NMR-Based Metabonomic Studies of Human Pancreatic Cancer

Rowland Lewis Storey

Submitted in accordance with the requirements for the degree of MD

The University of Leeds
Faculty of Medicine & Health

April 2016

I (Rowland Storey) confirm that this work is my own and that appropriate credit has been given where reference has been made to the work of others.

Acknowledgements

Many people have been invaluable in facilitating my project at various points. I am grateful to Mr Andrew Smith for taking me on in the first instance and holding my position open following an un-planned years return to clinical training half way through my MD. I am most grateful to Dr Julie Fisher who has been a true inspiration during my study in Leeds. Julie was always available to support my progress, encourage me to move on to the next step and to complete my thesis write-up. We were all greatly saddened to hear of Julie's early death following a difficult period of short illness. Along with many others I will always remember Julie as someone who remained so humble despite immense personal and professional achievement. She took great interest not only in our research but also each and every person she happened to work with irrespective of professional status. It was a pleasure and honour to work under her guidance. Julie will very much be missed but by no means forgotten.

I am grateful to Professor Jayne who acted as co-supervisor. It was an experience to attend Professor Jayne's research meetings and to meet his many research fellows along with Sarah Perry and Dr Tom Hughes.

My project would of course not have been possible without the kind donation of samples provided by the patients under the care of Mr Andrew Smith, Mr Amer Aldouri and Mr Christian Macutkiewicz.

Finally but by no means least I wish to also thank my parents who have always been there when I needed them most.

Abstract

Pancreatic cancer is one of the most fatal malignancies in the world with an overall 5-year survival rate of less than 5%. Diagnosing pancreatic cancer is not always straightforward. Clinical imaging of the pancreas can be misleading and currently available clinical biomarkers are few and lack sensitivity and specificity. Obtaining tissue or cytology from the pancreaticobiliary system to confirm a diagnosis is clinically invasive and may be inconclusive. Further non-invasive diagnostic biomarkers are clearly required. I describe the feasibility of nuclear magnetic resonance (NMR) spectroscopy as a modality for novel plasma and urine pancreaticobiliary biomarker discovery

Plasma and urine samples from 44 patients undergoing pancreatic resection for pancreaticobiliary malignancy along with a benign cohort of 45 patients were acquired. Spectra were obtained on a Varian NMR 500 MHz spectrometer. Unsupervised and supervised multivariate pattern recognition techniques were used for chemometric analysis. Model validation was assessed through permutation and cross validation techniques

Plasma metabonomic profiling identified clear separation between malignant and benign pancreaticobiliary disease with an overall sensitivity and specificity of 64.9 and 73.5% respectively. Sensitivity and specificity among non-jaundiced patients rose to 75 and 75.8% respectively. Suppressed metabolites among cancer patients included VLDL, valine and acetate. Up-regulated metabolites included isobutyrate, 3-hydroxybutyrate, lactate, acetoacetate, pyruvate, glucose and taurine. Urinary metabonomic profiling failed to satisfactorily discriminate between benign and malignant disease.

Plasma nuclear magnetic resonance metabonomic profiling has significant potential for future pancreaticobiliary biomarker development. Plasma bilirubin is an important confounding factor, which must be accounted for in all future pancreaticobiliry metabonomic studies

Contents

Figures

Tables

Abbreviations

AIP	Autoimmune pancreatitis
AMP	Adenosine monophosphate
BB	Benign biliary
Bil	Bilirubin
BMI	Body mass index
BMP	Bone morphogenic protein
B_0	External magnetic field
BR	Benign plasma
BU	Benign urine
CA 19-9	Carbohydrate antigen 19-9
CAPS	International cancer of the pancreas screening
CBD	Common bile duct
CEA	Carcinoembryonic antigen
CEACAM1	Carcinoembryonic antigen-related cell adhesion molecule 1
COPD	Chronic obstructive airways disease
CP	Chronic pancreatitis
CPMG	Car-Purcell-Meiboom-Gill
CR	Cancer plasma
crExos	Circulating cancer cell derived exosomes
CT	Computed tomography
CU	Cancer urine
DFS	Disease free survival
ERCP	Endoscopic retrograde cholangiopancreatography
EUS	Endoscopic ultrasound

FAMMM	Familial atypical multiple mole melanoma
FID	Free induction decay
FPC	Familial pancreatic cancer
GB	Gallbladder
GIST	Gastrointestinal stromal tumour
Gluc	Glucose
GPCI	Cell surface proteoglycan glypican-1
HDL	High density lipoprotein
HMP	Human Metabolome Project
HNOCC	Hereditary non-polyposis colorectal cancer
HTN	Hypertension
HV	Healthy volunteers
IHD	Ischaemic heart disease
IPMN	Intraductal papillary mucinous neoplasm
KRAS	Kirsten rat sarcoma viral oncogene homolog
Lact	Lactate
Lap chole	Laparoscopic cholecystectomy
LDL	Low density lipoprotein
LDP	Laparoscopic distal pancreatectomy
MCN	Mucinous cystic neoplasm
MIC	Macrophage inhibitory cytokine
miRNA	Micro ribonucleic acid
MS	Mass spectrometry
NAD	Nicotinamide adenine dinucleotide
NMR	Nuclear magnetic resonance spectroscopy
NOESY	Nuclear Overhauser effect spectroscopy

ODP	Open distal pancreatectomy
Open chole	Open cholecystectomy
OPLS	Orthogonal PLS
OS	Overall survival
Pan	Pancreatic intraepithelial neoplasia
PB	Pancreaticobiliary
PC	Pancreatic cancer
PCA	Principle component analysis
PDA	Pancreatic ductal adenocarcinoma
PET CT	Positron emission CT
PJ	Peutz-Jager's syndrome
PLS	Partial least squares analysis
PLS-DA	PLS-disciminant analysis
PMH	Past medical history
PPM	Parts per million
PPPD	Pylorus preserving pancreaticoduodenectomy
REC	Research Ethics Committee
rf	Radiofrequency
rpm	Revolutions per minute
SCN	Serous cystic neoplasm
splen	Splenectomy
TGF-β	Transforming growth factor beta
TA USS	Trans-abdominal ultrasound scan
TP	Total pancreatectomy
VLDL	Very low-density lipoprotein
V_0	Lamor frequency

1 Introduction

1.1 Pancreatic cancer

1.1.1 Disease burden on society

Pancreatic cancer (PC) is one of the most fatal malignancies in the world. The overall 5-year survival rate of less than 5% is the lowest of the 21 most-common malignancies in the UK. Around 70,000 new cases of PC are diagnosed each year in the European Union with around 8000 occurring in the UK. PC is the fourth most common cause of cancer related death in men (after lung, prostate and colorectal) and women (after lung, breast and colorectal) within the UK (1). The lifetime risk of developing PC is 1 in 77 for men and 1 in 79 for women (2). The incidence of PC increases with age with around three-quarters of cases occurring in the over 65 (3). The poor prognosis is due to several factors including delayed presentation, high metastatic potential and resistance to chemoradiotherapy. The early presenting features of PC are difficult to recognize and are vague and non-specific. As such 80% of cases are advanced and non-operable at presentation (4). Although surgical resection offers the only chance for long-term survival patients rarely survive beyond five-years. The median survival following surgery is 11-20 months with five-year survival ranging from 7-25% (5, 6). Patients with irresectable locally advanced disease have a median survival of 6-11 months. Patients who have metastatic disease have a median survival of only 2 - 6 months (7, 8).

1.1.2 Pancreatic cancer current diagnostic limitations

Following the clinical suspicion of PC, imaging modalities such as trans-abdominal ultrasound (TA USS), endoscopic ultrasound (EUS), computed tomography (CT) or endoscopic retrograde cholangiopancreatography (ERCP) may be deployed. TA USS is often the first investigation in a jaundiced patient in the UK. Common bile duct dilation (> 7mm, 10mm in patients post cholecystectomy) together with pancreatic duct dilation (> 2mm) is an indirect sign of a pancreatic head lesion (double-duct sign). TA USS is however operator dependent and limited by overlying bowel gas and

patient habitus (9). CT is the most widely available and best-validated imaging modality for diagnosing and staging patients with pancreatic cancer. A pancreas CT protocol involves triphasic (arterial phase, late arterial phase, and venous phase) cross-sectional imaging. The triphasic CT protocol allows for selective visualization of important arterial (celiac axis, superior mesenteric and peri-pancreatic arteries) and venous structures (superior mesenteric, splenic, and portal vein), thereby providing an assessment of vascular invasion by the tumour (10). Positron emission computered tomography (PET/CT) may be used to assess for metastatic disease. In a retrospective study the sensitivity of detecting metastatic disease for PET/CT alone, standard CT alone, and the combination of PET/CT and standard CT was 61%, 57%, and 87%, respectively (11). EUS was introduced in the 1980's to overcome difficulties in visualization of the pancreas on TA USS. EUS may also be used to guide tissue collection for cytological or histological analysis (12). The diagnostic accuracy of fine needle aspiration at EUS is around 71%. The diagnostic yield of cytology from common bile duct brushings at ERCP lies between 23 and 41% (13, 14).

Although pancreatic ductal adenocarcinoma (PDA) accounts for 90% of exocrine PC additional cell lines include adenosquamous, colloid, hepatoid, medullary, signet ring and undifferentiated carcinoma. Identifying these cell lines pre-operatively is difficult but potentially advantageous due to their varying clinical characteristics (prognosis and chemosensitivity). Colloid carcinomas often arise within type 1 intra-ductal papillary-mucinous neoplasms and have an improved five year survival of 57% following resection (15, 16). The prognosis from adenosquamous, hepatoid, medullary and undifferentiated tumors is worse than that of PDA (17-21).

1.1.3 Chemotherapy for pancreatic cancer

PC is highly desmoplastic with prominent stroma (22). The stromal compartment plays an active role in promoting invasion and growth of PC and is a physical barrier for drug delivery and chemoresistance (23). As such adjuvant chemotherapy following surgery is associated with only a modest

improvement in overall survival. In the European Study Group for Pancreatic Cancer-1 (ESPAC-1) trial adjuvant chemotherapy improved survival from a median of 14 with observation alone to only 19.7 months ($p = 0.0005$) (24). It has been suggested that neoadjuvant therapy has no added advantage in terms of surgical resectability or survival among patients presenting with resectable tumours (25). Advantages of neoadjuvant therapy include facilitating negative margin resection among those with borderline resectable disease. Neoadjuvant therapy may in addition highlight aggressive tumour types (with disease progression during chemotherapy) for which surgery is avoided (25, 26). For metastatic disease gemcitabine has historically been the agent of choice. In 2010 the phase II/III PRODIGE 4/ACCORD 11 trial was instrumental in changing perceptions towards palliative chemotherapy. Patients were randomly assigned to FOLFIRINOX (oxaliplatin, leucovorin, irinotecan and fluorouracil bolus followed by infusional fluorouracil) or gemcitabine alone. Median overall survival was substantially longer on the FOLFIRINOX arm (11.1 month vs. 6.8 months; $p < 0.001$) (27).

1.2 Precursors to pancreatic cancer

PDA is believed to arise from precursor lesions that develop into invasive carcinoma through a multistep carcinogenic process. This process is known to take at least 10-years (28). This provides a large window of opportunity for screening among the high-risk population. In a study by Biankin *et al*. an average of 26 mutations per-patient were identified among a cohort of 142 PC patients who collectively displayed 2,016 non-silent mutations (29).

