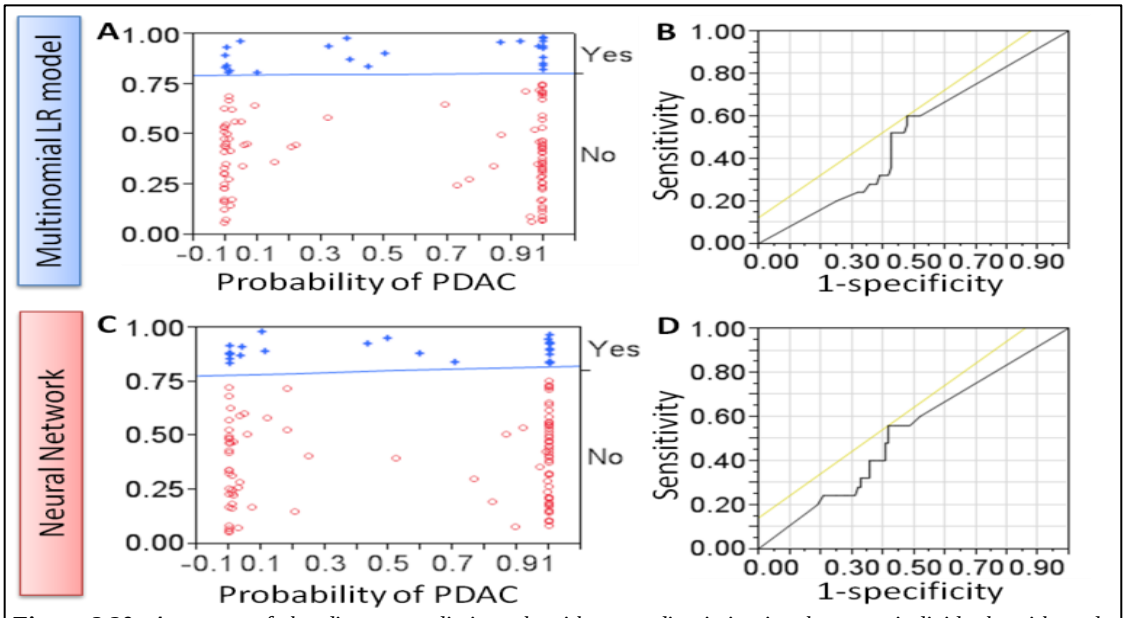


Figure 3.20- *Accuracy of the disease-predicting algorithms at discriminating between individuals with and without diabetes. [**A, C**] Logistic plots showing the distribution of individuals with diabetes (**blue**) and without diabetes (**red**) against the probability of PDAC. [**B, D**] ROC-curves showing the accuracy of the algorithms for predicting the presence of diabetes. ROC-AUC: multinomial logistic regression (**M-LR**) =0.51 and neural network (**NN**) =0.53*

3.5.3 Impact of smoking on the serum levels of candidate CCGFs

Univariate non-parametric analysis was performed to assess the impact of smoking on the serum concentrations of candidate CCGFs. Results indicated that smoking does not appear to influence the serum concentrations of IL-4, IL-17, or G-CSF. However, there seemed to be a statistically significant decrease in the serum levels of IP-10 in smokers and when all data were combined (Wilcoxon's Test, p=0.04 and p=0.01 respectively). However, further analysis using the probability values generated by the two algorithms yielded insignificant results on univariate and ROC analyses (Wilcoxon's Test, p≥0.05; ROC <0.7). Again, these results should be considered in view of the fact that there is a large number of missing data for smoking status, especially in the HC control subgroup.

Table 3.22- *Correlation between smoking, individual CCGFs, and disease-predicting formulae*

Parameters	Characteristic	Univariate Analysis, p-value		
		Controls	PDAC	All data
IL-4	Smoking	0.53	0.52	0.36
IL-17		0.23	0.54	0.66
G-CSF		0.37	0.49	0.24
IP-10		0.06	**0.04 (↓)**	**0.01 (↓ in PDAC)**
M-LR algorithm		0.53	0.05	0.09
NN algorithm		0.15	0.49	0.61

Abbreviations*: Interleukin (**IL**); Granulocyte colony stimulating factor (**G-CSF**); Interferon-gamma inducible protein-10 (**IP-10**); Multinomial Logistic Regression (**M-LR**); Artificial neural network (**NN**)*

Note: *Wilcoxon's Test was used for the univariate analysis of non-parametric variables between two groups*

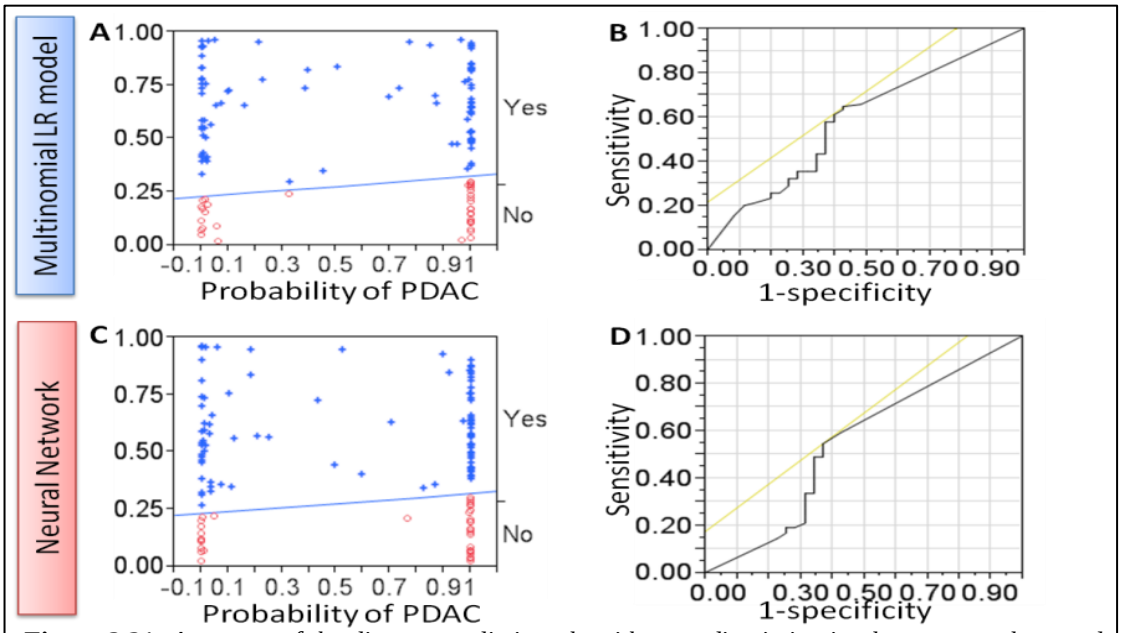


Figure 3.21- *Accuracy of the disease-predicting algorithms at discriminating between smokers and non-smokers. [**A, C**] Logistic plots showing the distribution of smokers (**blue**) and non-smokers (**red**) against the probability of PDAC. [**B, D**] ROC-curves showing the accuracy of the algorithms for identifying smokers. ROC-AUC: multinomial logistic regression (**M-LR**) =0.58 and neural network (**NN**) =0.54*

3.6 Discussion

Despite evidence from a number of robust studies describing the intimate and complex relationships between inflammatory mediators, tumour microenvironment, tumour-associated macrophages, inflammation, and cancer in general[150, 187, 191, 196, 243-246], the exact roles of cytokines, chemokines, and growth factors (CCGFs) in pancreatic cancer continue to elude researchers. It is known that the growth, proliferation, and metastasis of cancer cells are governed by many different mechanisms and pathways[162, 167, 169, 196, 247]. However, the communication within and between these mechanisms and pathways are invariably dependent upon various stimulatory and inhibitory mediators, especially CCGFs[167, 169, 179, 248-249].

We hypothesized that the initiation and continuation of cellular growth, proliferation, and metastasis in pancreatic cancer would require the production and inhibition of a number of specific CCGFs. It follows therefore, that there should be a different serum CCGF profile between individuals with pancreatic cancer and individuals with benign pancreatic inflammatory diseases such as chronic pancreatitis and benign biliary obstruction. Results from the current study confirmed this as the concentrations of fourteen CCGFs were differentially observed in the serum of individuals with pancreatic cancer compared to individuals in the CP, DC, and HC control subgroups.

Results from the *Discovery Phase* (***section 3.4.2.1***) showed two important findings. First, the serum concentrations of some CCGFs appeared to be directly correlated only in individuals with PDAC (e.g. IL-17 and IL-10), which could indicate that some CCGFs may be involved in a common signalling pathway in cancer. Alternatively, it may an artefact of the heterogeneity of the combined Control group where a correlation between CCGFs in one control subgroup may be masked by the absence of this correlation in the other two control subgroups. With in mind, a stepwise regression model was employed to select CCGFs that were independently expressed in PDAC compared to others. This led to the second important finding that no single CCGF was differentially observed in all individuals with pancreatic cancer compared to Controls (i.e. there is no perfect marker). Rather, we report that each of the four independent CCGFs appeared to be responsible for identifying a unique sub-group of patients in the PDAC or Control groups, which may mean that there are

four distinctly different cytokine mediated processes in these individuals. We report that IL-4, IL-17, and IP-10 were responsible for identifying a subset of patients with PDAC where as G-CSF may play a role in identifying individuals with pancreatic inflammatory disease.

3.6.1 CCGFs for the diagnosis of pancreatic cancer

For the past few decades, researchers have sought to identify a good and useful biomarker for pancreatic cancer. Considering the current lack of success in finding an accurate standalone marker of PDAC and in view of the evidence suggesting that the growth, proliferation, survival, and metastasis of PDAC is likely to be influenced by several independent pathways and processes (***section 1.2.3*** *and* ***1.4.2***), one might be tempted to conclude that a "magic marker" for PDAC is extremely unlikely to exist. Indeed, results from ***section 3.4.1*** would support this predicament. Despite observing that a relatively large number of CCGFs were differentially expressed in patients with PDAC compared to controls, only IFN-γ achieved a similar sensitivity and specificity to the widely used biomarker of PDAC, CA19-9, which means that, as expected[250], CCGFs are not sufficiently accurate as individual markers of pancreatic cancer.

In consideration of the above findings, the current study utilised two independent models to combine a carefully selected panel of CCGFs- the Multinomial Logistic Regression model and the Neural Network model. The diagnostic accuracies for these models were very encouraging with ROC-AUCs of ≥0.99 for discriminating between PDAC against control subjects. Furthermore, in view of the results from ***section 3.4.3.1*** and those from the validation of the NN CCGF algorithm, which demonstrated a perfect classification for the independent validation sample set, we conclude that a combined CCGF marker consisting of IL-4, IL-17, G-CSF, and IP-10 may be used as a potential diagnostic biomarker for PDAC.

Finally, we report that the addition of CA19-9 into the combined CCGF algorithms may marginally improves the diagnostic accuracy of both combined CCGF algorithms (M-LR and NN). However, results in ***section 3.4.6*** would suggest that CA19-9 should not be used in place of any of the four CCGFs

3.6.2 The roles of Il-4, IL-17, G-CSF, and IP-10 in PDAC and pancreatic inflammatory diseases

3.6.2.1 Interleukin 4

IL-4 is a cytokine that is responsible for the activation of a number of immune cells (e.g. B-lymphocytes) and plays an important role in promoting the proliferation of T-lymphocytes[251]. In addition, studies have shown that IL-4 is involved in the differentiation of CD4+ T-cells into Type II T-Helper (Th2) cells[139, 252-254].