The most common pre-neoplastic lesion seen histologically among 80% of patients with PDA is pancreatic intraepithelial neoplasia (PanIN) (30). PanIN may arise within regions of acinar-to-ductal metaplasia (31). PanIN may be classified as PanIN1A, PanIN1B, PanIN2 and PanIN3 depending on the grade of architectural and nuclear atypia (32). PanIN3 lesions represent carcinoma in-situ and may harbour mitotic figures and exhibit local invasion.

1.3 Pancreatic cancer screening & the high risk population

There is no current evidence to support screening the population as a whole as the overall risks (in terms of diagnostic intervention) would out-weigh the potential benefits (33). Screening is however of potential benefit to a higher risk population with a lifetime risk of over 5% and/or relative risk greater than five-fold of developing PC (34). The international cancer of the pancreas screening (CAPS) group has defined ‘‘successful screening” as the detection and treatment of T1N0M0 margin negative pancreatic cancer and high-grade dysplastic precursor lesions including pancreatic intraepithelial neoplasia-3 (PanIN-3), intraductal papillary mucinous neoplasm (IPMN) with high-grade dysplasia, and mucinous cystic neoplasm (MCN) with high-grade dysplasia (34).

1.3.1 Risk factors for pancreatic cancer

1.3.1.1 Generic risk

Smoking is estimated to be the cause of 25-30% of PC cases in the UK (35). Current cigarette smokers and former smokers who have quit for less than 5-years have a higher risk of pancreatic cancer than non-smokers (odds ratio: 1:71 and 1:78 respectively) (36). Overweight and obese individuals have an increased risk (odds ratio: 1.8 and 1.22 for males and females, respectively) (37). Patients with diabetes are at higher risk for PC (odds ratio: 1:76) (38) . New onset of diabetes may also be an early indicator of PC (39). Diets rich in red meat and dairy are associated with an increased cancer risk (40) along with exposure to ionizing radiation, insecticides, nickel, acrylamide, halogenated hydrocarbons and chlorinated hydrocarbon solvents (41-43). Chronic pancreatitis is associated with a 7.2 fold increased risk (44).

1.3.1.2 Genetic risk

It is estimated that 10% of patients suffering from PC have a hereditary component to their disease (45, 46). Peutz-Jager’s (PJ), familial atypical multiple mole melanoma (FAMMM), familial breast and ovarian cancer syndrome, hereditary non-polyposis colorectal cancer (HNPCC), hereditary

pancreatitis and familial pancreatic cancer (FPC) are all syndromes which are known to predispose to PC. PJ is an autosomal dominant syndrome characterised by a STK11 gene mutation (47). The cumulative lifetime risk of PC among these patients is 36% (48). FAMMM is an autosomal dominant syndrome with variable penetrance developing following p16/CDKN2A gene mutation (49). Patients with FAMMM suffer from multiple benign melanocytic nevi, dysplastic nevi and melanoma (50). The lifetime risk of developing PC is increased 13-22 fold among these patients (51, 52). Familial breast and ovarian cancer syndrome is an autosomal dominant syndrome due to germline mutation in either the BRCA1 and/or BRCA2 gene. Mutation carriers are at a higher risk of developing breast, ovarian, gastrointestinal and prostatic cancer (53, 54). The risk of developing PC is increased 3-10 fold among BRCA2 and 2.3-3.6 fold among those carrying the BRACA1 mutation (54). HNPCC occurring due to mismatch repair gene (MLH1, MSH2, MSH6 and PMS2) mutation is characterized by early-onset colorectal and extra-colonic malignancy (55) including an 8.6 fold increase in the risk of developing PC (56). Hereditary pancreatitis is a rare inherited autosomal dominant disorder with incomplete penetrance (57). The lifetime risk of developing PC is 40% among these patients (58). (58). FPC describes families with at least two first-degree relatives with confirmed exocrine PC that do not fulfil the criteria of other inherited tumour syndromes (59, 60).

1.3.1.3 Pancreatic cystic lesions

As a result of the widespread use of cross-sectional imaging, pancreatic cystic lesions are being discovered in increasing frequency (61). In an Italian study pancreatic cystic lesions were identified on 1.2% of 24,039 MRI or CT scans arranged for alternative pathology (62). The vast majority of lesions are asymptomatic. Pancreatic cystic lesions may rarely present with pain, exocrine insufficiency or obstructive jaundice (63). Accurate classification of the cystic lesion is imperative due to the potential for malignant change among specific lesions. This is often challenging and may be incorrect in up-to a third of cases (64). The most commonly recognised lesions are serous

cystic neoplasms, mucinous cystic neoplasms and intraductal papillary mucinous neoplasms.

1.3.1.3.1 Serous cystic neoplasms (SCN's)

SCN's account for 20-30% of all diagnosed pancreatic cystic tumours. SCN's occur most commonly in women with a peak incidence in those over 70 years old (65). Malignant transformation is rare (66). SCN's may be micro or macrocystic. Macrocystic lesions consist of unilocular or bilocular cysts greater than 2 cm in diameter, which can be difficult to differentiate from either mucinous cystic neoplasms or pancreatic pseudocysts. SCN's do not communicate with the pancreatic duct and have a honeycomb appearance on CT with a central area of calcification (65). Although the presence of glycogen containing cells on cytological examination of EUS cyst fluid is diagnostic, negative cytology does not exclude a diagnosis. Cytology alone correctly identifies only 38% of SCN's (67). A fluid amylase level of <250 U/L along with an absent history of pancreatitis may be used to distinguish pseudocysts and non-pseudocysts with a sensitivity and specificity of 44% and 98% respectively (67). Due to the low risk of malignant potential SCN's are often managed non-operatively (68) . Up to 10% of SCN's are mistaken for solid tumours such as neuroendocrine or solid pseudopapillary (69).

1.3.1.3.2 Mucinous cystic neoplasms (MCN'S)

MCN's occur mainly in women with a peak incidence in the fifth decade of life (65). Up-to 17.5% of MCN's (in contrast to SCN's) undergo malignant transformation (70). Findings on cross sectional imaging are that of a large unilocular cyst with peripheral calcification. MCN's do not communicate with the main pancreatic duct and occur predominantly within the body or tail of the pancreas. One pathogenic feature of MCN's along with IPMN's (also potentially pre-malignant) is the presence of mucin containing goblet cells at cytological examination. To improve diagnostic accuracy cyst carcinoembryonic antigen (CEA) and carbohydrate antigen 19-9 (CA 19-9) may be measured. CEA glycoprotein is secreted by the epithelium of mucinous but not serous pancreatic cysts (71). CEA however is of little use

in differentiating a mucinous lesion with dysplastic or malignant change from a benign mucinous cyst (72, 73). Due to the high incidence of invasive carcinoma within MCN's surgical resection for all patients who are medically fit is recommended (74).

1.3.1.3.3 Intraductal papillary mucinous neoplasms (IPMN's)

IPMN's occur predominantly in males most commonly in their seventh decade of life (75). In contrast to MCN's they communicate with either the main pancreatic duct or its side branch (76). Invasive carcinoma occurs in up-to 50% of main duct IPMN's versus 11% of branch duct IPMN's (77). On cross sectional imaging dilatation of either the main pancreatic duct or its side branch is typically seen (76). Surgical resection is recommended for all otherwise fit patients with main duct IPMN's (74). The 2012 consensus guidelines only recommended surgical treatment in patients with branch duct IPMN when one of the following are present: lesion >3 cm, mural nodules, main pancreatic duct dilatation >6 mm, symptomatic lesions or positive cytology (74).

1.4 Biomarkers & pancreatic cancer

For the reasons described above biomarkers are of crucial importance in the recognition and treatment of patients with suspected PC. A biomarker (biological marker), as defined by the National Institutes of Health Biomarker Working Group is a characteristic that is objectively measured and evaluated as an indicator of a normal biological process, pathogenic processes, or a pharmacologic response to a therapeutic intervention (15). A biomarker may be diagnostic, prognostic or predictive. The ideal diagnostic biomarker would be non-invasive, inexpensive, and potentially an effective screening test to identify high-risk premalignant lesions or early invasive cancers at a curable stage. Effective tests require a high degree of sensitivity and specificity to minimize unnecessary interventions for false positive results. Prognostic markers predict the natural history of disease. Predictive markers facilitate personalized treatment as they may be used to predict a response to a particular therapy such as chemotherapy (78).

1.4.1 Currently available biomarkers

1.4.1.1 Carcinoembryonic antigen

The glycoprotein carcinoembryonic antigen (CEA) which is used clinically as a biomarker for colorectal cancer was discovered in 1965 and was the first biomarker for PC (79). CEA has a sensitivity and specificity of 54% and 79% respectively for detecting PC. Low sensitivity along with known elevation in breast, gastric and colorectal cancer limit its use for PC diagnosis (80).

1.4.1.2 Carbohydrate antigen 19-9

The serum carbohydrate antigen 19-9 (CA 19-9) is the only biomarker for PC in routine clinical use today. CA 19-9 was first described in the colorectal cancer cell line SW1116 in 1979 (81). CA 19-9, also known as sialyl Lewis-a is a glycolipid expressed on the surface of cancer cells. CA 19-9 is derived during aberrant production of disialyl Lewis-a, which acts as a ligand for monocytes and macrophages (82).

1.4.1.2.1 Serum CA 19-9 as a diagnostic marker

In order to evaluate the utility of CA 19-9 in screening an asymptomatic population Kim *et al.* assessed CA 19-9 serum levels among 70,940 asymptomatic individuals. Among 1,063 individuals with an elevated CA 19-9 (> 37 U/ml) only 4 patients with PC were identified. Although sensitivity and specificity were 100 and 98.5% respectively, the positive predictive value of the test was only 0.9% (83). Routine serum CA 19-9 is therefore a poor test for screening an asymptomatic population.

The utility of serum CA 19-9 as a diagnostic biomarker significantly improves when deployed among a symptomatic patient cohort or a cohort presenting with a pancreatic mass. In a systematic review of 26 case-series (including a total of 2,283 symptomatic patients) Goonentilleke *et al.* described a sensitivity and specificity of 79 and 82% respectively along with positive and negative predictive values of 72 and 81% respectively (84).

1.4.1.2.2 Serum CA 19-9 to assess pancreatic cancer stage and resectability

Pre-operative serum CA 19-9 is known to vary with disease stage. Kim *et al.* studied 114 PC patients who underwent either pancreatic resection (n=72) or palliative bypass (n=42). Mean pre-operative serum CA 19-9 was 40.05 U/mL for patients with stage IA, 469.64 U/mL stage IIA, 747.79 U/mL stage IIB, 709 U/mL stage III and 3239 U/mL stage IV pancreatic cancer (85). A pre-operative CA 19-9 greater than 150 U/mL is known to correlate with an 88% positive predictive value for surgical irrresectability (86)

1.4.1.2.3 Serum CA 19-9 as a marker of prognosis and disease recurrence

In a study of 129 surgically resected PC patients Berger *et al.* stratified patients into four groups based on pre-operative CA 19-9 (undetectable, normal < 37 U/mL, 38-200 U/mL and > 200 U/mL). Patients with undetectable or levels < 37 U/mL had an improved median survival (32 and 35 months resepectively) compared to those with levels between 38-200 U/mL (22 months) and greater than 200 U/mL (16 months) (87). Smith *et al.* in 2008 similarly described significant survival variation with pre-operative CA 19-9. Among a cohort of 109 patients with surgically resected tumours median survival was found to be 10.4 versus 22.1 months among patients with a pre-operative serum CA 19-9 of less than or greater than 150 U/mL, respectively (88).

Post-operative CA 19-9 may also be similarly used to predict survival. Kondo *et al.* in a study including 109 surgically resected patients described that a CA 19-9 of <37 U/mL, < 200 U/mL and >500 U/mL measured 2-4 weeks post-operatively were associated with a 49%, 38% and 0% three-year survival rate respectively (89).

1.4.1.2.4 Limitations of serum CA 19-9

Serum CA 19-9 as a biomarker has several drawbacks. Firstly serum CA 19-9 is only elevated in 80-85% of pancreatic cancer patients (84) and may

often fail to detect early stage malignancies (90). CA 19-9 is in addition related to the Lewis blood group antigen, and only patients belonging to the Le (α-β+) or (α+β-) blood groups will express CA 19-9. CA 19-9 as a biomarker is as such ineffective among the 5-10% of the population who are Lewis blood group antigen negative (82, 84).

Serum CA 19-9 may also be elevated in non-malignant conditions such as pancreatitis, cirrhosis and cholangitis. Elevation may also occur in other gastrointestinal cancers such as gastric, oesophageal and cholangiocarcinoma (84, 91).

One significant confounding factor is up-regulation of CA 19-9 among patients with hyperbilirubinemia due to either benign or malignant disease. This is clearly an important issue as painless obstructive jaundice is the most common presentation of pancreatic cancer (84, 92). In order to differentiate malignancy from benign disease it has been suggested that the upper limit (or cut-off level) for CA 19-9 be increased beyond 37 U/mL. Alternative strategies include to re-assess serum CA 19-9 post biliary drainage. Re-assessing CA 19-9 however may lead to diagnostic delay. Marrelli *et al.* (among a cohort of 87 patients with pancreatico-biliary malignancy and 41 benign controls) through increasing the upper cut-off limit of CA 19-9 to 90 U/mL was able to describe an increase in CA 19-9 specificity for the diagnosis of malignant disease to 95%. Sensitivity however declined to 61%. In the same study serum CA 19-9 levels were elevated (>37 U/mL) among 61% of benign cases and 86% malignant cases. Following biliary drainage CA 19-9 levels decreased in nearly all benign cases (41 of 42, 98%) but only in 19 of 38 (50%) patients with malignant biliary obstruction (93).

1.4.2 Novel biomarkers for pancretic cancer

Many novel biomarkers have been described through the use of various analytical techniques. I will describe a broad overview of such markers present within tissue and/or body fluid. A further detailed review of novel

biomarker discovery through NMR and metabonomic techniques are described within section 1.8.

1.4.2.1 Novel biomarker markers in tissue

1.4.2.1.1 Kirsten rat sarcoma viral oncogene homolog (KRAS)

The KRAS gene resides within the Ras family of oncogenes. Proteins produced from these genes are GTPases, which play an important role in cell division, cell differentiation and apoptosis. The KRAS mutation can be identified in up-to 90% of PC cases (94). PanIN-I, PanIN-II and PanIN-III lesions are also known to harbour this mutation in 36%, 44% and 87% of cases respectively (95). The diagnostic utility of KRAS is limited by its lack of sensitivity and specificity. Similar mutations have also been seen in chronic pancreatitis, gastric and colorectal cancer (96).

1.4.2.1.2 Cyclin-dependent kinase inhibitor $P16^{INK4A}$

The loss of tumour suppressor $p16^{INK4A}$ function has been documented in up-to 95% of PC cases (97, 98). $P16^{INK4A}$ mutations are however not specific for PC and have been observed in familial cases of malignant melanoma and breast carcinoma (97). Loss of p16INK4A however occurs at earlier stages of PC development along with PanIN (98, 99).

1.4.2.1.3 Tumour suppressor p53

The tumour suppressor p53 is the most frequently mutated gene in human cancers. The p53 mutation has been identified in up-to 70% of PC's (100). The p53 mutation appears late in PC development which limits it's usefulness for early disease detection (101).

1.4.2.1.4 Smad4 gene

The Smad4 protein mediates transforming growth factor beta (TGF-β) signalling to regulate cell growth and differentiation. SMAD4 is inactivated in 55% of pancreatic cancers. Deletion of this protein correlates with a poor prognosis which is often due to metastatic disease (102).

1.4.2.2 Novel biomarkers in body fluid

1.4.2.2.1 Macrophage inhibitory cytokine-1 (MIC-1)

MIC-1 is a member of the transforming growth factor-β super-family (103). Macrophages are one of the sentinel cells of the innate immune system and provide a defense mechanism against cancer cells (104). MIC-1 presence in the tumor microenvironment limits the secretion of tumor necrosis factor-α (by activated macrophages) thereby reducing macrophage tumour killing activity (103). Koopmann *et al.* in 2006 identified serum MIC-1 as being significantly superior to CA19-9 in differentiating patient's with pancreatic cancer from healthy controls. 50 patients with resectable pancreatic cancer, 50 patients with chronic pancreatitis and 50 age/sex matched healthy controls were included in the study. It was however not possible to distinguish pancreatic cancer from those patients suffering from chronic pancreatitis (105).

1.4.2.2.2 Carcinoembryonic antigen-related cell adhesion molecule 1 (CEACAM1)

CEACAM1 is a cell-cell adhesion molecule found on leukocytes, epithelia, and endothelia. Multiple cellular activities have been attributed to CEACAM1 including differentiation and arrangement of tissue structure, angiogenesis, apoptosis, tumor suppression and modulation of the innate and adaptive immune responses. CEACAM1 lacks sensitivity and specificity to be used as a diagnostic biomarker alone. Simeone *et al*. in 2007 described CEACAM1 expression in the serum of 91% (74/81) of PC patients, 24% (15/61) of normal controls, and 66% (35/53) of patients with chronic pancreatitis (106).

1.4.2.2.3 Micro ribonucleic acids (miRNAs)

miRNAs were first discovered in 1993 and comprise of a class of highly conserved short (17-25 nucleotide) non-coding RNA products that regulate gene expression at the post-transcriptional level (107). miRNAs have shown promise for use as a biomarker for PC diagnosis, prognosis and chemosensitivity. Ma *et al.* (108) in a meta-analysis review (including 538 cancerous and 206 non-cancerous controls) described varying tissue miRNA

expression including seven up-regulated (including miR-21, miR-155, and miR-221) and three down-regulated (miR-217, miR-148a, and miR-375) miRNAs. High levels of tissue miR-21, miR-23a, and miR-27a have more recently being associated with shorter survival times after surgical resection (109).

Wang *et al.* in 2009 first described the detection of miRNAs within circulating blood plasma (110). Through profiling 28 pancreatic cancer and 19 control samples for miR-21, miR-210, miR-155 and miR-196, a sensitivity and specificity of 64% and 89% respectively was achieved for differentiating benign from neoplastic disease (110). Li *et al.* profiled sera from patients with pancreatic cancer (n=41), chronic pancreatitis (n=35), neuroendocrine tumours (n=18) and 19 healthy controls. Serum miR-1290 facilitated the detection of low stage pancreatic cancer from controls. Higher levels of miR-1290 in addition correlated with poorer outcome among patients undergoing pancreaticoduodenectomy (111) .

1.4.2.2.4 Glypican-1

Melo *et al.* in 2015 through the use of mass spectrometry described identification of the cell surface proteoglycan glypican-1 (GPCI), specifically enriched on circulating cancer cell derived exosomes (crExos). Exosomes are lipid-bilayer-enclosed extracellular vesicles (containing proteins and nucleic acids) which are secreted by all cells and circulated within the bloodstream. Significantly higher levels of GPCI positive crExos from the serum of all 190 patients suffering from pancreatic adenocarcinoma when compared to 100 healthy donors ($p<0.0001$) were described (112). These findings are clearly very significant, however it must also be noted that levels were also significantly raised in 75% of a separate patient cohort of 32 patients suffering from breast carcinoma within the same study. This clearly raises potential limitations with biomarker specificity.

Through comparing patients with stage I-IV pancreatic cancer to healthy donors and to those with benign pancreatic disease (n=26), the receiver-

operating characteristic curves revealed near perfect classification with an area under the curve of 1.0 exhibiting a sensitivity and specificity of 100%. Identical results were achieved through validation on an independent patient cohort composing of six patients suffering from chronic pancreatitis, 56 PDAC and 20 healthy donors. In addition GPCI positive crExos levels were also described to correlate with disease burden. Levels were significantly raised among patients with distal metastatic disease over nodal disease over localised disease. Levels of GPCI positive crExos were also interestingly significantly raised among a cohort of patients suffering from pancreatic cancer precursor diseases (IPMNs). Patient numbers among this cohort were however small (n=5) (112).

1.4.2.3 Novel biomarkers in stool

Bone morphogenetic proteins (BMP) constitute a large subgroup within the transforming growth factor beta (TGF- β) super-family. BMP'*s* are involved in a variety of developmental processes including tumour suppression (113). Kisiel *et al.* described elevated methylated BMP3 (mBMP3) in stool from PDA versus controls (non-neoplastic colonic epithelial controls) with a sensitivity and specificity of 51% and 90% respectively (113).

1.5 Metabonomics & metabolomics

Metabolomics is a newly emerging field of 'omics' research concerned with the high-throughput identification, quantification and characterization of small molecule (<1500 Da) metabolites in the metabolome (114). Metabonomics is "the quantitative measurement of the dynamic multi-parametric metabolic response of living systems to pathophysiological stimuli or genetic modification" (115). 'Metabolomics' places emphasis on metabolic profiling at a cellular or organ level and is primarily concerned with normal endogenous metabolism, whereas 'Metabonomics' extends metabolic profiling to include information about perturbations of metabolism caused by environmental factors (including diet and toxins) and disease processes. These terms are often however used interchangeably. The metabolome can be defined as the complete complement of all small molecule (<1500 Da) metabolites found in

a specific cell, organ or organism. It is a close counterpart to the genome, the transcriptome and the proteome. Together these four ‘omes’ constitute the building blocks of systems biology with the study of metabolites being the final downstream product of biological systems and hence closest to the phenotype. Thanks to the Human Genome Project most of the human genome, transcriptome and proteome are now know. Unfortunately the same cannot be said of the human metabolome. The Human Metabolome Project (HMP) was launched in 2004 as part of an effort to identify and quantify all detectable metabolites (>1 μM) in the human body (116). Over time metabolomics has evolved from a little-known branch of analytical chemistry to a mainstream enterprise.

Metabolomics can be used in a variety of applications including biomarker identification, drug discovery or development and clinical toxicology (117). The two main technologies deployed for metabolome analysis are mass spectrometry (MS) and nuclear magnetic resonance spectroscopy (NMR). NMR has the advantage of being fast with minimal sample preparation and is cost-effective. Bio-fluid samples may be kept close to their native state and variability due to sample preparation (which is not labour intensive) is kept to a minimum. NMR samples are never in direct contact with the equipment and contamination between samples is minimal. NMR is highly reproducible and each sample may be re-run with only minor changes in results (118). MS is however widely used and is the standard technique in the pharmaceutical industry. MS is more sensitive than NMR but usually requires more extensive sample preparation, which unlike NMR can result in metabolite loss (119). Sensitivity is the largest weakness of NMR spectroscopy, which presents itself as a detection limit in the micromolar range rather than the nanomolar range as for MS. In addition a large water peak signal is always encountered during bio-fluid analysis, which may obscure metabolites. Low molecular weight metabolites can also be obscured by the broad envelope of high molecular weight resonances of proteins encountered during NMR acquisition. Both problems can be substantially addressed by application of appropriate pulse sequences (120).