Interestingly, existing literature appears to suggest that IL-4 plays a paradoxical role in cancer[251]. Several studies have demonstrated that IL-4 exhibits anti-tumour properties by inducing anti-tumour immune response, indeed, the potential of IL-4 as an anti-tumour agent was demonstrated in a number of solid tumours including renal and colorectal cancers[255-259], although results from later studies on the therapeutic potential of IL-4 were disappointing[260-262]. There is increasing contrary evidence indicating that IL-4 acts to protect tumour cells from apoptosis[251]. Concordant to the results from the current study, evidence from a number of clinical studies have demonstrated that the IL-4 is up regulated in patients with pancreatic cancer as well as other types of solid cancers including renal cell, lung, colon, and breast cancers[263]. Furthermore, a study by Onishi *et al.* reported that the IL-4 expression at the tumour site is associated with the stage and grade of renal cell carcinoma[264].

Despite these observations, the exact biological mechanisms by which IL-4 exert its tumour promoting ability in pancreatic cancer is undetermined. However, recent evidence from *in vitro* and *in vivo* studies have demonstrated that high levels of IL-4 in the tumour microenvironment can cause the induction of cathepsin activity in tumour-associated macrophages and therefore may play an indirect role in promoting the remodelling of the tumour microenvironment and ultimately facilitating tumour growth, proliferation, and metastasis[265].

3.6.2.2 Interleukin 17

IL-17 belongs to a relatively new subclass of cytokines and is predominantly but not exclusively produced by T-helper 17 (Th17) cells which are a newly designated subset of CD4+ T-cells[266]. Other immune cells known to secrete IL-17 include natural killer T cells, CD8+ T cells, and lymph tissue inducer cells[266]. The role of IL-17 in malignant diseases is not completely characterised. However, studies have

reported that IL-17 plays a role in promoting carcinogenesis and tumour growth by facilitating angiogenesis and stimulating remodelling to the extracellular matrix[267-268]. Indeed, *in vivo* studies have demonstrated that, IL-17 can stimulate the production of VEGF, prostaglandin E1, and prostaglandin E2 in fibroblasts thereby enhancing tumour angiogenesis[267]. Furthermore, recent evidence indicated that IL-17 is associated with an increased MMP expression, which in turn facilitates angiogenesis through the destruction of the extracellular matrix[266].

The biology behind IL-17 induced angiogenesis in pancreatic cancer specifically has not been previously reported. However, recent studies have indicated that the production of IL-17 (and subsequent angiogenesis) may be regulated by the STAT3 transcription factor[169]. It has been proposed that the activation of the STAT3 pathway by stimuli from cytokine receptors such as IL-6R and IL-10R promotes the gene transcription and subsequent production of a number of inflammatory cytokines, including IL-6, IL-10, and IL-17, which are responsible for promoting tumour survival, proliferation, and angiogenesis[169]. Furthermore, studies have indicated that many of these up regulated cytokines are also activators of the STAT3 pathway and therefore forming a positive feedback loop for their own production[169].

In view of the results from ***sections 3.4.1.5*** *and* ***3.4.2.1*** demonstrating that IL-17 is an independent indicator of PDAC and that is expression is closely correlated to that of IL-10, it is likely that IL-17 may be involved as an activator and/or a downstream product of the STAT3 transcription factor pathway in certain individuals with pancreatic cancer.

3.6.2.3 G-CSF

G-CSF, also called CSF3 is one of four members of the colony-stimulating factor (CSF) family proteins[269]. Other members of the CSF family including macrophage colony stimulating factor (M-CSF), granulocyte macrophage colony stimulating factor (GM-CSF), and Multi colony stimulating factor (M-CSF)[269]. G-CSF is a 19.6 kDa glycosylated protein consisting of 174 amino acid residues and is mainly produced by macrophages[270].

The role of G-CSF in pancreatic cancer is unclear. However, studies have shown that in the acute inflammatory setting, G-CSF plays an essential role in promoting the proliferation of neutrophil colonies, inducing the differentiation of precursor cells to neutrophils, and stimulating the activity of mature neutrophils[269-270]. Indeed, results from ***section 3.4.1*** showed that the serum concentration of G-CSF is elevated in certain individuals with CP and biliary obstruction compared to PDAC. We therefore propose that G-CSF is not an indicator of pancreatic cancer. Rather, it is an indicator of inflammatory diseases where an elevated G-CSF serum concentration reflects the increased neutrophil activity in benign inflammatory diseases.

3.6.2.4 IP-10

IP-10 is a dimerized 10kDa protein consisting of 98 amino acids, which is predominantly secreted by macrophages in response to interferon gamma[271]. IP-10 belongs to the CXCL chemokine subfamily and is responsible for the recruitment of monocytes and T-lymphocytes to its site of production[271].

Whilst there is a current lack of evidence associating IP-10 with pancreatic cancer, existing literature indicates that IP-10 may play a dual role as an anti-tumour agent for cancer in general[272]. First, IP-10 is responsible for the chemotaxis of tumour infiltrating T-lymphocytes[273], in particular NK T-cells, and secondly, studies have shown that IP-10 have a role in the prevention of angiogenesis[272, 274-275]. A study by Musha *et al.* demonstrated that Th1 cells along the invasive margin of colorectal cancer produce IFN-gamma, which stimulates cancer cells and macrophages to produce IP-10 and ultimately resulting in the infiltration of CXCR3 expressing T-cells at the invasive margin[273].

In view of the pro-invasive and aggressive nature of pancreatic cancer and the observation in ***section 3.4.1*** that a subset of individuals with PDAC exhibits a higher serum level of IP-10, it is possible that the increase in serum level of IP-10 may be the direct result of host immune response against PDAC along the invasive margin.

Chapter 4

Final Discussion, Limitations, and Future Directions

4.1 Final Discussion and Conclusions

The diagnosis of pancreatic cancer is seldom straightforward as patients often present with non-specific, systemic symptoms such as pain, weight loss, and jaundice[55]. There are two major problems to the current system for the diagnosis of pancreatic cancer. The first problem is that most patients seek help from their doctors only when they have systemic symptoms (tip of the clinical iceberg). These systematic symptoms are often indications of advanced disease and therefore, it is not surprising that up to 80% of patients have metastatic disease on presentation. The second problem is that patients are referred for further investigation by CE-CT based on the vague systemic symptoms, which can also be present in many other benign diseases such as peptic ulcer, gall stone obstruction, and cholangitis to name but a few. Therefore, the use of an accurate diagnostic biomarker of pancreatic cancer can significantly reduce the number of patients with non-malignant diseases undergoing an unnecessary invasive and potentially harmful diagnostic procedure.

In the past few decades, researchers have sought to identify proteomic biomarkers for the diagnosis of pancreatic cancer via a number of different approaches. In particular, researchers are increasingly interested in newer biomarker discovery techniques such as mass spectrometry and protein microarray, which can simultaneously identify multiple differentially expressed proteins in pancreatic cancer. However, the progress of biomarker research can often be impeded by the verification techniques such as western blotting, ELISA, and immunohistochemistry, which are only capable of verifying the expression of one protein per experiment. The invention of novel experimental methods, such as infrared detection of western blots and the development of multiplex assays, have partially addressed this issue. Nonetheless, these methods are invariably more expensive (per experiment) compared to traditional methods and in particular, the use of multiplexing assays is often limited by the availability, sensitivity, and specificity of the detection antibodies.

Currently, Carbohydrate Antigen 19-9 (CA19-9) is the most widely used biomarker for PDAC. However, with reported sensitivities and specificities of 0.87 and 0.73 from the current series, CA19-9 is neither adequately specific nor adequately sensitive as a diagnostic marker. Recent studies have proposed several novel standalone markers with high discriminatory power for PDAC and healthy subjects;

however, the diagnostic accuracy of the majority of these markers were either inferior to CA19-9 or performed poorly in distinguishing between PDAC and chronic pancreatitis [100, 111, 276-283].

Considering the current lack of success in finding an accurate standalone marker of pancreatic cancer and in view of evidence suggesting that the growth, proliferation, survival, and metastasis of PDAC is likely to be influenced by a number of independent pathways, it is clear that a single standalone marker for pancreatic cancer is extremely unlikely to exist. Therefore, a possible solution would be to devise a mathematical algorithm, which would combine a panel of carefully selected markers each capable of identifying a unique signalling pathway or characteristic found in PDAC. Indeed, a number of studies have attempted to improve the diagnostic accuracy of existing and novel biomarkers for pancreatic cancer using various classification-modelling methods [276, 284-286]. In particular, a recent study by Zhang *et al.* reported encouraging results using a panel of four salivary transcriptomic biomarkers (*KRAS, MBD3L2, ACRV1, and DPM1*) combined using multinomial logistic regression with reported sensitivity and specificity of 90% and 95%, respectively, for discriminating between patients with PDAC against individuals with chronic pancreatitis and against healthy volunteers[286].

Nevertheless, despite the successes reported by studies combining multiple markers, their accuracies in the presence of pancreaticobiliary diseases other than chronic pancreatitis remain to be validated particularly in light of recent evidence from Yan *et al.*[225] suggested that the diagnostic accuracy of novel markers might be significantly reduced when analysed against individuals with benign biliary obstruction.

4.1.1 Conclusions from the current study

In chapters 2 and 3 of this thesis, the diagnostic accuracies of three iTRAQ-identified proteins and twenty-seven CCGFs as individual and combined markers of resectable pancreatic cancer were explored. Results from both chapters were encouraging.

4.1.1.1 Validation of iTRAQ results

In accordance with previous studies on the expression of VDBP and RBP-4 various types of cancers, results from chapter 2 indicate that the serum concentration of VDBP, RBP-4, and FINC were significantly decreased in pancreatic cancer compared to healthy controls and chronic pancreatitis[203, 212, 220, 231, 287] (***sections 2.4.2.6, 2.4.2.7,*** *and* ***2.4.2.10***). Furthermore, the three proteins achieved a statistically significant accuracy for diagnosing PDAC against the HC and CP. However, in support of the study presented by Yan *et al.*[225], the diagnostic accuracies of the markers were significantly decreased when faced with patients with biliary obstruction (***section 2.4.2.8***). Therefore, it can be concluded that whilst VDBP, RBP-4, and FINC were relatively accurate at discriminating between PDAC against CP and HC, their diagnostic accuracy were far inferior to that of CA19-9 and therefore not sufficiently accurate to be used for the diagnosis of pancreatic cancer. In addition, results from the current study would suggest that biliary obstruction is a disease-related confounding factor for the diagnostic accuracy of VDBP, RBP-4, and FINC.

Subsequent analyses sought to improve the diagnostic accuracies by combining the markers into a single marker using the M-LR modelling method. Results were encouraging for PDAC against CP and HC with a reported sensitivity of 0.85 and a specificity of 0.95. However, the confounding effect of biliary obstruction was immediately obvious when all control subgroups were considered (sensitivity=0.70; specificity=0.80; ***section 2.4.2.10***).

Phase III of chapter 2 was designed to validate the iTRAQ results with a different sample set- the UKCTOCS pre-diagnostic serum samples. Western blotting results from this experiment indicated that in the 0-0.5 year pre-diagnosis category, there is a trend for decreased serum concentrations of VDBP and RBP-4 in the PDAC group compared to their matched controls. However, this tendency did not reach statistical significance due to the relatively small sample size (n=10 per group). Furthermore, that the serum concentrations of VDBP and RBP-4 in pre-pancreatic cancer patients

were comparable with the matched controls in the other time categories (***section 2.4.3***). Interestingly, we noted that certain patients exhibited a gradual decrease in serum VDBP and RBP-4 nearer to diagnosis. However, the significance of this trend must be further analysed before any conclusions can be made.