1.6 NMR spectroscopy

1.6.1 Classical and Quantum NMR Model

The nuclear magnetic resonance (NMR) phenomenon was first reported in 1946 (121). NMR exploits the behaviour of atomic nuclei in an externally applied magnetic field. Magnetic resonance occurs because of the quantum mechanical property of 'spin', a source of angular momentum intrinsic to nuclei with an odd mass number. Examples of such nuclei include hydrogen-1 (^{1}H), carbon-13 (^{13}C), fluorine-19 (^{19}F) and phosphorous-31 (^{31}P). The spinning nucleus like an electric current creates a tiny magnetic field. When placed in a strong magnetic field (B_0) the magnetic nucleus tries to align like a compass needle in the Earth's magnetic field. As the nucleus is spinning and has angular momentum the torque exerted by the external field results in a circular motion called precession (analogous to a spinning top in the Earth's gravitational field). The rate of precession is proportional to the external magnetic and nuclear magnetic field strength (Figure 1.1) and is termed the Lamor frequency (v_0) (122).

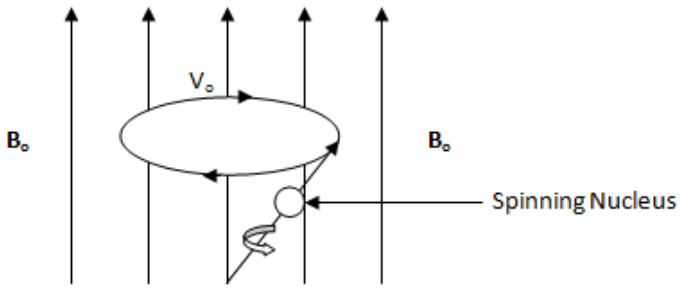

Figure 1.1 Classical NMR model

For "spin ½" nuclei there are two quantum states, which can be visualised as having the spin axis directed "up" or "down". In the absence of an external magnetic field these two states possess equal energy and at thermal equilibrium exactly half of a population of nuclei will be in the "up" state and the remainder in the "down" state. In the presence of an external magnetic field the "up" state is aligned with the magnetic field and is lower in energy than the "down" state which opposes the external magnetic field (122).

Among a population of nuclei in thermal equilibrium, slightly more than half will reside in the lower “up” energy state.

When a perpendicular radiofrequency (rf) pulse is applied the populations of the nuclear spins are disturbed and the bulk magnetisation is rotated from the z-axis (aligned with B_0) towards the x-y plane (Figure 1.2).

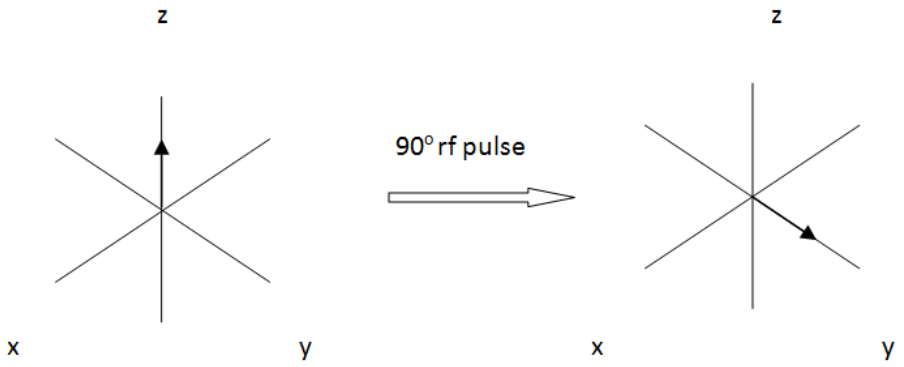

Figure 1.2 Effect of a 90° RF pulse on bulk magnetisation

The spins subsequently return back to equilibrium, thus allowing energy levels to return back to the Boltzmann distribution, with a slight excess in the lower energy state (122, 123). This resonant frequency can be measured by applying a radio frequency signal to the sample and varying the frequency until absorbance of energy is detected (122).

1.6.2 NMR pulsed Fourier transformation

Early NMR spectrometers (continuous wave spectrometry) recorded a spectrum by slowly changing the radio frequency signal fed into a coil near the sample. When the frequency passed through a resonant frequency for a given nucleus a “peak” in the spectrum was recorded. Modern NMR spectrometers operate in “pulsed Fourier-transformation” mode. The collection of nuclei in a sample, are exposed to a strong radiofrequency pulse to stimulate precession in unison. Over time individual nuclei get out of synch and the signal dies down. This “echo” of the pulse is called the free induction decay (FID). Fourier transformation of the FID (Figure 1.3) converts

if from a signal as a function of time to a plot of intensity as a function of frequency (123).

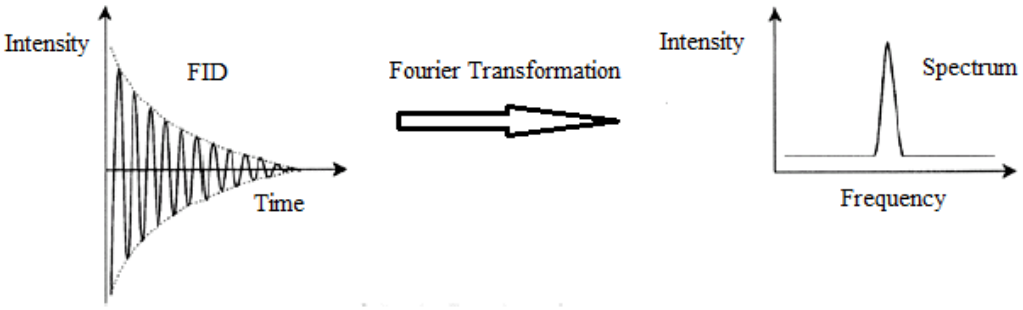

Figure 1.3 Fourier transformation of the FID

1.6.3 NMR chemical shift

The resonant frequency varies not only according to the type of nucleus but also on the position of that atom within a molecule (chemical environment). Bonding electrons create their own small magnetic field, which modifies the external magnetic field in the vicinity of the nucleus. This variation is called chemical shift and is measured in parts per million (ppm) (122, 123).

1.6.4 NMR pulse sequences

To visualise metabolites within bio-fluids with high water content the large dominant water signal must be suppressed (Figure 1.4). This is done through NMR pulse sequencing (which is also useful to suppress high molecular weight molecules such as protein). An understanding of longitudinal and transverse relaxation is fundamental to the understanding of NMR pulse sequencing and is described in section 1.6.5.

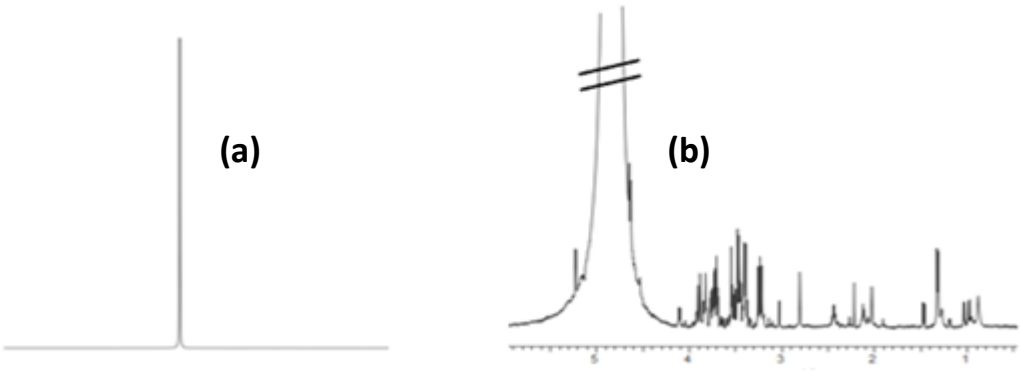

Figure 1.4 Suppression of the dominant water peak (a) through pulse sequencing (b)

1.6.5 NMR longitudinal & transverse relaxation

The dissipation of energy from the spins to the surrounding lattice and return of bulk magnetisation to the z-axis is known as longitudinal relaxation. The longitudinal relaxation time constant is defined as T_1 (123). If all spins experience an identical magnetic field, precession at the same frequency would occur. However, each spin will experience a different magnetic field due to the inhomogeneity of B_0 along with local magnetic field variation existing within the sample (123). Each sample can be thought of as being divided into several small regions (isochromats). Within each isochromat the magnetic field is uniform and the addition of these fields give rise to the total magnetisation. Unlike longitudinal relaxation, transverse relaxation is an entropic process as energy is transferred between spins rather than lost to the surrounding lattice (Figure 1.5). Given time the isochromats fan out as the bulk magnetisation reduces. Transverse relaxation is characterised by the time constant T_2 (122, 123).

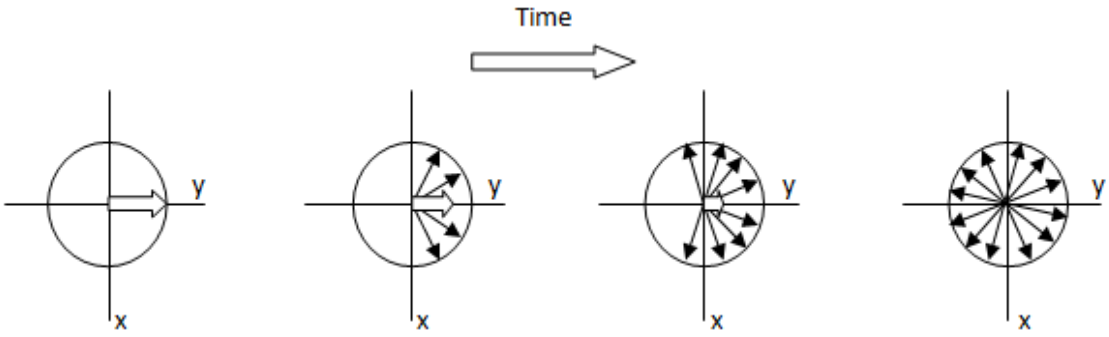

Figure 1.5 Fanning out of individual spins resulting in net zero magnetisation in the x-y plane (transverse relaxation)

1.6.6 NMR spin-echo pulse sequence

The different relaxation pathways for spins mean that T_2 is always less than or equal to T_1. The spin echo pulse sequence was devised to measure an accurate T_2 value of a sample through attempting to reduce the effect of field inhomogeneity (122, 123). The initial 90^0_x pulse pushes the magnetisation onto the y-axis where inhomogeneity of the static field causes isochromats to fan-out during the time period τ. A second pulse rotates all the isochromats by 180^0 to the –y-axis. This allows precessing isochromats to catch up with the average magnetisation vector, which is refocused (Figure 1.6). The second echo after 2τ is an exponential decay and is Fourier transformed to produce a spectrum.

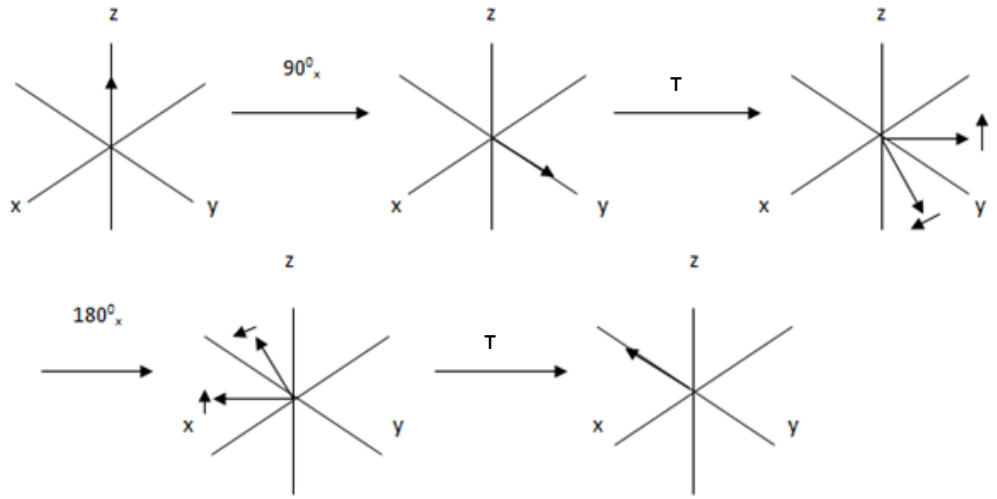

Figure 1.6 Magnetisation during the spin-echo pulse sequence and refocusing of the magnetic vectors

NMR spectra can be edited according to molecular size due to differences in T_2 duration. Pulse sequences take advantage of this to suppress or enhance various molecules. To suppress large molecules such as protein, which have short T_2 times a short τ delay is selected. The Carr-Purcell-Meiboom-Gill (CPMG) pulse sequence uses many repetitions of short τ delays and 180^0_x pulses. This technique removes signals from high molecular weight molecules (such as proteins) enabling those for small molecules to be resolved and quantified (Figure 1.7) (122, 123).

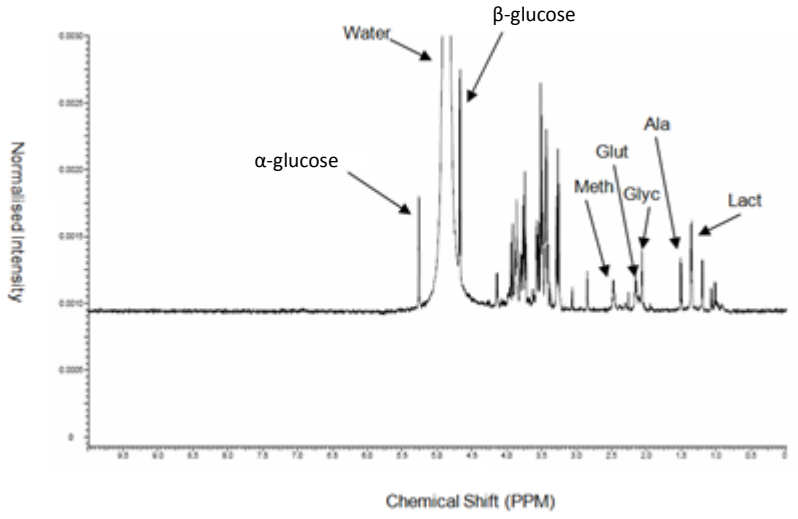

Figure 1.7 CPMG spectrum of plasma

For urine data acquisition the 1D Nuclear Overhauser effect spectroscopy (NOESY) sequence is commonly used. This applies a pre-saturation pulse to the water signal to saturate the spins prior to acquisition and analysis (123).

1.7 NMR Data analysis

1.7.1 Chemometrics

The science of metabonomics generates huge complex data tables that are hard to summarize with conventional statistics. Chemometrics utilises robust methods of modelling and analysis which take complicated chemical/biological data tables and produce interpretable and reliable statistical models (124). Principle component analysis (PCA) and partial least squares (PLS) analysis techniques are commonly used and will be discussed further. Prior to chemometric analysis NMR spectral data must be reduced, normalised and scaled (Figure 1.8).

1.7.2 Data reduction

Data reduction is commonly achieved through a process known as “binning”. Spectra are divided into smaller regions known as “bins” (125). Each of the “bins” are then integrated to achieve a numerical value that reflects the concentration of the species giving rise to the signal in that bin.

1.7.3 Normalisation

In metabonomic studies there will be samples where the total metabolite concentration will vary due to factors unrelated to the parameters being investigated. This is a feature of urine, which varies greatly in concentration. In the constant sum method the integral values of bins are summed to give a total value and then each bin integral is divided by this value. This allows signal intensities to be compared between samples (125).

1.7.4 Scaling

Metabolites of high concentration are not always the most informative. Scaling regulates the importance of each variable to avoid overlooking variation in lower concentration metabolites (126). Through subtracting the

mean from the entire sample set from each normalised variable the data may be mean centred (127). Division of centred data by the standard deviation of the whole sample set is known as Unit Variance and ensures all variables have the potential to influence the model (126, 127). To reduce the influence of spectra noise Pareto Scaling may be used whereby centred data is divided by the square root of the standard deviation of the whole sample (126, 127).

1.7.5 Principle component analysis (PCA)

PCA forms the basis for multivariate data analysis. PCA's main function is to reduce dimensions of a large unmanageable multivariate dataset (X) into a few manageable dimensions. The approximated data facilitates the identification of any clustering of samples within the overall dataset as well as outliers. The PCA process starts with the transformation of a multivariate table of data into multidimensional co-ordinate space. For *n* observations (samples), a *k* dimensional space is constructed, where *k* is the number of variables. Each sample is represented as a single point co-ordinate according to each variable point within the *k* dimensional space. The first principle component, the linear combination of the original variables represents the largest variation in the swarm of points. The second component represents the second largest variation and so on (124). The number of principle components in a model is determined by the difference in degree of fit and predictive ability of a model. The goodness of it is estimated by R^2. When the number of components increases, R^2 tends towards 1 as every value is predicted.

Any two principle components can form a plane, onto which observation can be projected as a scores plot. Observations that lie close to each other have similar multivariate hence metabonomic profiles. Strong outliers can be visualised by Hotelling's distribution which provides a tolerance region for the data (95% confidence interval) represented by an ellipse on the scores plot (124). Analogous to the scores plot, a loading plot may be generated which displays the weight or influence of individual variables in the model (124).

The loadings plot facilitates the identification of areas of the NMR spectrum accounting for maximum metabolic variation between samples (Figure 1.8).

1.7.6 Partial least squares discriminant analysis (PLS-DA)

Unlike PCA, partial least squares-discriminant analysis (PLS-DA) is a supervised method for multivariate analysis, which incorporates class membership to improve data separation. In addition to an X-matrix of observations (samples) and variables (bins), a Y-matrix consisting of class membership (case/control) is created. Orthogonal PLS (OPLS) may be used to enhance the interpretation of PLS by displaying all information into a single component for visual display (128).

For validation purposes PLS-DA may be used to predict class membership of samples from the X-matrix data. One-third of samples can be randomly excluded followed by generation of a new model. This model is then used to predict class allocation of the excluded samples. Permutation testing is an alternative method. In a permutation test the class labels of case and control are permuted or randomly assigned to different classes. With incorrect class labelling a new classification model is calculated. Permutation testing demonstrates how models in which the Y-variables (class membership) are randomised compare to the original PLS-DA model. An R^2 Y intercept value below 0.3-0.4 and a Q^2 Y less than 0.05 indicates a valid model (128) .

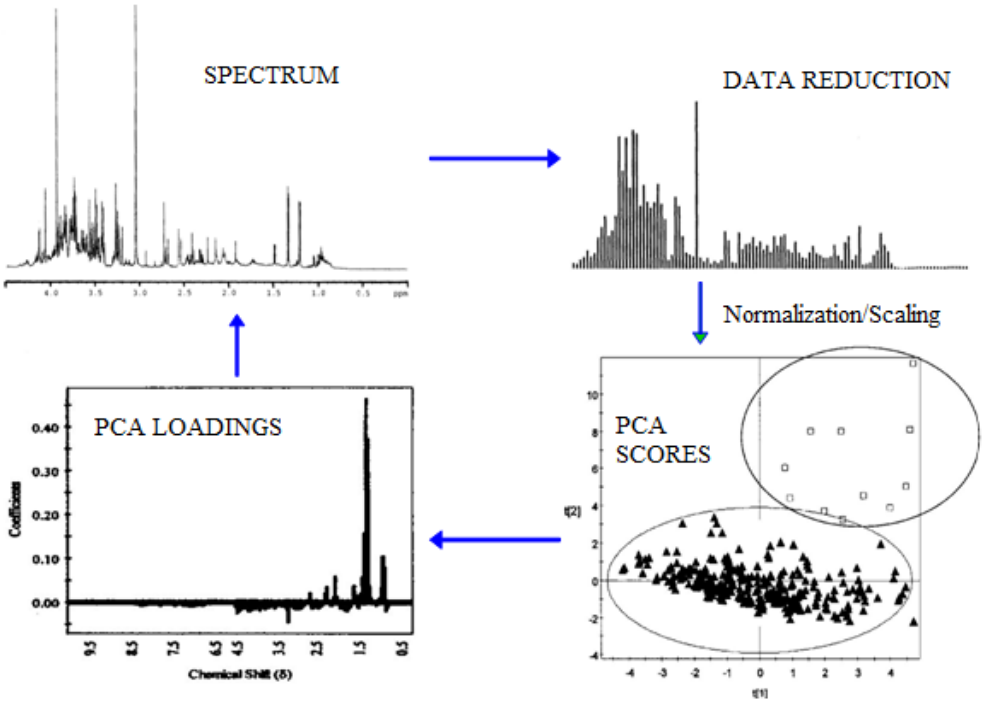

Figure 1.8 Chemometric analysis through data reduction ("binning") with subsequent scores and loadings plot

1.8 Prior NMR metabonomic studies of pancreatic cancer

1.8.1 Literature search

A literature search (title and abstract) of Ovid Medline (R) (1948–2014), Embase (1974-2014), PubMed, Web of Science and Sci-Finder electronic databases was performed up to and including 11th October 2014. The search was conducted using the MeSH terms: nuclear magnetic resonance spectroscopy, metabolomic, metabonomic, metabolic profiling (AND) pancreatic cancer. Studies comparing the metabolomic profile of human biological samples of pancreatic cancer with a control arm by NMR spectroscopy were included. Studies analysing the proteome, in-vivo imaging studies, animal studies, studies without a control arm and studies that reported the same patient population were excluded. Extracted data included primary author, date of publication, sample modality and number, the analytical platform used and statistically different metabolites between cancer and control arms of the study. The primary outcome measure was the

identification of metabolites found to have statistically different abundances between cancer and control samples.

1.8.2 Study identification and review

Eight studies met the inclusion criteria for systematic review. The biological samples used for analysis among the eight pancreatic cancer studies included serum (n=3), plasma (n=1), urine (n=2), bile (n=1) and pancreatic fluid (n=1). ^{1}H NMR was the metabolic platform used in all studies (Table 1-1). The metabolites identified to be either up or down regulated respectively are highlighted in Table 1-2.