Although there are potential advantages for using the UKCTOCS pre-diagnostic serum samples for the discovery and validation of potential markers for the screening of PDAC, there are also a number of disadvantages. Firstly, these samples were from female participants only and therefore results derived from these samples would require validation in male subjects. Secondly, ~80% of all patients with pancreatic cancer have advanced disease on diagnosis therefore even though a serum sample may be labelled as pre-diagnostic, the individual may already be at an advanced of pancreatic cancer on sample collection, especially for samples collected in the weeks/months prior to diagnosis. This is particularly important because in the absence of data regarding the tumour staging for these patients, it would be difficult to determine whether an individual designated as pre-pancreatic cancer is likely to have already developed pancreatic cancer at the >1 year time-categories.

4.1.1.2 Discovery and validation of CCGF markers

In Chapter 3, we described the use of a multiplex cytokines assay to quantify the serum concentrations of 27 CCGFs. Furthermore, we assessed the diagnostic potential of these CCGFs as an individual and as a combined marker for pancreatic cancer. Results from the initial analysis identified fourteen differentially observed serum CCGFs in PDAC compared to the HC, CP, and DC control subgroups. In particular, IFN-γ was found to have a comparable diagnostic accuracy to the widely used pancreatic cancer biomarker, CA19-9 with sensitivities of 0.77 versus 0.87 and specificities of 0.89 and 0.73, respectively. Nevertheless, this level of accuracy is insufficient for the diagnosis of pancreatic cancer in a clinical setting.

Subsequent analysis sought to improve the diagnostic accuracy of CCGFs by combining independent markers of PDAC (including IL-4, IL-17, G-CSF, and IP-10) using two separate modelling methods: M-LR and NN modelling methods. Results from the *Discovery Phase* were encouraging with reported sensitivities of 0.97 and specificities of 0.98. Furthermore, analysis of the diagnostic accuracies of the M-LR

CCGF and NN CCGF algorithms in the *Validation Phase* yielded remarkable results. In particular, the NN CCGF algorithm achieved a perfect diagnostic accuracy.

Direct comparison between the diagnostic accuracy of the NN CCGF models against the existing biomarker for pancreatic cancer, CA19-9, in the *Discovery Phase* indicated that the NN CCGF model was the more accurate biomarker compared to CA19-9 with sensitivities of 0.97 versus 0.73 and specificities of 0.98 versus 0.90. This was reflected in the *Validation Phase* with sensitivities of 1.00 for the NN CCGF algorithm versus 0.50 for CA19-9 and specificities of 1.00 against 0.90, respectively.

In a final analysis, the diagnostic potential of combining CA19-9 with IL-4, IL-17, G-CSF, and IP-10 was assessed. Results showed that the NN CCGF-CA19-9 marker performed better in the *Discovery Phase* than the NN CCGF marker with reported sensitivities of 1.00 versus 0.97 and specificities of 0.98 versus 0.98, respectively. In addition, the CCGF-CA19-9 marker also achieved perfect diagnostic accuracy in the *Validation Phase*.

In view of the results presented in Chapter 3, we report that a combined marker CCGF marker consisting of IL-4, IL-17, G-CSF, and IP-10 may serve as an accurate diagnostic marker of pancreatic cancer. Furthermore, we conclude that the addition of CA19-9 may improve the sensitivity of the NN CCGF algorithm, which may be more desirable in the clinical setting.

On a more technical note, we would like to emphasize that mathematical predication models combining multiple variables, such as the M-LR and the NN models, are susceptible to data overfitting. The reason for this is that whilst it would be entirely possible to include all 27 CCGFs in the prediction model and the resulting accuracy would likely to be extremely high, it would also mean that the model no long describes the structure of the data, rather, it would be fitting to individual points. Therefore, models with a large number of predictors tend not to generalise well. Chapter 3 was designed to account for this effect by the use of an independent validation sample set. However, based on the highly accurate predictive ability of the combined algorithms, we can conclude that there is minimal data overfitting present in our study.

4.1.2 General Limitations

Although the results from the current study were promising, conclusions drawn from the current study have several limitations. For example, due to limitations in sample availability, the current study did not examine the serum CCGF profiles for other inflammatory diseases and other cancers and therefore there is doubt to whether the disease-predicting algorithm can be generalised for other benign and malignant non-pancreatic diseases. Furthermore, as highlighted in a recent review article by Schrohl *et al.*, pre-analytical parameters may have an impact on the protein concentrations of serum samples, such as the type of container used for serum collection, delays in sample processing, and repeated freeze-thaw cycles[288]. However, the impact of these parameters on the current study were minimise as serum samples were collected, processed, and stored according to a strict standard operations procedure at the Division of Surgery and Oncology, University of Liverpool. Despite this, the use of serum samples from the Liverpool Pancreatic Cancer database was not without disadvantages. In particular, PDAC serum samples were collected from individuals with resectable PDAC only and therefore it is unclear as to whether the disease-predicting algorithms could be generalised to include individuals with metastatic pancreatic cancer.

4.1.2.1 Limitation of sample size

The samples size of 180 was decided arbitrarily as there was no previous data, with which to base our sample size calculation on. However, in view of the results from the current study, we now estimate that the sample size that would be required to show a sensitivity of 0.9 and a specificity of 0.9 with 95% confidence (**W**) is approximately 280 patients[289].

Power calculation based on the lowest limit for sensitivity (**SN**) =0.9:

$$True\ Postive\ (TP) + False\ Negative(FN) = Z^2\ \times \frac{(SN \times (1 - SN))}{W^2}$$

$$TP + FN = 1.96^2\ \times \frac{(0.9 \times (1 - 0.9))}{0.05^2}$$

$$Sample\ Size = \frac{TP + FN}{1 - p} = \frac{138.2976}{0.5} = 276.5952 = approx.\ 280\ patients$$

Similarly, power calculation based on the lowest limit for specificity (**SP**) =0.9:

$$True\ Postive\ (TN) + False\ Negative(FP) = Z^2\ \times \frac{(SP \times (1 - SP))}{W^2}$$

$$TN + FP = 1.96^2\ \times \frac{(0.9 \times (1 - 0.9))}{0.05^2}$$

$$Sample\ Size = \frac{TN + FP}{1 - p} = \frac{138.2976}{0.5} = 276.5952 = approx.\ 280\ patients$$

***Note** "**Z**" is 1.96, which represents the standard deviation of the mean for 95% of sample population*

It is clear from the power calculation that the current thesis is significantly underpowered and that more samples for both the PDAC and control groups would be required to improve the strength of our study.

4.2 Future directions

The current thesis has identified a number of potential biomarker for pancreatic cancer including VDBP, RBP-4, FINC, IL-4, IL-17, G-CSF, and IP-10. However, there is a current lack of literature directly associating these proteins to pancreatic cancer. Therefore, we propose that future studies should aim to determine the roles of these proteins in pancreatic cancer.

In view of the limitations described in ***section 4.1.2***, we propose that future studies should aim to increase the sample size of the study and to analysis the diagnostic accuracy of the disease-predicting algorithms against cancers of other origins and other inflammatory diseases. Furthermore, the impact of sample collection methodology, delays in sample processing, and repeated freeze-thaw on the serum concentrations of candidate biomarkers should be assessed.

Moreover, the UKCTOCS samples (particularly samples within the 0-1 time category) should be quantified for serum CCGF and the resulting data used to validate the disease-predicting algorithms described in the current thesis and furthermore to identify potential pre-diagnostic CCGF markers for pancreatic cancer.

Finally, considering that we now have CCGF data for ninety PDAC serum samples, it would be interesting to conduct a retrospective study with the aim of identifying a serum CCGF prognostic marker for pancreatic cancer.

Chapter 5

References

References

1. Encyclopaedia Britannica. Anatomy of the Pancreas. [Image] [2nd July, 2010]. Available at: http://www.britannica.com/EBchecked/topic/440971/pancreas.

2. Bailey H, Love M. Surgical Anatomy. In: Mann CV, Russell RCG, Williams NS, editors. Bailey & Love's Short practice of surgery. 22nd / edited by Charles V. Mann, R.C.G. Russell, Norman S. Williams. ed. London: Chapman & Hall; 1995. p. 750-1.

3. Gray H, Lewis WH. Anatomy Pancreas. In: Gray H, editor. Anatomy of the human body. 20th / thoroughly rev. and re-edited by Warren H. Lewis. ed. Philadelphia: Lea & Febiger; 1918. p. 1199-204.

4. Corson JD, Williamson RCN, Bradley JA. Surgical Anatomy. In: Williamson RCN, Worthington TR, editors. Surgery. London: Mosby; 2001. p. 14.1-.3.

5. Office for National Statistics, (2008). *Cancer Statistics registrations: registration of cancer diagnosed in 2006*. MB1. No. 37. London: Office for National Statistics.

6. Office for National Statistics, (2008). *Mortality statistics: Deaths registered in 2007*. DR_07. London: Office of National Statistics.

7. Raimondi S, Maisonneuve P, Lowenfels AB. Epidemiology of pancreatic cancer: an overview. *Nat Rev Gastroenterol Hepatol*. 2009;6(12):699-708.

8. Iodice S, Gandini S, Maisonneuve P, Lowenfels AB. Tobacco and the risk of pancreatic cancer: a review and meta-analysis. *Langenbecks Arch Surg*. 2008;393(4):535-45.

9. Howes N, Lerch MM, Greenhalf W, Stocken DD, Ellis I, Simon P, et al. Clinical and genetic characteristics of hereditary pancreatitis in Europe. *Clin Gastroenterol Hepatol*. 2004;2(3):252-61.

10. Lowenfels AB, Maisonneuve P, Cavallini G, Ammann RW, Lankisch PG, Andersen JR, et al. Pancreatitis and the risk of pancreatic cancer. International Pancreatitis Study Group. *N Engl J Med*. 1993;328(20):1433-7.

11. Malka D, Hammel P, Maire F, Rufat P, Madeira I, Pessione F, et al. Risk of pancreatic adenocarcinoma in chronic pancreatitis. *Gut*. 2002;51(6):849-52.

12. Raimondi S, Lowenfels AB, Morselli-Labate AM, Maisonneuve P, Pezzilli R. Pancreatic cancer in chronic pancreatitis; aetiology, incidence, and early detection. *Best Pract Res Clin Gastroenterol*. 2010;24(3):349-58.

13. Berrington de Gonzalez A, Sweetland S, Spencer E. A meta-analysis of obesity and the risk of pancreatic cancer. *Br J Cancer*. 2003;89(3):519-23.

14. Fryzek JP, Schenk M, Kinnard M, Greenson JK, Garabrant DH. The association of body mass index and pancreatic cancer in residents of southeastern Michigan, 1996-1999. *Am J Epidemiol*. 2005;162(3):222-8.

15. Patel AV, Rodriguez C, Bernstein L, Chao A, Thun MJ, Calle EE. Obesity, recreational physical activity, and risk of pancreatic cancer in a large U.S. Cohort. *Cancer Epidemiol Biomarkers Prev*. 2005;14(2):459-66.

16. Khasawneh J, Schulz MD, Walch A, Rozman J, Hrabe de Angelis M, Klingenspor M, et al. Inflammation and mitochondrial fatty acid beta-oxidation link obesity to early tumor promotion. *Proc Natl Acad Sci U S A*. 2009;106(9):3354-9.

17. Hassan MM, Bondy ML, Wolff RA, Abbruzzese JL, Vauthey JN, Pisters PW, et al. Risk factors for pancreatic cancer: case-control study. *Am J Gastroenterol*. 2007;102(12):2696-707.

18. Huxley R, Ansary-Moghaddam A, Berrington de Gonzalez A, Barzi F, Woodward M. Type-II diabetes and pancreatic cancer: a meta-analysis of 36 studies. *Br J Cancer*. 2005;92(11):2076-83.

19. Luo J, Iwasaki M, Inoue M, Sasazuki S, Otani T, Ye W, et al. Body mass index, physical activity and the risk of pancreatic cancer in relation to smoking status and history of diabetes: a large-scale population-based cohort study in Japan--the JPHC study. *Cancer Causes Control*. 2007;18(6):603-12.