Table 1-1 Prior metabonomic studies of pancreatic cancer utilising ^{1}H NMR

Author & Year	Sample	Pathology	Cancer (n)	Control	Control (n)	Up-regulated	Down regulated	Other findings/comments	Ref.
Bezabeh T, 2009	Bile	PC	4	BB	10	D-glucuronic acid			(129)
				CP	3				
Zhang L, 2012	Plasma	PC	19	HV	20	PC vs controls: N-acetyl glycoprotein, VLDL, lipid glyceryl, dimethylamine, acetone	PC vs. controls: 3-hydroxybutyrate, citrate, lactate, LDL, HDL, valine, lysine, leucine, isoleucine, histidine, glutamine, glutamate and alanine		(130)
		CP	20			CP vs. controls:glucose, lactate, creatine, formate, lipid glyceryls, tyrosine, phenylalanine, lysine, histidine, glutamine, glutamate, alanine	CP vs. controls:LDL, VLDL, 3-hydroxybutyrate and acetone		
						PC vs. CP: NAG, VLDL, dimethylamine, acetone	PC vs. CP: 3-hydroxybutyrate, creatine, lactate, citrate, LDL, HDL, lipid glyceryl, formate, valine, tyrosine, phenylalanine, lysine, isoleucine, histidine, glutamine, glutamate, and alanine		
Bathe OF, 2010	Serum	PC	43	BB	41	Glutamate, glucose	Creatine, glutamine		(131)
OuYang D, 2011	Serum	PC	17	HV	23	Isoleucine, triglyceride, leucine, creatinine	3-hydroxybutyrate, 3-hydroxyisovalerate, lactate, trimethylamine-N-oxide		(132)
Tesiram YA, 2012	Serum	PC	14	HV	12	Choline, taurine, glucose, triglycerides			(133)

Author & Year	Sample	Pathology	Cancer (n)	Control	Control (n)	Up-regulated	Down regulated	Other findings/comments	Ref.
Davis VW, 2012	Urine	PDA	32	HV	32	Acetone, hypoxanthine, O-acetylcarnitine, dimethylamine, choline, 1-methylnicotinamide, threonine, fucose, cis-aconitate, 4-pyridoxate, glucose, trimethylamine-N-oxide, aminobutyrate, tryptophan, xylose, trans-aconitate, 4-hydroxyisobutyrate, taurine	Trigonelline, methanol	11 metabolites no longer significantly raised in post-op period	(134)
Napoli C, 2012	Urine	PDA	33	HV	54	Acetoacetate, acetylated compounds, glucose, leucine, 2-phenylacetamide	Citrate, creatinine, glycine, hippurate, 3-hydroxyisovalerate		(135)
Wang J, 2011	Pancreatic fluid	PC	10	HV	5			Triplet peak of the methyl ethoxy group (CH3CH2O-) at chemical shift of 1.19 ppm in alcoholic chronic pancreatitis	(136)
		CP	5						
		Cirrhosis	15						

Table 1-2 Prior metabonomic studies of pancreatic cancer and metabolite identification through ^{1}H NMR analysis

Metabolite	Study (ref)					
	Zhang L (130)	**Bathe OF** (131)	**OuYang D (132)**	**Tesiram YA** (133)	**Davis VW** (134)	**Napoli C** (135)
Glucose		↑		↑	↑	↑
Citric Acid/Citrate	↓					
Lactate	↓		↓			
Acetone	↑				↑	
Isoleucine	↓		↑			
Leucine	↓		↑			↑
Lysine	↓					
Threonine					↑	
Tryptophan					↑	
Valine	↓					
Alanine	↓					
Glutamine	↓	↓				
Glutamate	↓	↑				
Taurine				↑	↑	
VLDL	↑					
LDL	↓					
HDL	↓					
Triglyceride			↑	↑		
Choline				↑	↑	
N-acetyl glycoprotein	↑					
Dimethylamine	↑				↑	
1-methylnicotinamide					↑	
2-phenylacetamide						↑
Trimethylamine-N-oxide			↓		↑	
Aminobutyrate					↑	
3-hydroxybutyrate	↓		↓			
4-hydroxyisobutyrate					↑	
3-hydroxyisovalerate			↓			
Acetoacetate						↑
O-acetylcarnitine					↑	
Cis-aconitate					↑	
4-pyridoxate					↑	
Creatine		↓				
Creatinine			↑			
Hypoxanthine					↑	
Xylose					↑	

Metabolite	Study (ref)					
	Zhang L (130)	**Bathe OF** (131)	**OuYang D (132)**	**Tesiram YA** (133)	**Davis VW** (134)	**Napoli C** (135)
Trans-aconitate					↑	
Ttigonelline					↓	
Methanol					↓	

As displayed in Table 1-1 and Table 1-2 it is clear that metabolite variation exists between pancreatic cancer and control patients. The most commonly identified metabolites (in two or more studies) are glucose, lactate, acetone, isoleucine, leucine, glutamine, glutamate, taurine, triglyceride, choline, dimethylamine, trimethylamine-N-oxide and 3-hydroxybutyrate (Table 1-2). Concordant findings are described for glucose (up-regulated), lactate (down regulated), acetone (up-regulated), glutamine (down-regulated), taurine (up-regulated), triglyceride (up-regulated), choline (up-regulated), dimethylamine (up-regulated) and 3-hydroxybutyrate (down-regulated) in cancer states. Contradictory findings are however apparent for isoleucine, leucine, glutamate and trimethlamine-N-oxide (Table 1-2). Contradictory results may be due to differences in sample collection, poor documentation of histology or control patient recruitment, the biological medium investigated or specific ^{1}H NMR technique (Table 1-3).

Table 1-3 Prior metabonomic study design and patient enrolment

		Bezabeth T (129)	**Zhang L (130)**	**Bathe OF (131)**	**OuYang D (132)**	**Tesiram YA (133)**	**Davis VW (134)**	**Napoli C 127)**	**Wang J (136)**
Benign/controls	**Cohort**	**CBD calculi (6), CP (3), post-liver transplant (1)**	**Healthy controls (20), CP (20)**	**Benign PB (43)**	**Healthy controls (23)**	**Healthy controls (12)**	**Healthy controls (32) age/sex matched**	**Healthy controls (54) age matched**	**CP (5)**
	Exclusions	-	Any PMH	Acute or sepsis	No PMH or recent illness	-	Infection, renal failure, breast feeding, pregnancy	**No PMH**	
	Age	Y	Y	Y	Y	Y	Y	Y	Y
	Sex	Y	Y	Y	Y	Y	Y	Male only	Y
	Ethnicity	-	-	-	-	-	-	-	-
	BMI	-	Y	-	-	-	-	Yes	-
	PMH	-	Y	Y	Y	Y	Y	Y	-
	Drug History	-	-	-	Y	-	Y	Y	-
	Bilirubin	Y	Y	Y	Y	-	Y	-	-
Cancer cohort	**Cell-line**	**PC (4), papillary cancer (1), cholangiocarcinoma (1), myeloma (1)**	**PC (19)**	**PDA (56)**	**PC (17)**	PDA (14)	**PDA (32)**	**PDA (33)**	**PC (10)**
	Exclusion	-	-	Acute inflammation or sepsis	No PMH or recent illness	-	Infection, renal failure, pregnancy		
	Grade	-	-	-	-	-	Y	Y	-
	Staging	-	-	-	-	Y	Y	Y	-
	Age	Y	Y	Y	Yes	Y	Y	Y	Yes
	Sex	Y	Y	Y	Yes	Y	Y	Male only	Yes

	Ethnicity	-	-	-	-	-	-	-	-
	BMI	-	Y	-	-	-	-	-	-
	PMH	-	-	Y	Y	-	-	Y	-
	Drug History	-	-	-	Y	-	-	Y	-
	Bilirubin	Y	-	Y	-	-	-	-	-
	Resectability	-	-	Y	-	-	Yes	-	-
	Survival	-	-	-	-	-	-	-	-

1.9 Plan of investigation

The main aim of this work is to determine any variation in the metabolic profile among patients with pancreatic carcinoma for potential diagnostic purposes. Through the use of ^{1}H NMR spectroscopy both plasma and urine will be profiled. Clear attention will be made to document the histological cell type, staging and grading among pancreatic cancer patients along with known confounders such as sex and race. In contrast to previous studies the metabolic profiles of patients with pre-malignant conditions will be compared against those with both confirmed malignancy and benign disease.

1.9.1 Hypothesis testing

A hypothesis is a statement about a population to be tested. The hypothesis of “no difference” within a population is called the null hypothesis(137). Within the context of this work a null hypothesis would be that there is no statistical difference between the mean integral peak in the NMR spectrum among patients with, versus those without PC. Acceptance or rejection of the null hypothesis is determined by a *p*-value, which is related to the means of two groups. The *p*-value is the probability of obtaining a test statistic test result at least as extreme as the one that was actually observed, assuming that the null hypothesis is true. A *p*-value of ≤0.05 is taken as statistically significant (137).

1.9.2 Biomarker model generation

We describe wherever possible any generated metabolic model for pancreatic cancer in terms of sensitivity, specificity, positive and negative predictive values through internal model validity testing.

2 Experimental methods

2.1 Ethical approval for study

This study was approved by the National Research Ethics Service (REC GS11/10064). The study conformed to the ethical guidelines outlined in the 1975 Declaration of Helsinki. Written informed consent was obtained for all patients.

2.2 Patient selection

2.2.1 Patient enrolment

Fifty-five patients undergoing pancreatic resection for presumed malignancy were enrolled into the study. Only patients undergoing surgery were approached for inclusion due to availability of post-operative histology. To act as a control arm for the study patients undergoing surgery for benign pancreaticobiliary disease were enrolled. Exclusion criteria included paediatric pancreatic resections, patients with a history of known diabetes mellitus and those with evidence of recent (within one week) or active infection at the time of surgery. A total of 34 patients with known benign disease and 55 with presumed malignant disease were recruited. Among those with presumed malignant disease only 10 patients were cytology or biopsy proven pre-operatively with a further two displaying highly suspicious cytology for malignancy (cellular atypia and pleomorphic nuclei). Eleven patients who underwent surgery for presumed malignant disease were shown on post-operative histology review to have either benign or pre-malignant disease. As such two initial groups of confirmed malignancy (n=44) and confirmed benign/pre-malignant (n=45) were formed. Each plasma sample was coded by the prefix “C” for confirmed cancer versus “B” for confirmed benign followed by “R” for plasma or “U” for urine followed by the subsequent sample number.

2.2.2 Confirmed malignancy patient cohort

A total of 44 patients with confirmed malignancy were recruited into the study. 27/44 (61%) of patients were male with an overall median age of 66 (40-81)

years. Among the cancer resection cohort twenty-eight patients underwent surgical resection (19 pylorus-preserving pancreaticoduodenectomy (PPPD), 6 distal pancreatectomy, one enucleation for early neuroendocrine tumour, one total pancreatectomy and one wide excision of gallbladder fossa and common bile duct excision for gallbladder carcinoma). Sixteen patients at time of laparotomy were found to have either locally advanced (n=11) or metastatic (n=5) disease. Of these thirteen underwent palliative single (n=3) or double bypass (n=8). Data regarding patient demographics, clinical presentation, pre-operative serum bilirubin, operative intervention, histology, chemotherapy and recurrence were collected (Table 2-1). Past-medical and prescription history is displayed in Table 2-2.

2.2.3 Confirmed benign patient cohort

A total of 45 patients with confirmed benign disease were recruited. 19/45 (42%) of patients were male with an overall median age of 53 (20-85) years. Thirty-one underwent cholecystectomy for gallstone disease. Among pancreatic resections performed for presumed pre-malignant or malignant change four had histology confirming dysplastic change whereas six had benign histology. Two patients underwent pancreaticojejunostomy for chronic pancreatitis, a further patient duodenal resection for low-grade gastro-intestinal stromal tumour and one patient open local excision of a retroperitoneal paraganglionoma. Patient demographics are displayed in Table 2-3 and patient past medical and prescription history in Table 2-4.

2.2.4 Cancer versus benign cohort age and sex variation

Patients among the confirmed cancer cohort were older than those with benign disease. A greater proportion of the benign cohort were female although statistical significance was not reached (Table 2-5)

2.2.5 CA 19-9

Among 55 patients undergoing surgery for potential malignancy, pre-operative serum CA 19-9 levels were available at the time of patient recruitment in 70.9% (n=39) patients. Among 55 patients undergoing surgery

44 (80%) were for histologically proven neoplastic disease and 11 (20%) for benign disease. Among 44 patients with neoplastic disease pre-operative serum CA 19-9 levels were available in 28 patients. Using an upper limit CA 19-9 of 37 u/mL, levels were raised pre-operatively in 20 (71.4%) of patients. Pre-operative CA 19-9 levels were within the normal range among all eleven patients who underwent surgery for histologically benign disease. Among this small cohort sensitivity and specificity of 71.4% and 100% respectively with positive and negative predictive values of 71.4% and 57.9% respectively were achieved for CA 19-9 as a diagnostic pre-operative biomarker.

Among the 28 patients with available pre-operative CA 19-9, 17 (60.7%) underwent pancreatic resection with 11 (39.3%) having irresectable disease. CA 19-9 was raised over and above 150 u/mL in 6 (35.3%) patients who underwent resection and was less than 150 u/mL among 3 (27.3%) patients with irresectable disease. Using CA 19-9 > 150 u/mL as a marker of irresectability failed to reach statistical significance among our patient cohort (p=0.1201).

Table 2-1 Cancer resection patient cohort

Sample Code		Sex	Age (yrs)	Ethnicity	ERCP stent (y/n)	Pre-op bilirubin (micromol/L)	Ca 19-9 (u/ml)	Operation	Diagnosis	Differentiation	Histology/Cytology	Adj Chemo (y/n)	Recurrence	DFS (mths)	OS (mths)
CR	CU														
1	1	F	73	Caucasian	y	10	156	PPPD	PDA	Mod-poor	T3N1V1R1	y	12	12	15
2	2	F	48	Caucasian	y	40	–	PPPD	PDA	Mod	T3N1V1R1	y	9	9	14
3	3	F	63	Caucasian	y	55	196	PPPD	PDA	Mod-poor	T3N1V1R1	y	13	13	17
Spoiled	4	F	80	Caucasian	y	60	240	Palliative	PDA	Mod	Adenocarcinoma	n	Locally advanced	0	15
5	5	M	65	Caucasian	y	53	13711	Palliative	PDA	Mod	Adenocarcinoma	y	Metastatic disease at surgery	0	6
6	6	F	71	Caucasian	y	129	100	PPPD	Ampullary carcinoma	Poor	T4N1V1R1	y	9	9	18
7	7	M	58	Caucasian	y	14	436	PPPD	PDA	Mod-poor	T3N1V1R1	y	–	26	–
8	8	M	73	Caucasian	y	116	1069	Palliative	PDA	Mod	Adenocarcinoma	y	Locally advanced	0	–
9	9	M	65	Caucasian	y	48	432	Palliative	PDA	Mod	Adenocarcinoma	n	Metastatic disease at surgery	0	6
10	10	F	78	Caucasian	y	19	–	PPPD	Adenocarcinoma from side branch IPMN	Mod-poor	T3N1V1R1	n	16	16	23
11	11	F	66	Caucasian	y	86	2774	Palliative	PDA	Mod	Adenocarcinoma	0	Metastatic disease at surgery	0	2
12	12	M	65	Caucasian	y	53	437	Palliative	PDA	Mod-poor	Adenocarcinoma	y	Locally advanced	0	15
13	13	M	76	Pakistani	y	10	–	PPPD	Ampullary carcinoma	Well	T4N1V1R1	0	12	12	14
14	14	M	67	Caucasian	y	9	1271	Palliative	PDA	Well	Adenocarcinoma	n	Locally advanced	0	18
15	15	M	63	Caucasian	n	132	–	Palliative	Duodenal Carcinoma	Poor	Adenocarcinoma	y	Locally advanced	0	10
16	16	F	69	Caucasian	y	9	164	PPPD	PDA	Mod	T3N1V1R1	n	–	24	–
17	17	M	74	Caucasian	y	28	–	PPPD	PDA	Mod	T3N1V1R1	y	–	9	–
18	18	F	70	Caucasian	y	9	–	Palliative	Distal cholangiocarcinoma	Mod	Adenocarcinoma	n	Locally advanced	0	–
19	19	M	68	Caucasian	n	20	5	LDP + splen	Neuroendocrine	Well	T2N0V0R0	n	–	9	–

Sample Code		Sex	Age (yrs)	Ethnicity	ERCP stent (y/n)	Pre-op bilirubin (micromol/L)	Ca 19-9 (u/ml)	Operation	Diagnosis	Differentiation	Histology/Cytology	Adj Chemo (y/n)	Recurrence	DFS (mths)	OS (mths)
CR	CU														
20	20	M	80	Caucasian	n	7	17	Palliative	Duodenal Carcinoma	Mod	Adenocarcinoma	y	Locally advanced	0	8
21	21	M	62	Caucasian	n	10	171	PPPD	Acinar cell carcinoma	Mod	T3NOR1	y	–	9	–
22	22	M	59	Caucasian	y	7	62	Palliative	PDA	Poor	Adenocarcinoma	y	Metastatic disease at surgery	0	–
23	23	M	65	Caucasian	y	41	–	Palliative	PDA	Poor	Adenocarcinoma	y	Locally advanced	0	–
24	24	F	71	Caucasian	n	5	–	ODP + splen	Adenosquamous carcinoma	Poor	T3N1V1R1	n	–	8	–
25	25	F	61	Caucasian	n	9	–	Enucleation	Neuroendocrine	Well	T1NxV0R1	n	–	8	–
26	26	F	81	Caucasian	y	13	26	Wide excision GB, CBD resection	Gallbladder carcinoma	Poor	T2N1R1	n	–	8	–
27	27	M	66	Caucasian	y	9	91	PPPD	PDA	Mod	T3N1V1R1	y	–	7	–
28	28	F	40	Caucasian	y	37	–	Palliative	Neuroendocrine	Poor	Neuroendocrine	y	Metastatic disease at surgery	0	2
29	29	M	72	Caucasian	n	9	53	Palliative	PDA	Mod	Adeno CA	y	Locally advanced	0	–
30	30	M	45	Caucasian	n	15	–	LDP	Neuroendocrine	Well	T2N0MxR0	n	–	7	–
31	31	M	62	Caucasian	y	12	90	PPPD	PDA	Mod	T3N1V1R1	y	–	6	–
32	32	M	66	Caucasian	n	11	25	ODP	Pancreatic adenosquamous	Mod	T3N1V1R1	y	–	6	–
33	33	F	66	Caucasian	y	40	436	Palliative	PDA	Poor	Adenocarcinoma	y	Locally advanced	0	–
34	34	M	45	Caucasian	n	12	11	PPPD	PDA	Mod	T3N1V1R1	n	–	5	–
35	35	M	68	Caucasian	y	29	–	Palliative	PDA	Mod	Adenocarcinoma	y	Locally advanced	0	5
36	36	M	75	Caucasian	y	14	–	PPPD	PDA	Poor	T3N1V1R1	n	4	5	–
37	37	M	79	Caucasian	n	9	5	PPPD	Colloid carcinoma arising within main duct IPMN	Mod	T3N1V0R1	y	–	5	–
38	38	F	59	Caucasian	y	28	19	TP + splen	Colloid carcinoma arising within main duct IPMN	Mod	T3N1V0R0	y	–	5	–
39	39	M	73	Caucasian	PTC	10	213	PPPD	Ampullary carcinoma	Mod	T4N1V1R0	y	–	5	–
40	40	M	76	Caucasian	y	6	66	PPPD	Ampullary carcinoma	Mod	T3N1R1	y	–	4	–

Sample Code		Sex	Age (yrs)	Ethnicity	ERCP stent (y/n)	Pre-op bilirubin (micromol/L)	Ca 19-9 (u/ml)	Operation	Diagnosis	Differentiation	Histology/Cytology	Adj Chemo (y/n)	Recurrence	DFS (mths)	OS (mths)
CR	CU														
41	41	M	50	Caucasian	y	12	–	PPPD	Distal cholangiocarcinoma	Mod	T3N1V1R1	n	–	4	–
42	42	F	64	Caucasian	n	7	–	ODP + splen + gastrectomy + transverse colectomy	PDA	Poor	T3N1V1R1	n	–	4	–
43	43	F	54	Caucasian	y	24	–	PPPD	PDA	Poor	T3N0V1R1	y	–	4	–
44	44	M	69	Caucasian	n	16	28	ODP + splen	PDA	Mod	T3N1V1R1	y	–	3	–

Table 2-2 Cancer resection patient cohort past medical and pharmaceutical history

Sample Code		Presenting features	Pancreatic tumour site	Tumour size (mm)	Past medical history	Drug history
CR	CU					
1	1	Jaundice	Uncinate	20.0	HTN, IHD, Hypothyroid	Lansoprazole, aspirin, biosprolol, nicorandil, levothyroxine, losaran
2	2	Jaundice	Head	24.0	Retinoblastoma, 1b malignant melanoma	Lansoprazole
3	3	Jaundice	Head	25.0	T2N1 Breast Cancer	–
Spoiled	4	Abdo pain + jaundice	Head	30.0	–	–
5	5	Jaundice	Head	28.0	–	Glicalzide, ramipril, amlodipine, omeprazole
6	6	Abdo pain	Head	28.0	T1N0 Breast Cancer	Anastrazole, aspirin, budesonide, carbocistine, omeprazole, simvastatin
7	7	Jaundice	Head/neck	40.0	Gastro-oesophageal reflux, HTN	Loratidine, atenolol
8	8	Jaundice	Head	34.0	–	–
9	9	Jaundice	Head	40.0	Epilepsy	Tegretol
10	10	Jaundice	Head	50.0	HTN	Atenolol, valsartan, methyldopa
11	11	Jaundice	Head	32.0	HTN	Lisinopril
12	12	Jaundice	Head	30.0	–	–
13	13	Jaundice	Ampulla	15.0	HTN, IHD	Aspirin, biosoprolol, levothyroxine, ramipril

Sample Code		Presenting features	Pancreatic tumour site	Tumour size (mm)	Past medical history	Drug history
CR	CU					
14	14	Jaundice	Neck	35.0	–	–
15	15	Gastric outflow obst	Head	Undefined	–	–
16	16	Jaundice	Ampulla	17.0	Polymyalgia rheumatica	Prednisolone, pantoprazole
17	17	Jaundice	Uncinate	35.0	HTN	Bendroflumethiazide, enalapril
18	18	Jaundice	Uncinate	15.0		–
19	19	Abdo pain	Body/tail	28.0	HTN, IHD	Salbutamol, ramipril, bisoprolol, simvastatin, aspirin, furosemide, naproxin
20	20	GI bleed	Duodenal	30.0	T4N0M1 mid-sigmoid & asc colon cancer (resected)	Aspirin, ferrous sulphate, quinine, gabapentin, pantoprazole, lope amide, folic acid
21	21	Abdo pain	Head	33.0	–	–
22	22	Abdo pain	Ampulla	15.0	–	–
23	23	Jaundice	Head	20.0	Diffuse large B cell lymphoma (prior R CHOP)	Gliclazide, metformin
24	24	Abdo pain	Tail	30.0	–	–
25	25	Hypoglycaemia	Head	15.0	COPD	Diazoxide, seretide, tiotropium
26	26	Jaundice	Gallbladder	10.0	HTN	Bendroflumethiazide, doxazosin, simvastatin
27	27	Jaundice	Head	32.0	HTN	Ranitidine, aspirin
28	28	Jaundice	Ampulla	16.0	–	–