20. Chari ST, Leibson CL, Rabe KG, Timmons LJ, Ransom J, de Andrade M, et al. Pancreatic cancer-associated diabetes mellitus: prevalence and temporal association with diagnosis of cancer. *Gastroenterology*. 2008;134(1):95-101.

21. Hidalgo M. Pancreatic cancer. *N Engl J Med*. 2010;362(17):1605-17.

22. Chu D, Kohlmann W, Adler DG. Identification and screening of individuals at increased risk for pancreatic cancer with emphasis on known environmental and genetic factors and hereditary syndromes. *JOP*. 2010;11(3):203-12.

23. Giardiello FM, Brensinger JD, Tersmette AC, Goodman SN, Petersen GM, Booker SV, et al. Very high risk of cancer in familial Peutz-Jeghers syndrome. *Gastroenterology*. 2000;119(6):1447-53.

24. Ames BN, Gold LS, Willett WC. The causes and prevention of cancer. *Proc Natl Acad Sci U S A*. 1995;92(12):5258-65.

25. Wolpin BM, Chan AT, Hartge P, Chanock SJ, Kraft P, Hunter DJ, et al. ABO blood group and the risk of pancreatic cancer. *J Natl Cancer Inst*. 2009;101(6):424-31.

26. Ghaneh P, Costello E, Neoptolemos JP. Biology and management of pancreatic cancer. *Gut*. 2007;56(8):1134-52.

27. Hezel AF, Kimmelman AC, Stanger BZ, Bardeesy N, Depinho RA. Genetics and biology of pancreatic ductal adenocarcinoma. *Genes Dev*. 2006;20(10):1218-49.

28. Maitra A, Fukushima N, Takaori K, Hruban RH. Precursors to invasive pancreatic cancer. *Adv Anat Pathol*. 2005;12(2):81-91.

29. Sipos B, Frank S, Gress T, Hahn S, Kloppel G. Pancreatic intraepithelial neoplasia revisited and updated. *Pancreatology*. 2009;9(1-2):45-54.

30. Brugge WR, Lauwers GY, Sahani D, Fernandez-del Castillo C, Warshaw AL. Cystic neoplasms of the pancreas. *N Engl J Med*. 2004;351(12):1218-26.

31. Sommers SC, Murphy SA, Warren S. Pancreatic duct hyperplasia and cancer. *Gastroenterology*. 1954;27(5):629-40.

32. Hruban RH, Adsay NV, Albores-Saavedra J, Compton C, Garrett ES, Goodman SN, et al. Pancreatic intraepithelial neoplasia: a new nomenclature and classification system for pancreatic duct lesions. *Am J Surg Pathol*. 2001;25(5):579-86.

33. Longnecker DS, Adsay NV, Fernandez-del Castillo C, Hruban RH, Kasugai T, Klimstra DS, et al. Histopathological diagnosis of pancreatic intraepithelial neoplasia and intraductal papillary-mucinous neoplasms: interobserver agreement. *Pancreas*. 2005;31(4):344-9.

34. Roggin KK, Chennat J, Oto A, Noffsinger A, Briggs A, Matthews JB. Pancreatic cystic neoplasm. *Curr Probl Surg*. 2010;47(6):459-510.

35. Crippa S, Fernandez-Del Castillo C, Salvia R, Finkelstein D, Bassi C, Dominguez I, et al. Mucin-producing neoplasms of the pancreas: an analysis of distinguishing clinical and epidemiologic characteristics. *Clin Gastroenterol Hepatol*. 2010;8(2):213-9.

36. Le Borgne J, de Calan L, Partensky C. Cystadenomas and cystadenocarcinomas of the pancreas: a multiinstitutional retrospective study of 398 cases. French Surgical Association. *Ann Surg*. 1999;230(2):152-61.

37. Li T, Ji Y, Zhi XT, Wang L, Yang XR, Shi GM, et al. A comparison of hepatic mucinous cystic neoplasms with biliary intraductal papillary neoplasms. *Clin Gastroenterol Hepatol*. 2009;7(5):586-93.

38. Siech M, Tripp K, Schmidt-Rohlfing B, Mattfeldt T, Widmaier U, Gansauge F, et al. Cystic tumours of the pancreas: diagnostic accuracy, pathologic observations and surgical consequences. *Langenbecks Arch Surg*. 1998;383(1):56-61.

39. Sugiyama M, Atomi Y, Kuroda A. Two types of mucin-producing cystic tumors of the pancreas: diagnosis and treatment. *Surgery*. 1997;122(3):617-25.

40. Fasanella KE, McGrath K. Cystic lesions and intraductal neoplasms of the pancreas. *Best Pract Res Clin Gastroenterol*. 2009;23(1):35-48.

41. Adsay NV, Conlon KC, Zee SY, Brennan MF, Klimstra DS. Intraductal papillary-mucinous neoplasms of the pancreas: an analysis of in situ and invasive carcinomas in 28 patients. *Cancer*. 2002;94(1):62-77.

42. Augustin T, Vandermeer TJ. Intraductal papillary mucinous neoplasm: a clinicopathologic review. *Surg Clin North Am*. 2010;90(2):377-98.

43. Fan F, Lai EC, Xie F, Yang JM, Xu F, Kan T, et al. Intraductal papillary mucinous neoplasms of the pancreas--predictors of malignancy. *Hepatogastroenterology*. 2010;57(99-100):635-9.

44. Wasif N, Bentrem DJ, Farrell JJ, Ko CY, Hines OJ, Reber HA, et al. Invasive intraductal papillary mucinous neoplasm versus sporadic pancreatic adenocarcinoma: a stage-matched comparison of outcomes. *Cancer*. 2010;116(14):3369-77.

45. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell*. 2000;100(1):57-70.

46. Almoguera C, Shibata D, Forrester K, Martin J, Arnheim N, Perucho M. Most human carcinomas of the exocrine pancreas contain mutant c-K-ras genes. *Cell*. 1988;53(4):549-54.

47. Hruban RH, van Mansfeld AD, Offerhaus GJ, van Weering DH, Allison DC, Goodman SN, et al. K-ras oncogene activation in adenocarcinoma of the human pancreas. A study of 82 carcinomas using a combination of mutant-enriched polymerase chain reaction analysis and allele-specific oligonucleotide hybridization. *Am J Pathol*. 1993;143(2):545-54.

48. Santos E, Nebreda AR. Structural and functional properties of ras proteins. *FASEB J*. 1989;3(10):2151-63.

49. Thayer SP, di Magliano MP, Heiser PW, Nielsen CM, Roberts DJ, Lauwers GY, et al. Hedgehog is an early and late mediator of pancreatic cancer tumorigenesis. *Nature*. 2003;425(6960):851-6.

50. Strobel O, Rosow DE, Rakhlin EY, Lauwers GY, Trainor AG, Alsina J, et al. Pancreatic duct glands are distinct ductal compartments that react to chronic injury and mediate Shh-induced metaplasia. *Gastroenterology*. 2010;138(3):1166-77.

51. Miyamoto Y, Maitra A, Ghosh B, Zechner U, Argani P, Iacobuzio-Donahue CA, et al. Notch mediates TGF alpha-induced changes in epithelial differentiation during pancreatic tumorigenesis. *Cancer Cell*. 2003;3(6):565-76.

52. Lomberk G, Fernandez-Zapico ME, Urrutia R. When developmental signaling pathways go wrong and their impact on pancreatic cancer development. *Curr Opin Gastroenterol*. 2005;21(5):555-60.

53. Leach SD. Epithelial differentiation in pancreatic development and neoplasia: new niches for nestin and Notch. *J Clin Gastroenterol*. 2005;39(4 Suppl 2):S78-82.

54. Bailey H, Love M. Signs and symptoms. In: Mann CV, Russell RCG, Williams NS, editors. Bailey & Love's Short practice of surgery. 22nd / edited by Charles V. Mann, R.C.G. Russell, Norman S. Williams. ed. London: Chapman & Hall; 1995. p. 760.

55. Corson JD, Williamson RCN, Bradley JA. Signs and symptoms. In: Williamson RCN, Worthington TR, editors. Surgery. London: Mosby; 2001. p. 14.8.

56. Corson JD, Williamson RCN, Bradley JA. Investigation. In: Williamson RCN, Worthington TR, editors. Surgery. London: Mosby; 2001. p. 14.8-.9.

57. BSG. Guidelines for the management of patients with pancreatic cancer periampullary and ampullary carcinomas. *Gut*. 2005;54 Suppl 5:v1-16.

58. Minniti S, Bruno C, Biasiutti C, Tonel D, Falzone A, Falconi M, et al. Sonography versus helical CT in identification and staging of pancreatic ductal adenocarcinoma. *J Clin Ultrasound*. 2003;31(4):175-82.

59. Catalano C, Laghi A, Fraioli F, Pediconi F, Napoli A, Danti M, et al. Pancreatic carcinoma: the role of high-resolution multislice spiral CT in the diagnosis and assessment of resectability. *Eur Radiol*. 2003;13(1):149-56.

60. Helmstaedter L, Riemann JF. Pancreatic cancer--EUS and early diagnosis. *Langenbecks Arch Surg*. 2008;393(6):923-7.

61. UICC. TNM Classification of Malignant Tumours. 7th ed. Sobin LH, Gospodarowicz MK, Wittekind C, editors: Wiley-Blackwell; 2009.

62. Nakagohri T, Kinoshita T, Konishi M, Takahashi S, Gotohda N. Nodal involvement is strongest predictor of poor survival in patients with invasive adenocarcinoma of the head of the pancreas. *Hepatogastroenterology*. 2006;53(69):447-51.

63. Pawlik TM, Gleisner AL, Cameron JL, Winter JM, Assumpcao L, Lillemoe KD, et al. Prognostic relevance of lymph node ratio following pancreaticoduodenectomy for pancreatic cancer. *Surgery*. 2007;141(5):610-8.

64. Reber HA. Lymph node involvement as a prognostic factor in pancreatic cancer. *Int J Pancreatol*. 1990;7(1-3):125-7.

65. Riediger H, Keck T, Wellner U, zur Hausen A, Adam U, Hopt UT, et al. The lymph node ratio is the strongest prognostic factor after resection of pancreatic cancer. *J Gastrointest Surg*. 2009;13(7):1337-44.

66. Sierzega M, Popiela T, Kulig J, Nowak K. The ratio of metastatic/resected lymph nodes is an independent prognostic factor in patients with node-positive pancreatic head cancer. *Pancreas*. 2006;33(3):240-5.

67. Bailey H, Love M. Treatment. In: Mann CV, Russell RCG, Williams NS, editors. Bailey & Love's Short practice of surgery. 22nd / edited by Charles V. Mann, R.C.G. Russell, Norman S. Williams. ed. London: Chapman & Hall; 1995. p. 761-2.

68. Corson JD, Williamson RCN, Bradley JA. Treatment. In: Williamson RCN, Worthington TR, editors. Surgery. London: Mosby; 2001. p. 14.9-.21.

69. Ghaneh P, C M, Neoptolemos JP. Management of PDAC. In: Williams C, editor. London: BMJ Books; 2003. p. 247-72.

70. Rykowski JJ, Hilgier M. Efficacy of neurolytic celiac plexus block in varying locations of pancreatic cancer: influence on pain relief. *Anesthesiology*. 2000;92(2):347-54.

71. Suleyman Ozyalcin N, Talu GK, Camlica H, Erdine S. Efficacy of coeliac plexus and splanchnic nerve blockades in body and tail located pancreatic cancer pain. *Eur J Pain*. 2004;8(6):539-45.

72. Wong GY, Schroeder DR, Carns PE, Wilson JL, Martin DP, Kinney MO, et al. Effect of neurolytic celiac plexus block on pain relief, quality of life, and survival in patients with unresectable pancreatic cancer: a randomized controlled trial. *JAMA*. 2004;291(9):1092-9.

73. Bruno MJ, Haverkort EB, Tijssen GP, Tytgat GN, van Leeuwen DJ. Placebo controlled trial of enteric coated pancreatin microsphere treatment in patients with unresectable cancer of the pancreatic head region. *Gut*. 1998;42(1):92-6.

74. Burris HA, 3rd, Moore MJ, Andersen J, Green MR, Rothenberg ML, Modiano MR, et al. Improvements in survival and clinical benefit with gemcitabine as first-line therapy for patients with advanced pancreas cancer: a randomized trial. *J Clin Oncol*. 1997;15(6):2403-13.

75. Sultana A, Smith CT, Cunningham D, Starling N, Neoptolemos JP, Ghaneh P. Meta-analyses of chemotherapy for locally advanced and metastatic pancreatic cancer. *J Clin Oncol*. 2007;25(18):2607-15.

76. Smith DB, Neoptolemos JP. Capecitabine in carcinoma of the pancreas. *Expert Opin Pharmacother*. 2006;7(12):1633-9.

77. Hess V, Salzberg M, Borner M, Morant R, Roth AD, Ludwig C, et al. Combining capecitabine and gemcitabine in patients with advanced pancreatic carcinoma: a phase I/II trial. *J Clin Oncol*. 2003;21(1):66-8.

78. Hess V, Pratsch S, Potthast S, Lee L, Winterhalder R, Widmer L, et al. Combining gemcitabine, oxaliplatin and capecitabine (GEMOXEL) for patients with advanced pancreatic carcinoma (APC): a phase I/II trial. *Ann Oncol*. 2010.

79. Herrmann R, Bodoky G, Ruhstaller T, Glimelius B, Bajetta E, Schuller J, et al. Gemcitabine plus capecitabine compared with gemcitabine alone in advanced pancreatic cancer: a randomized, multicenter, phase III trial of the Swiss Group for Clinical Cancer Research and the Central European Cooperative Oncology Group. *J Clin Oncol*. 2007;25(16):2212-7.

80. Cunningham D, Chau I, Stocken DD, Valle JW, Smith D, Steward W, et al. Phase III randomized comparison of gemcitabine versus gemcitabine plus capecitabine in patients with advanced pancreatic cancer. *J Clin Oncol*. 2009;27(33):5513-8.

81. Fatima J, Schnelldorfer T, Barton J, Wood CM, Wiste HJ, Smyrk TC, et al. Pancreatoduodenectomy for ductal adenocarcinoma: implications of positive margin on survival. *Arch Surg*. 2010;145(2):167-72.

82. Campbell F, Smith RA, Whelan P, Sutton R, Raraty M, Neoptolemos JP, et al. Classification of R1 resections for pancreatic cancer: the prognostic relevance of tumour involvement within 1 mm of a resection margin. *Histopathology*. 2009;55(3):277-83.

83. Neoptolemos JP, Dunn JA, Stocken DD, Almond J, Link K, Beger H, et al. Adjuvant chemoradiotherapy and chemotherapy in resectable pancreatic cancer: a randomised controlled trial. *Lancet*. 2001;358(9293):1576-85.

84. Neoptolemos JP, Stocken DD, Friess H, Bassi C, Dunn JA, Hickey H, et al. A randomized trial of chemoradiotherapy and chemotherapy after resection of pancreatic cancer. *N Engl J Med*. 2004;350(12):1200-10.

85. Ghaneh P, Neoptolemos JP. Conclusions from the European Study Group for Pancreatic Cancer adjuvant trial of chemoradiotherapy and chemotherapy for pancreatic cancer. *Surg Oncol Clin N Am*. 2004;13(4):567-87, vii-viii.

86. Oettle H, Post S, Neuhaus P, Gellert K, Langrehr J, Ridwelski K, et al. Adjuvant chemotherapy with gemcitabine vs observation in patients undergoing curative-intent resection of pancreatic cancer: a randomized controlled trial. *JAMA*. 2007;297(3):267-77.

87. Chua YJ, Cunningham D. Adjuvant treatment for resectable pancreatic cancer. *J Clin Oncol*. 2005;23(20):4532-7.

88. Garcea G, Dennison AR, Pattenden CJ, Neal CP, Sutton CD, Berry DP. Survival following curative resection for pancreatic ductal adenocarcinoma. A systematic review of the literature. *JOP*. 2008;9(2):99-132.

89. Quiros RM, Brown KM, Hoffman JP. Neoadjuvant therapy in pancreatic cancer. *Cancer Invest*. 2007;25(4):267-73.

90. Neoptolemos JP, Buchler MW, Stocken DD, Ghaneh P, Smith CT, Bassi C, et al. ESPAC-3(v2): A multicenter, international, open-label, randomized, controlled phase III trial of adjuvant 5-fluorouracil/folinic acid (5-FU/FA) versus gemcitabine (GEM) in

patients with resected pancreatic ductal adenocarcinoma. *J Clin Oncol*. 2009;27(18s):suppl; abstr LBA4505.

91. BDWG. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther*. 2001;69(3):89-95.

92. Tonack S, Aspinall-O'Dea M, Neoptolemos JP, Costello E. Pancreatic cancer: proteomic approaches to a challenging disease. *Pancreatology*. 2009;9(5):567-76.

93. Koprowski H, Herlyn M, Steplewski Z, Sears HF. Specific antigen in serum of patients with colon carcinoma. *Science*. 1981;212(4490):53-5.

94. Tempero MA, Uchida E, Takasaki H, Burnett DA, Steplewski Z, Pour PM. Relationship of carbohydrate antigen 19-9 and Lewis antigens in pancreatic cancer. *Cancer Res*. 1987;47(20):5501-3.

95. Duffy MJ. CA 19-9 as a marker for gastrointestinal cancers: a review. *Ann Clin Biochem*. 1998;35 (Pt 3):364-70.

96. Duffy MJ, Sturgeon C, Lamerz R, Haglund C, Holubec VL, Klapdor R, et al. Tumor markers in pancreatic cancer: a European Group on Tumor Markers (EGTM) status report. *Ann Oncol*. 2010;21(3):441-7.

97. Goonetilleke KS, Siriwardena AK. Systematic review of carbohydrate antigen (CA 19-9) as a biochemical marker in the diagnosis of pancreatic cancer. *Eur J Surg Oncol*. 2007;33(3):266-70.

98. Lamerz R. Role of tumour markers, cytogenetics. *Ann Oncol*. 1999;10 Suppl 4:145-9.

99. Steinberg W. The clinical utility of the CA 19-9 tumor-associated antigen. *Am J Gastroenterol*. 1990;85(4):350-5.

100. Morris-Stiff G, Teli M, Jardine N, Puntis MC. CA19-9 antigen levels can distinguish between benign and malignant pancreaticobiliary disease. *Hepatobiliary Pancreat Dis Int*. 2009;8(6):620-6.

101. Duffy MJ, van Dalen A, Haglund C, Hansson L, Klapdor R, Lamerz R, et al. Clinical utility of biochemical markers in colorectal cancer: European Group on Tumour Markers (EGTM) guidelines. *Eur J Cancer*. 2003;39(6):718-27.

102. Sturgeon CM, Duffy MJ, Hofmann BR, Lamerz R, Fritsche HA, Gaarenstroom K, et al. National Academy of Clinical Biochemistry Laboratory Medicine Practice Guidelines for use of tumor markers in liver, bladder, cervical, and gastric cancers. *Clin Chem*. 2010;56(6):e1-48.

103. Reni M, Cereda S, Balzano G, Passoni P, Rognone A, Fugazza C, et al. Carbohydrate antigen 19-9 change during chemotherapy for advanced pancreatic adenocarcinoma. *Cancer*. 2009;115(12):2630-9.

104. Tsavaris N, Kosmas C, Papadoniou N, Kopteridis P, Tsigritis K, Dokou A, et al. CEA and CA-19.9 serum tumor markers as prognostic factors in patients with locally advanced (unresectable) or metastatic pancreatic adenocarcinoma: a retrospective analysis. *J Chemother*. 2009;21(6):673-80.

105. Waraya M, Yamashita K, Katagiri H, Ishii K, Takahashi Y, Furuta K, et al. Preoperative serum CA19-9 and dissected peripancreatic tissue margin as determiners of long-term survival in pancreatic cancer. *Ann Surg Oncol*. 2009;16(5):1231-40.

106. Shibata K, Iwaki K, Kai S, Ohta M, Kitano S. Increased levels of both carbohydrate antigen 19-9 and duke pancreatic monoclonal antigen type 2 reflect postoperative prognosis in patients with pancreatic carcinoma. *Pancreas*. 2009;38(6):619-24.

107. Ferrone CR, Finkelstein DM, Thayer SP, Muzikansky A, Fernandez-delCastillo C, Warshaw AL. Perioperative CA19-9 levels can predict stage and survival in patients with resectable pancreatic adenocarcinoma. *J Clin Oncol*. 2006;24(18):2897-902.

108. Duffy MJ. Role of tumor markers in patients with solid cancers: A critical review. *Eur J Intern Med*. 2007;18(3):175-84.

109. Grote T, Logsdon CD. Progress on molecular markers of pancreatic cancer. *Curr Opin Gastroenterol*. 2007;23(5):508-14.

110. Bussom S, Saif MW. Methods and rationale for the early detection of pancreatic cancer. Highlights from the "2010 ASCO Gastrointestinal Cancers Symposium". Orlando, FL, USA. January 22-24, 2010. *JOP*. 2010;11(2):128-30.

111. Joergensen MT, Brunner N, De Muckadell OB. Comparison of circulating MMP-9, TIMP-1 and CA19-9 in the detection of pancreatic cancer. *Anticancer Res*. 2010;30(2):587-92.

112. Takano S, Sogawa K, Yoshitomi H, Shida T, Mogushi K, Kimura F, et al. Increased circulating cell signalling phosphoproteins in sera are useful for the detection of pancreatic cancer. *Br J Cancer*. 2010;103(2):223-31.

113. Klose J. Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals. *Humangenetik*. 1975;26(3):231-43.

114. O'Farrell PH. High resolution two-dimensional electrophoresis of proteins. *J Biol Chem*. 1975;250(10):4007-21.

115. O'Farrell PZ, Goodman HM, O'Farrell PH. High resolution two-dimensional electrophoresis of basic as well as acidic proteins. *Cell*. 1977;12(4):1133-41.

116. Celis JE, Gromov P. 2D protein electrophoresis: can it be perfected? *Curr Opin Biotechnol*. 1999;10(1):16-21.

117. Gorg A, Weiss W, Dunn MJ. Current two-dimensional electrophoresis technology for proteomics. *Proteomics*. 2004;4(12):3665-85.

118. Weiss W, Gorg A. High-resolution two-dimensional electrophoresis. *Methods Mol Biol*. 2009;564:13-32.

119. Weiss W, Weiland F, Gorg A. Protein detection and quantitation technologies for gel-based proteome analysis. *Methods Mol Biol*. 2009;564:59-82.

120. Gorg A, Postel W, Gunther S. The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis*. 1988;9(9):531-46.

121. Bjellqvist B, Ek K, Righetti PG, Gianazza E, Gorg A, Westermeier R, et al. Isoelectric focusing in immobilized pH gradients: principle, methodology and some applications. *J Biochem Biophys Methods*. 1982;6(4):317-39.

122. Gorg A, Obermaier C, Boguth G, Harder A, Scheibe B, Wildgruber R, et al. The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis*. 2000;21(6):1037-53.

123. Drews O, Reil G, Parlar H, Gorg A. Setting up standards and a reference map for the alkaline proteome of the Gram-positive bacterium Lactococcus lactis. *Proteomics*. 2004;4(5):1293-304.

124. Gorg A, Obermaier C, Boguth G, Csordas A, Diaz JJ, Madjar JJ. Very alkaline immobilized pH gradients for two-dimensional electrophoresis of ribosomal and nuclear proteins. *Electrophoresis*. 1997;18(3-4):328-37.

125. Wildgruber R, Reil G, Drews O, Parlar H, Gorg A. Web-based two-dimensional database of Saccharomyces cerevisiae proteins using immobilized pH gradients from pH 6 to pH 12 and matrix-assisted laser desorption/ionization-time of flight mass spectrometry. *Proteomics*. 2002;2(6):727-32.

126. Viswanathan S, Unlu M, Minden JS. Two-dimensional difference gel electrophoresis. *Nat Protoc*. 2006;1(3):1351-8.

127. Shadforth IP, Dunkley TP, Lilley KS, Bessant C. i-Tracker: for quantitative proteomics using iTRAQ. *BMC Genomics*. 2005;6:145.

128. Wu WW, Wang G, Baek SJ, Shen RF. Comparative study of three proteomic quantitative methods, DIGE, cICAT, and iTRAQ, using 2D gel- or LC-MALDI TOF/TOF. *J Proteome Res*. 2006;5(3):651-8.

129. Tonack S, Aspinall-O'Dea M, Jenkins RE, Elliot V, Murray S, Lane CS, et al. A technically detailed and pragmatic protocol for quantitative serum proteomics using iTRAQ. *J Proteomics*. 2009;73(2):352-6.

130. Burnette WN. "Western blotting": electrophoretic transfer of proteins from sodium dodecyl sulfate--polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Anal Biochem*. 1981;112(2):195-203.

131. Mathews ST, Plaisance EP, Kim T. Imaging systems for westerns: chemiluminescence vs. infrared detection. *Methods Mol Biol*. 2009;536:499-513.

132. Bio-Rad, (2009). *Imaging of Chemiluminescent Western Blots: Comparison of Digital Imaging and X-ray Film, Rev A.* No. tech note 5809. Hercules, USA.

133. Engvall E, Perlmann P. Enzyme-linked immunosorbent assay (ELISA). Quantitative assay of immunoglobulin G. *Immunochemistry*. 1971;8(9):871-4.

134. Van Weemen BK, Schuurs AH. Immunoassay using antigen-enzyme conjugates. *FEBS Lett*. 1971;15(3):232-6.

135. Yalow RS, Berson SA. Immunoassay of endogenous plasma insulin in man. *J Clin Invest*. 1960;39:1157-75.

136. Chowdhury F, Williams A, Johnson P. Validation and comparison of two multiplex technologies, Luminex and Mesoscale Discovery, for human cytokine profiling. *J Immunol Methods*. 2009;340(1):55-64.

137. dupont NC, Wang K, Wadhwa PD, Culhane JF, Nelson EL. Validation and comparison of luminex multiplex cytokine analysis kits with ELISA: determinations of a panel of nine cytokines in clinical sample culture supernatants. *J Reprod Immunol*. 2005;66(2):175-91.

138. Madro A, Celinski K, Slomka M. The role of pancreatic stellate cells and cytokines in the development of chronic pancreatitis. *Med Sci Monit*. 2004;10(7):RA166-70.

139. Schmitz-Winnenthal H, Pietsch DH, Schimmack S, Bonertz A, Udonta F, Ge Y, et al. Chronic pancreatitis is associated with disease-specific regulatory T-cell responses. *Gastroenterology*. 2010;138(3):1178-88.

140. Bachem MG, Zhou Z, Zhou S, Siech M. Role of stellate cells in pancreatic fibrogenesis associated with acute and chronic pancreatitis. *J Gastroenterol Hepatol*. 2006;21 Suppl 3:S92-6.

141. Behrman SW, Fowler ES. Pathophysiology of chronic pancreatitis. *Surg Clin North Am*. 2007;87(6):1309-24, vii.

142. Talukdar R, Tandon RK. Pancreatic stellate cells: new target in the treatment of chronic pancreatitis. *J Gastroenterol Hepatol*. 2008;23(1):34-41.

143. Masamune A, Watanabe T, Kikuta K, Shimosegawa T. Roles of pancreatic stellate cells in pancreatic inflammation and fibrosis. *Clin Gastroenterol Hepatol*. 2009;7(11 Suppl):S48-54.

144. Jura N, Archer H, Bar-Sagi D. Chronic pancreatitis, pancreatic adenocarcinoma and the black box in-between. *Cell Res*. 2005;15(1):72-7.

145. Porta C, Larghi P, Rimoldi M, Totaro MG, Allavena P, Mantovani A, et al. Cellular and molecular pathways linking inflammation and cancer. *Immunobiology*. 2009;214(9-10):761-77.

146. Germano G, Allavena P, Mantovani A. Cytokines as a key component of cancer-related inflammation. *Cytokine*. 2008;43(3):374-9.

147. Korc M. Pancreatic cancer-associated stroma production. *Am J Surg*. 2007;194(4 Suppl):S84-6.

148. Garcea G, Dennison AR, Steward WP, Berry DP. Role of inflammation in pancreatic carcinogenesis and the implications for future therapy. *Pancreatology*. 2005;5(6):514-29.

149. Farrow B, Evers BM. Inflammation and the development of pancreatic cancer. *Surg Oncol*. 2002;10(4):153-69.

150. Colotta F, Allavena P, Sica A, Garlanda C, Mantovani A. Cancer-related inflammation, the seventh hallmark of cancer: links to genetic instability. *Carcinogenesis*. 2009;30(7):1073-81.

151. de Visser KE, Coussens LM. The inflammatory tumor microenvironment and its impact on cancer development. *Contrib Microbiol*. 2006;13:118-37.

152. Solinas G, Germano G, Mantovani A, Allavena P. Tumor-associated macrophages (TAM) as major players of the cancer-related inflammation. *J Leukoc Biol*. 2009;86(5):1065-73.

153. Vonlaufen A, Phillips PA, Xu Z, Goldstein D, Pirola RC, Wilson JS, et al. Pancreatic stellate cells and pancreatic cancer cells: an unholy alliance. *Cancer Res*. 2008;68(19):7707-10.

154. Bachem MG, Schunemann M, Ramadani M, Siech M, Beger H, Buck A, et al. Pancreatic carcinoma cells induce fibrosis by stimulating proliferation and matrix synthesis of stellate cells. *Gastroenterology*. 2005;128(4):907-21.

155. Mantovani A, Allavena P, Sica A, Balkwill F. Cancer-related inflammation. *Nature*. 2008;454(7203):436-44.

156. Schernhammer ES, Kang JH, Chan AT, Michaud DS, Skinner HG, Giovannucci E, et al. A prospective study of aspirin use and the risk of pancreatic cancer in women. *J Natl Cancer Inst*. 2004;96(1):22-8.

157. Chan AT, Ogino S, Fuchs CS. Aspirin and the risk of colorectal cancer in relation to the expression of COX-2. *N Engl J Med*. 2007;356(21):2131-42.

158. Flossmann E, Rothwell PM. Effect of aspirin on long-term risk of colorectal cancer: consistent evidence from randomised and observational studies. *Lancet*. 2007;369(9573):1603-13.

159. Koehne CH, Dubois RN. COX-2 inhibition and colorectal cancer. *Semin Oncol*. 2004;31(2 Suppl 7):12-21.

160. Westra WH, Baas IO, Hruban RH, Askin FB, Wilson K, Offerhaus GJ, et al. K-ras oncogene activation in atypical alveolar hyperplasias of the human lung. *Cancer Res*. 1996;56(9):2224-8.

161. Mantovani A. Molecular pathways linking inflammation and cancer. *Curr Mol Med*. 2010;10(4):369-73.

162. Karin M. NF-kappaB and cancer: mechanisms and targets. *Mol Carcinog*. 2006;45(6):355-61.

163. Prasad S, Ravindran J, Aggarwal BB. NF-kappaB and cancer: how intimate is this relationship. *Mol Cell Biochem*. 2010;336(1-2):25-37.

164. Foo SY, Nolan GP. NF-kappaB to the rescue: RELs, apoptosis and cellular transformation. *Trends Genet*. 1999;15(6):229-35.

165. Escarcega RO, Fuentes-Alexandro S, Garcia-Carrasco M, Gatica A, Zamora A. The transcription factor nuclear factor-kappa B and cancer. *Clin Oncol (R Coll Radiol)*. 2007;19(2):154-61.

166. Karin M. NF-kappaB as a critical link between inflammation and cancer. *Cold Spring Harb Perspect Biol*. 2009;1(5):a000141.

167. Bollrath J, Greten FR. IKK/NF-kappaB and STAT3 pathways: central signalling hubs in inflammation-mediated tumour promotion and metastasis. *EMBO Rep*. 2009;10(12):1314-9.

168. Rayet B, Gelinas C. Aberrant rel/nfkb genes and activity in human cancer. *Oncogene*. 1999;18(49):6938-47.

169. Yu H, Pardoll D, Jove R. STATs in cancer inflammation and immunity: a leading role for STAT3. *Nat Rev Cancer*. 2009;9(11):798-809.

170. Schindler C, Levy DE, Decker T. JAK-STAT signaling: from interferons to cytokines. *J Biol Chem*. 2007;282(28):20059-63.

171. Chen Z, O'Shea JJ. Th17 cells: a new fate for differentiating helper T cells. *Immunol Res*. 2008;41(2):87-102.

172. Heinrich PC, Behrmann I, Haan S, Hermanns HM, Muller-Newen G, Schaper F. Principles of interleukin (IL)-6-type cytokine signalling and its regulation. *Biochem J*. 2003;374(Pt 1):1-20.

173. Heinrich PC, Behrmann I, Muller-Newen G, Schaper F, Graeve L. Interleukin-6-type cytokine signalling through the gp130/Jak/STAT pathway. *Biochem J*. 1998;334 (Pt 2):297-314.

174. Yu H, Kortylewski M, Pardoll D. Crosstalk between cancer and immune cells: role of STAT3 in the tumour microenvironment. *Nat Rev Immunol*. 2007;7(1):41-51.

175. Aggarwal BB, Gehlot P. Inflammation and cancer: how friendly is the relationship for cancer patients? *Curr Opin Pharmacol*. 2009;9(4):351-69.

176. Catlett-Falcone R, Landowski TH, Oshiro MM, Turkson J, Levitzki A, Savino R, et al. Constitutive activation of Stat3 signaling confers resistance to apoptosis in human U266 myeloma cells. *Immunity*. 1999;10(1):105-15.

177. Darnell JE, Jr., Kerr IM, Stark GR. Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Science*. 1994;264(5164):1415-21.

178. Rebouissou S, Amessou M, Couchy G, Poussin K, Imbeaud S, Pilati C, et al. Frequent in-frame somatic deletions activate gp130 in inflammatory hepatocellular tumours. *Nature*. 2009;457(7226):200-4.

179. Lin WW, Karin M. A cytokine-mediated link between innate immunity, inflammation, and cancer. *J Clin Invest*. 2007;117(5):1175-83.

180. Chen J, Liu XS. Development and function of IL-10 IFN-gamma-secreting CD4(+) T cells. *J Leukoc Biol*. 2009;86(6):1305-10.

181. Lee H, Herrmann A, Deng JH, Kujawski M, Niu G, Li Z, et al. Persistently activated Stat3 maintains constitutive NF-kappaB activity in tumors. *Cancer Cell*. 2009;15(4):283-93.

182. Liao D, Johnson RS. Hypoxia: a key regulator of angiogenesis in cancer. *Cancer Metastasis Rev*. 2007;26(2):281-90.

183. Garcea G, Doucas H, Steward WP, Dennison AR, Berry DP. Hypoxia and angiogenesis in pancreatic cancer. *ANZ J Surg*. 2006;76(9):830-42.

184. Semenza GL. Defining the role of hypoxia-inducible factor 1 in cancer biology and therapeutics. *Oncogene*. 2010;29(5):625-34.

185. Bachem MG, Zhou S, Buck K, Schneiderhan W, Siech M. Pancreatic stellate cells--role in pancreas cancer. *Langenbecks Arch Surg*. 2008;393(6):891-900.

186. Kleeff J, Beckhove P, Esposito I, Herzig S, Huber PE, Lohr JM, et al. Pancreatic cancer microenvironment. *Int J Cancer*. 2007;121(4):699-705.

187. Lorusso G, Ruegg C. The tumor microenvironment and its contribution to tumor evolution toward metastasis. *Histochem Cell Biol*. 2008;130(6):1091-103.

188. Smyth MJ, Dunn GP, Schreiber RD. Cancer immunosurveillance and immunoediting: the roles of immunity in suppressing tumor development and shaping tumor immunogenicity. *Adv Immunol*. 2006;90:1-50.

189. Porta C, Subhra Kumar B, Larghi P, Rubino L, Mancino A, Sica A. Tumor promotion by tumor-associated macrophages. *Adv Exp Med Biol*. 2007;604:67-86.

190. Mantovani A, Sica A. Macrophages, innate immunity and cancer: balance, tolerance, and diversity. *Curr Opin Immunol*. 2010;22(2):231-7.

191. Mantovani A, Sica A, Allavena P, Garlanda C, Locati M. Tumor-associated macrophages and the related myeloid-derived suppressor cells as a paradigm of the diversity of macrophage activation. *Hum Immunol*. 2009;70(5):325-30.

192. Condeelis J, Pollard JW. Macrophages: obligate partners for tumor cell migration, invasion, and metastasis. *Cell*. 2006;124(2):263-6.

193. Pollard JW. Tumour-educated macrophages promote tumour progression and metastasis. *Nat Rev Cancer*. 2004;4(1):71-8.

194. Condeelis J, Segall JE. Intravital imaging of cell movement in tumours. *Nat Rev Cancer*. 2003;3(12):921-30.

195. Allavena P, Sica A, Solinas G, Porta C, Mantovani A. The inflammatory micro-environment in tumor progression: the role of tumor-associated macrophages. *Crit Rev Oncol Hematol*. 2008;66(1):1-9.

196. Allavena P, Garlanda C, Borrello MG, Sica A, Mantovani A. Pathways connecting inflammation and cancer. *Curr Opin Genet Dev*. 2008;18(1):3-10.

197. Masamune A, Shimosegawa T. Signal transduction in pancreatic stellate cells. *J Gastroenterol*. 2009;44(4):249-60.

198. Barkan D, Green JE, Chambers AF. Extracellular matrix: a gatekeeper in the transition from dormancy to metastatic growth. *Eur J Cancer*. 2010;46(7):1181-8.

199. Gaudard M, Ramsey P, Stephens M, Wright L. M-LR and NN Modelling methods. In: Gaudard M, editor. Visual Six Sigma: Making Data Analysis Lean. Cary: John Wiley Sons Inc; 2010. p. 391-484.

200. Hirschfeld J. Immune-electrophoretic demonstration of qualitative differences in human sera and their relation to the haptoglobins. *Acta Pathol Microbiol Scand*. 1959;47:160-8.

201. Daiger SP, Schanfield MS, Cavalli-Sforza LL. Group-specific component (Gc) proteins bind vitamin D and 25-hydroxyvitamin D. *Proc Natl Acad Sci U S A*. 1975;72(6):2076-80.

202. Yamamoto N. Structural definition of a potent macrophage activating factor derived from vitamin D3-binding protein with adjuvant activity for antibody production. *Mol Immunol*. 1996;33(15):1157-64.

203. Chishimba L, Thickett DR, Stockley RA, Wood AM. The vitamin D axis in the lung: a key role for vitamin D-binding protein. *Thorax*. 2010;65(5):456-62.

204. Gomme PT, Bertolini J. Therapeutic potential of vitamin D-binding protein. *Trends Biotechnol*. 2004;22(7):340-5.

205. White P, Cooke N. The multifunctional properties and characteristics of vitamin D-binding protein. *Trends Endocrinol Metab*. 2000;11(8):320-7.

206. Bertile F, Schaeffer C, Le Maho Y, Raclot T, Van Dorsselaer A. A proteomic approach to identify differentially expressed plasma proteins between the fed and prolonged fasted states. *Proteomics*. 2009;9(1):148-58.

207. Jorgensen CS, Christiansen M, Norgaard-Pedersen B, Ostergaard E, Schiodt FV, Laursen I, et al. Gc globulin (vitamin D-binding protein) levels: an inhibition ELISA assay for determination of the total concentration of Gc globulin in plasma and serum. *Scand J Clin Lab Invest*. 2004;64(2):157-66.

208. Gressner OA, Gao C, Siluschek M, Kim P, Gressner AM. Inverse association between serum concentrations of actin-free vitamin D-binding protein and the histopathological extent of fibrogenic liver disease or hepatocellular carcinoma. *Eur J Gastroenterol Hepatol*. 2009;21(9):990-5.

209. Schiodt FV. Gc-globulin in liver disease. *Dan Med Bull*. 2008;55(3):131-46.

210. Antoniades CG, Berry PA, Bruce M, Cross TJ, Portal AJ, Hussain MJ, et al. Actin-free Gc globulin: a rapidly assessed biomarker of organ dysfunction in acute liver failure and cirrhosis. *Liver Transpl*. 2007;13(9):1254-61.

211. Theodosiou M, Laudet V, Schubert M. From carrot to clinic: an overview of the retinoic acid signaling pathway. *Cell Mol Life Sci*. 2010;67(9):1423-45.

212. Fabris C, Piccoli A, Meani A, Farini R, Vianello D, Del Favero G, et al. Study of retinol-binding protein in pancreatic cancer. *J Cancer Res Clin Oncol*. 1984;108(2):227-9.

213. Goodman DS. Retinol-binding protein, prealbumin, and vitamin A transport. *Prog Clin Biol Res*. 1976;5:313-30.

214. Goodman DS. Plasma retinol-binding protein. *Ann N Y Acad Sci*. 1980;348:378-90.

215. Basu TK, Sasmal P. Plasma vitamin A, retinol-binding protein, and prealbumin in postoperative breast cancer patients. *Int J Vitam Nutr Res*. 1988;58(3):281-3.

216. Drozdz M, Gierek T, Jendryczko A, Piekarska J, Pilch J, Polanska D. Zinc, vitamins A and E, and retinol-binding protein in sera of patients with cancer of the larynx. *Neoplasma*. 1989;36(3):357-62.

217. Haines AP, Thompson SG, Basu TK, Hunt R. Cancer, retinol binding protein, zinc and copper. *Lancet*. 1982;1(8262):52-3.

218. Imamine T, Okuno M, Moriwaki H, Ninomiya M, Nishiwaki S, Noma A, et al. Plasma retinol transport system and taste acuity in patients with obstructive jaundice. *Gastroenterol Jpn*. 1990;25(2):206-11.

219. Milano G, Cooper EH, Goligher JC, Giles GR, Neville AM. Serum prealbumin, retinol-binding protein, transferrin, and albumin levels in patients with large bowel cancer. *J Natl Cancer Inst*. 1978;61(3):687-91.

220. Basu TK, Rowlands L, Jones L, Kohn J. Vitamin A and retinol-binding protein in patients with myelomatosis and cancer of epithelial origin. *Eur J Cancer Clin Oncol*. 1982;18(4):339-42.

221. Pearlstein E, Gold LI, Garcia-Pardo A. Fibronectin: a review of its structure and biological activity. *Mol Cell Biochem*. 1980;29(2):103-28.

222. Romberger DJ. Fibronectin. *Int J Biochem Cell Biol*. 1997;29(7):939-43.

223. Pankov R, Yamada KM. Fibronectin at a glance. *J Cell Sci*. 2002;115(Pt 20):3861-3.

224. University College London, (2008). *Protocol for the United Kingdom Collaborative Trial of Ovarian Cancer Screening (UKCTOCS)*. Contract No.: 22488978 London.

225. Yan L, Tonack S, Smith R, Dodd S, Jenkins RE, Kitteringham N, et al. Confounding effect of obstructive jaundice in the interpretation of proteomic plasma profiling data for pancreatic cancer. *J Proteome Res*. 2009;8(1):142-8.

226. Kalkunte S, Brard L, Granai CO, Swamy N. Inhibition of angiogenesis by vitamin D-binding protein: characterization of anti-endothelial activity of DBP-maf. *Angiogenesis*. 2005;8(4):349-60.

227. Kisker O, Onizuka S, Becker CM, Fannon M, Flynn E, D'Amato R, et al. Vitamin D binding protein-macrophage activating factor (DBP-maf) inhibits angiogenesis and tumor growth in mice. *Neoplasia*. 2003;5(1):32-40.

228. Yamamoto N, Suyama H, Nakazato H, Koga Y. Immunotherapy of metastatic colorectal cancer with vitamin D-binding protein-derived macrophage-activating factor, GcMAF. *Cancer Immunol Immunother*. 2008;57(7):1007-16.

229. Yamamoto N, Suyama H, Ushijima N. Immunotherapy of metastatic breast cancer patients with vitamin D-binding protein-derived macrophage activating factor (GcMAF). *Int J Cancer*. 2008;122(2):461-7.

230. Jing Y, Zhang J, Bleiweiss IJ, Waxman S, Zelent A, Mira YLR. Defective expression of cellular retinol binding protein type I and retinoic acid receptors alpha2, beta2, and gamma2 in human breast cancer cells. *FASEB J*. 1996;10(9):1064-70.

231. Kuppumbatti YS, Bleiweiss IJ, Mandeli JP, Waxman S, Mira YLR. Cellular retinol-binding protein expression and breast cancer. *J Natl Cancer Inst*. 2000;92(6):475-80.

232. Akbay E, Muslu N, Nayir E, Ozhan O, Kiykim A. SERUM RETINOL BINDING PROTEIN 4 LEVEL IS RELATED WITH RENAL FUNCTIONS IN TYPE 2 DIABETES. *J Endocrinol Invest*. 2010.

233. Yang Q, Graham TE, Mody N, Preitner F, Peroni OD, Zabolotny JM, et al. Serum retinol binding protein 4 contributes to insulin resistance in obesity and type 2 diabetes. *Nature*. 2005;436(7049):356-62.

234. Erikstrup C, Mortensen OH, Nielsen AR, Fischer CP, Plomgaard P, Petersen AM, et al. RBP-to-retinol ratio, but not total RBP, is elevated in patients with type 2 diabetes. *Diabetes Obes Metab*. 2009;11(3):204-12.

235. Graham TE, Yang Q, Bluher M, Hammarstedt A, Ciaraldi TP, Henry RR, et al. Retinol-binding protein 4 and insulin resistance in lean, obese, and diabetic subjects. *N Engl J Med*. 2006;354(24):2552-63.

236. Manolescu DC, Sima A, Bhat PV. All-trans retinoic acid lowers serum retinol-binding protein 4 concentrations and increases insulin sensitivity in diabetic mice. *J Nutr*. 2010;140(2):311-6.

237. Wolf G. Serum retinol-binding protein: a link between obesity, insulin resistance, and type 2 diabetes. *Nutr Rev*. 2007;65(5):251-6.

238. Mahadevan D, Von Hoff DD. Tumor-stroma interactions in pancreatic ductal adenocarcinoma. *Mol Cancer Ther*. 2007;6(4):1186-97.

239. Ryschich E, Khamidjanov A, Kerkadze V, Buchler MW, Zoller M, Schmidt J. Promotion of tumor cell migration by extracellular matrix proteins in human pancreatic cancer. *Pancreas*. 2009;38(7):804-10.

240. Akiyama SK, Olden K, Yamada KM. Fibronectin and integrins in invasion and metastasis. *Cancer Metastasis Rev*. 1995;14(3):173-89.

241. Erkan M, Kleeff J, Gorbachevski A, Reiser C, Mitkus T, Esposito I, et al. Periostin creates a tumor-supportive microenvironment in the pancreas by sustaining fibrogenic stellate cell activity. *Gastroenterology*. 2007;132(4):1447-64.

242. Erkan M, Reiser-Erkan C, Michalski CW, Deucker S, Sauliunaite D, Streit S, et al. Cancer-stellate cell interactions perpetuate the hypoxia-fibrosis cycle in pancreatic ductal adenocarcinoma. *Neoplasia*. 2009;11(5):497-508.

243. Grivennikov SI, Greten FR, Karin M. Immunity, inflammation, and cancer. *Cell*. 2010;140(6):883-99.

244. Ben-Baruch A. Inflammation-associated immune suppression in cancer: the roles played by cytokines, chemokines and additional mediators. *Semin Cancer Biol*. 2006;16(1):38-52.

245. Lopez-Novoa JM, Nieto MA. Inflammation and EMT: an alliance towards organ fibrosis and cancer progression. *EMBO Mol Med*. 2009;1(6-7):303-14.

246. Lu H, Ouyang W, Huang C. Inflammation, a key event in cancer development. *Mol Cancer Res*. 2006;4(4):221-33.

247. Solinas G, Marchesi F, Garlanda C, Mantovani A, Allavena P. Inflammation-mediated promotion of invasion and metastasis. *Cancer Metastasis Rev*. 2010;29(2):243-8.

248. Borrello MG, Degl'Innocenti D, Pierotti MA. Inflammation and cancer: the oncogene-driven connection. *Cancer Lett*. 2008;267(2):262-70.

249. Rollins BJ. Inflammatory chemokines in cancer growth and progression. *Eur J Cancer*. 2006;42(6):760-7.

250. Schetter AJ, Heegaard NH, Harris CC. Inflammation and cancer: interweaving microRNA, free radical, cytokine and p53 pathways. *Carcinogenesis*. 2010;31(1):37-49.

251. Li Z, Chen L, Qin Z. Paradoxical roles of IL-4 in tumor immunity. *Cell Mol Immunol*. 2009;6(6):415-22.

252. Pakala SV, Kurrer MO, Katz JD. T helper 2 (Th2) T cells induce acute pancreatitis and diabetes in immune-compromised nonobese diabetic (NOD) mice. *J Exp Med*. 1997;186(2):299-306.

253. Hill NJ, Van Gunst K, Sarvetnick N. Th1 and Th2 pancreatic inflammation differentially affects homing of islet-reactive CD4 cells in nonobese diabetic mice. *J Immunol*. 2003;170(4):1649-58.

254. Seicean A, Popa D, Mocan T, Cristea V, Neagoe I. Th1 and Th2 profiles in patients with pancreatic cancer compared with chronic pancreatitis. *Pancreas*. 2009;38(5):594-5.

255. Golumbek PT, Lazenby AJ, Levitsky HI, Jaffee LM, Karasuyama H, Baker M, et al. Treatment of established renal cancer by tumor cells engineered to secrete interleukin-4. *Science*. 1991;254(5032):713-6.

256. Pericle F, Giovarelli M, Colombo MP, Ferrari G, Musiani P, Modesti A, et al. An efficient Th2-type memory follows CD8+ lymphocyte-driven and eosinophil-mediated rejection of a spontaneous mouse mammary adenocarcinoma engineered to release IL-4. *J Immunol*. 1994;153(12):5659-73.

257. Pippin BA, Rosenstein M, Jacob WF, Chiang Y, Lotze MT. Local IL-4 delivery enhances immune reactivity to murine tumors: gene therapy in combination with IL-2. *Cancer Gene Ther*. 1994;1(1):35-42.

258. Rodolfo M, Zilocchi C, Accornero P, Cappetti B, Arioli I, Colombo MP. IL-4-transduced tumor cell vaccine induces immunoregulatory type 2 CD8 T lymphocytes that cure lung metastases upon adoptive transfer. *J Immunol*. 1999;163(4):1923-8.

259. Stoppacciaro A, Paglia P, Lombardi L, Parmiani G, Baroni C, Colombo MP. Genetic modification of a carcinoma with the IL-4 gene increases the influx of dendritic cells relative to other cytokines. *Eur J Immunol*. 1997;27(9):2375-82.

260. Atkins MB, Vachino G, Tilg HJ, Karp DD, Robert NJ, Kappler K, et al. Phase I evaluation of thrice-daily intravenous bolus interleukin-4 in patients with refractory malignancy. *J Clin Oncol*. 1992;10(11):1802-9.

261. Gilleece MH, Scarffe JH, Ghosh A, Heyworth CM, Bonnem E, Testa N, et al. Recombinant human interleukin 4 (IL-4) given as daily subcutaneous injections--a phase I dose toxicity trial. *Br J Cancer*. 1992;66(1):204-10.

262. Stadler WM, Rybak ME, Vogelzang NJ. A phase II study of subcutaneous recombinant human interleukin-4 in metastatic renal cell carcinoma. *Cancer*. 1995;76(9):1629-33.

263. Shurin MR, Lu L, Kalinski P, Stewart-Akers AM, Lotze MT. Th1/Th2 balance in cancer, transplantation and pregnancy. *Springer Semin Immunopathol*. 1999;21(3):339-59.

264. Onishi T, Ohishi Y, Imagawa K, Ohmoto Y, Murata K. An assessment of the immunological environment based on intratumoral cytokine production in renal cell carcinoma. *BJU Int*. 1999;83(4):488-92.

265. Gocheva V, Wang HW, Gadea BB, Shree T, Hunter KE, Garfall AL, et al. IL-4 induces cathepsin protease activity in tumor-associated macrophages to promote cancer growth and invasion. *Genes Dev*. 2010;24(3):241-55.

266. Xu S, Cao X. Interleukin-17 and its expanding biological functions. *Cell Mol Immunol*. 2010;7(3):164-74.

267. Numasaki M, Fukushi J, Ono M, Narula SK, Zavodny PJ, Kudo T, et al. Interleukin-17 promotes angiogenesis and tumor growth. *Blood*. 2003;101(7):2620-7.

268. He D, Li H, Yusuf N, Elmets CA, Li J, Mountz JD, et al. IL-17 promotes tumor development through the induction of tumor promoting microenvironments at tumor sites and myeloid-derived suppressor cells. *J Immunol*. 2010;184(5):2281-8.

269. Metcalf D. The colony-stimulating factors and cancer. *Nat Rev Cancer*. 2010;10(6):425-34.

270. Hill CP, Osslund TD, Eisenberg D. The structure of granulocyte-colony-stimulating factor and its relationship to other growth factors. *Proc Natl Acad Sci U S A*. 1993;90(11):5167-71.

271. Booth V, Keizer DW, Kamphuis MB, Clark-Lewis I, Sykes BD. The CXCR3 binding chemokine IP-10/CXCL10: structure and receptor interactions. *Biochemistry*. 2002;41(33):10418-25.

272. Strieter RM, Burdick MD, Mestas J, Gomperts B, Keane MP, Belperio JA. Cancer CXC chemokine networks and tumour angiogenesis. *Eur J Cancer*. 2006;42(6):768-78.

273. Musha H, Ohtani H, Mizoi T, Kinouchi M, Nakayama T, Shiiba K, et al. Selective infiltration of CCR5(+)CXCR3(+) T lymphocytes in human colorectal carcinoma. *Int J Cancer*. 2005;116(6):949-56.

274. Yang J, Richmond A. The angiostatic activity of interferon-inducible protein-10/CXCL10 in human melanoma depends on binding to CXCR3 but not to glycosaminoglycan. *Mol Ther*. 2004;9(6):846-55.

275. Arenberg DA, Kunkel SL, Polverini PJ, Morris SB, Burdick MD, Glass MC, et al. Interferon-gamma-inducible protein 10 (IP-10) is an angiostatic factor that inhibits human non-small cell lung cancer (NSCLC) tumorigenesis and spontaneous metastases. *J Exp Med*. 1996;184(3):981-92.

276. Chang ST, Zahn JM, Horecka J, Kunz PL, Ford JM, Fisher GA, et al. Identification of a biomarker panel using a multiplex proximity ligation assay improves accuracy of pancreatic cancer diagnosis. *J Transl Med*. 2009;7:105.

277. Firpo MA, Gay DZ, Granger SR, Scaife CL, DiSario JA, Boucher KM, et al. Improved diagnosis of pancreatic adenocarcinoma using haptoglobin and serum amyloid A in a panel screen. *World J Surg*. 2009;33(4):716-22.

278. Joergensen MT, Heegaard NH, Schaffalitzky de Muckadell OB. Comparison of plasma Tu-M2-PK and CA19-9 in pancreatic cancer. *Pancreas*. 2010;39(2):243-7.

279. Kim YC, Kim HJ, Park JH, Park DI, Cho YK, Sohn CI, et al. Can preoperative CA19-9 and CEA levels predict the resectability of patients with pancreatic adenocarcinoma? *J Gastroenterol Hepatol*. 2009;24(12):1869-75.

280. Malaguarnera M, Cristaldi E, Cammalleri L, Colonna V, Lipari H, Capici A, et al. Elevated chromogranin A (CgA) serum levels in the patients with advanced pancreatic cancer. *Arch Gerontol Geriatr*. 2009;48(2):213-7.

281. Marten A, Buchler MW, Werft W, Wente MN, Kirschfink M, Schmidt J. Soluble iC3b as an early marker for pancreatic adenocarcinoma is superior to CA19.9 and radiology. *J Immunother*. 2010;33(2):219-24.

282. Szajda SD, Snarska J, Jankowska A, Puchalski Z, Zwierz K. Isoenzymes A and B of N-acetyl-beta-D-hexosaminidase in serum and urine of patients with pancreatic cancer. *Hepatogastroenterology*. 2008;55(82-83):695-8.

283. Takayama R, Nakagawa H, Sawaki A, Mizuno N, Kawai H, Tajika M, et al. Serum tumor antigen REG4 as a diagnostic biomarker in pancreatic ductal adenocarcinoma. *J Gastroenterol*. 2010;45(1):52-9.

284. Bhattacharyya S, Siegel ER, Petersen GM, Chari ST, Suva LJ, Haun RS. Diagnosis of pancreatic cancer using serum proteomic profiling. *Neoplasia*. 2004;6(5):674-86.

285. Koopmann J, Rosenzweig CN, Zhang Z, Canto MI, Brown DA, Hunter M, et al. Serum markers in patients with resectable pancreatic adenocarcinoma: macrophage inhibitory cytokine 1 versus CA19-9. *Clin Cancer Res*. 2006;12(2):442-6.

286. Zhang L, Farrell JJ, Zhou H, Elashoff D, Akin D, Park NH, et al. Salivary transcriptomic biomarkers for detection of resectable pancreatic cancer. *Gastroenterology*. 2010;138(3):949-57 e1-7.

287. Kim BK, Lee JW, Park PJ, Shin YS, Lee WY, Lee KA, et al. The multiplex bead array approach to identifying serum biomarkers associated with breast cancer. *Breast Cancer Res*. 2009;11(2):R22.

288. Schrohl AS, Wurtz S, Kohn E, Banks RE, Nielsen HJ, Sweep FC, et al. Banking of biological fluids for studies of disease-associated protein biomarkers. *Mol Cell Proteomics*. 2008;7(10):2061-6.

289. Jones SR, Carley S, Harrison M. An introduction to power and sample size estimation. *Emerg Med J*. 2003;20(5):453-8.