Sample Code		Presenting features	Pancreatic tumour site	Tumour size (mm)	Past medical history	Drug history
CR	CU					
29	29	Abdo pain	Uncinate	48.0	HTN, IHD	Amlodipine, citirizine, citalopram , pravastatin, salbutamol, GTN
30	30	Incidental	Tail	35.0	Multiple endocrine neoplasia 1 - parathyroidectomy	–
31	31	Jaundice	Head	25.0	–	–
32	32	Incidental	Tail	48.0	Chronic lymphocytic leukaemia, IHD	Bisoprolol, simvastatin
33	33	Jaundice	Head	20.0	–	–
34	34	Incidental	Head	13.0	Ulcerative colitis	–
35	35	Chang bowel habit	Head	32.0	–	–
36	36	Jaundice	Head	37.0	–	–
37	37	Panc Insuff	Head	18.0	Asthma	–
38	38	Abdo pain and Jaundice	Head	25.0	COPD, Peripheral vascular disease, hypothyroid	Thyroxine
39	39	Jaundice	Ampulla	25.0	Prostate cancer, monoclonal gammopathy of unknown significance	Atorvastatin, bendroflumethiazide, amlodipine, omeprazole
40	40	Gastric outflow obst	Ampulla	25.0	Benign prostatic hypertrophy	Tamsulosin, finasteride
41	41	Jaundice	Ampulla	10.0	–	–
42	42	Abdomial pain	Body/tail	79.0	Squamous cell carcinoma of skin	–
43	43	Jaundice	Uncinate	30.0	Hypothyroid, asthma	Levothyroxine, seretide, salbutamol
44	44	Incidental	Neck	35.0	–	–

Table 2-3 Benign patient cohort

Sample Code		Sex	Age (yrs)	Ethnicity	Pre-op Bilirubin (micromol/L)	Operation	Pathology	Benign (0)/Pre-malignant histology (1)
BR	BU							
1	1	F	23	Caucasian	13	Lap chole	Gallstones/acute cholecystitis	0
2	2	F	50	Pakistani	38	PPPD	T2 AIP	0
3	No urine	F	25	Caucasian	8	Lap chole	Gallstones/chronic cholecystitis	0
4	No urine	F	54	Caucasian	6	Lap chole	Gallstones/chronic cholecystitis	0
No plasma	5	F	62	Caucasian	5	LDP + splen	Mixed type IPMN focal high grade dysplasia	1
6	6	F	59	Caucasian	8	PPPD	Side branch IPMN low and moderate dysplasia	1
No plasma	7	M	48	Caucasian	3	Pancreaticojejunostomy	Chronic pancreatitis	0
8	8	F	72	Caucasian	9	LDP + splen	Mixed type IPMN intermediate dysplasia	1
9	9	F	62	Caucasian	9	Lap chole	Gallstones/chronic cholecystitis	0
10	10	F	37	Caucasian	11	Lap chole	Gallstones/chronic cholecystitis	0
11	11	F	55	Caucasian	3	Lap chole	Gallstones/chronic cholecystitis	0
12	12	M	78	Caucasian	6	Open chole	Gallstones/chronic cholecystitis	0
No plasma	13	F	34	Caucasian	14	Lap chole	Gallstones/chronic cholecystitis	0
14	14	M	84	Caucasian	11	Lap chole	Gallstones/chronic cholecystitis	0
15	15	F	34	Caucasian	8	Open chole	Gallstones/chronic cholecystitis	0
16	16	M	79	Caucasian	21	Open chole	Gallstones/chronic cholecystitis	0
17	17	M	71	Caucasian	23	Lap chole	Gallstones/chronic cholecystitis	0
18	18	F	46	Caucasian	5	Lap chole	Gallstones/chronic cholecystitis	0
19	19	M	47	Caucasian	7	Lap chole	Gallstones/chronic cholecystitis	0
20	20	M	72	Caucasian	9	PPPD	Side branch IPMN no malignancy	0

Sample Code		Sex	Age (yrs)	Ethnicity	Pre-op Bilirubin (micromol/L)	Operation	Pathology	Benign (0)/Pre-malignant histology (1)
BR	BU							
21	21	F	24	Caucasian	7	Lap chole	Gallstones/chronic cholecystitis	0
22	22	F	76	Caucasian	7	ODP + splen	Serous cystadenoma	0
23	23	M	65	Caucasian	7	Open chole	Gallstones/chronic cholecystitis	0
24	24	F	29	Polish	4	Lap chole	Gallstones/chronic cholecystitis	0
25	25	M	42	Caucasian	7	Lap chole	Gallstones/chronic cholecystitis	0
26	26	F	21	Caucasian	13	Lap chole	Gallstones/chronic cholecystitis	0
27	27	M	46	Caucasian	7	LDP + splen	Benign cyst	0
28	28	F	62	Caucasian	4	Duodenal resection	Low grade GIST	1
29	29	F	58	Caucasian	10	Open chole + splenectomy	Gallstones/chronic cholecystitis	0
30	30	F	42	Pakistani	9	Lap chole	Gallstones/chronic cholecystitis	0
31	31	M	57	Polish	10	Lap chole	Gallstones/chronic cholecystitis	0
32	32	M	62	Pakistani	11	Lap chole	Gallstones/chronic cholecystitis	0
33	33	M	42	Caucasian	15	PPPD	IPMN intermediate grade dysplasia	1
34	34	M	20	Caucasian	15	Lap chole	Gallstones/chronic cholecystitis	0
35	35	F	67	Caucasian	9	Local excision	Paraganglioma	0
36	36	M	52	Caucasian	4	Hepaticojejunostomy	Chronic pancreatitis	0
37	37	M	42	Caucasian	14	Lap chole	Gallstones/chronic cholecystitis	0
38	38	F	68	Caucasian	7	ODP + splen	Serous microcystic adenoma	0
39	39	F	75	Caucasian	8	Lap chole	Gallstones/chronic cholecystitis	0
40	No urine	F	85	Caucasian	5	Lap chole	Gallstones/chronic cholecystitis	0
41	41	M	38	Caucasian	20	Lap chole	Gallstones/chronic cholecystitis	0
42	42	M	44	Caucasian	6	Lap chole	Gallstones/chronic cholecystitis	0

Sample Code		Sex	Age (yrs)	Ethnicity	Pre-op Bilirubin (micromol/L)	Operation	Pathology	Benign (0)/Pre-malignant histology (1)
BR	BU							
43	43	F	22	Caucasian	13	Lap chole	Gallstones/chronic cholecystitis	0
44	44	M	53	Caucasian	5	ODP + splen	Chronic pancreatitis	0
45	45	F	70	Caucasian	7	Lap chole	Gallstones/chronic cholecystitis	0

Table 2-4 Benign patient cohort past-medical and drug history

Sample Code		Operative Indication	Past medical history	Drug history
BR	BU			
1	1	Biliary colic	–	–
2	2	Presumed malignancy	–	–
3	No urine	Biliary colic	–	–
4	No urine	Biliary colic	–	–
No plasma	5	Main duct IPMN pancreatitis	HTN	Co-tenidone
6	6	Presumed malignancy	–	–
No plasma	7	Chronic pancreatitis dilated pancreatic duct	Chronic pancreatitis	–
8	8	Presumed malignancy	2013 Anterior resection (Dukes B rectal cancer)	–
9	9	Biliary colic	Hiatus hernia	Solifenacin, lansoprazole
10	10	Biliary colic	–	–
11	11	Biliary colic	–	–
12	12	Prior cholecystitis	IgA lambda myeloma 2009	–

Sample Code		Operative Indication	Past medical history	Drug history
BR	BU			
No plasma	13	Acute pancreatitis	–	–
14	14	Acute pancreatitis	–	–
15	15	Biliary colic	Breast ductal carcinoma in-situ (Excision and radiotherapy) 2013	–
16	16	Prior cholecystitis	Right thoracoplasty 1963 (TB)	–
17	17	Prior cholecystitis	CABG, atrial fibrillation	Warfarin, biosprolol, diltiazem, ramipril, simvastatin, salbutamol, seretide
18	18	Prior cholecystitis	Lichen planus, gastro-oesophageal reflux, asthma, osteoarthritis	Omeprazole, methotrexate, folic acid, amitriptyline, citalopram, MST, salbutamol
19	19	Biliary colic	–	–
20	20	Presumed malignancy	Prostate cancer (prostatectomy 2011)	–
21	21	Prior CBD calculi	2011 thyroidectmy Graves disease	Thyroxine
22	22	Presumed malignancy	Aortic stenosis	–
23	23	Prior cholecystitis	–	–
24	24	Biliary colic	Segment 5 liver focal nodular hyperplasia	–
25	25	Biliary colic	Laparotomy malrotation (child)	–
26	26	Biliary colic	Narcolepsy, epilepsy	Lamotrigine, Modafinil
27	27	Presumed malignancy	HTN, pancreatitis	Simvastatin, lisinopril, desloratidine, aspirin, metformin

Sample Code		Operative Indication	Past medical history	Drug history
BR	BU			
28	28	Duodenal tumour upper gastrointestinal bleed	–	–
29	29	Prior cholecystitis (incidental splenic lesion)	–	–
30	30	Biliary colic	–	–
31	31	Biliary colic	–	–
32	32	Biliary colic	Hyperthyroid	Carbimazole
33	33	Presumed malignancy	–	–
34	34	Acute pancreatitis	–	–
35	35	Presumed malignancy	Myocardial infarct, COPD	–
36	36	Chronic pancreatitis. CBD stricture recurrent cholangitis	Chronic pancreatitis (ETOH) recurrent CBD stents	–
37	37	Prior cholecystitis	HTN	Amlodiopne, bisoprolol, ramipril
38	38	Presumed malignancy	–	
39	39	Prior cholecystitis	HTN	Atenolol, bendroflumethiazide
40	No urine	Prior cholecystitis	–	–
41	41	Prior cholecystitis	–	–
42	42	Biliary colic	–	–
43	43	Biliary colic	–	–
44	44	Presumed malignancy	–	Citalopram
45	45	Prior CBD calculi	–	–

Table 2-5 Cancer versus benign cohort age and sex variation

	Cancer (n=44)	Benign (n=45)	Statistical Significance (p)
Age (yrs) mean +/- SD	65.9 +/- 9.7	52.3 +/- 18.3	0.0001
Sex M/F	27/17	19/26	0.0908

2.3 Sample collection and processing

A 5 ml blood sample was collected into a Lithium Heparinised tube via a venous cannula inserted for the purpose of general anaesthesia in each patient. A 5 ml urine sample was obtained from a urinary catheter placed for clinical need among major pancreatic resection patients. Patients managed without urinary catheter were asked to provide a sample pre-operatively.

All samples were processed within one-hour of collection. Blood samples were centrifuged at 2000 rpm for ten-minutes. The plasma was removed from the sample and stored in eppendorf tubes at -80°C prior to NMR data acquisition. Urine was transferred into eppendorf tubes and similarly stored at -80°C.

2.3.1 NMR sample preparation

Chemicals were purchased from Sigma-Aldrich Company Ltd. (Poole, Dorset, UK), unless otherwise stated. NMR tubes (S-5-500-7, Norell) were purchased from GPE Scientific Ltd. (Leighton Buzzard, Bedfordshire, UK).

2.3.1.1 Plasma samples

Individual samples were returned to room temperature (20°C) and centrifuged (Hettich Mikro 120 (C1204) centrifuge, angle rotor A1242) at 3000 rpm (11,992g) for one-minute to remove any sediment. 300 µL of plasma was added to 350 µL of deuterium oxide (D_2O). The mixture was vortexed for 8 seconds before transferring 600 µL into a 5 mm NMR tube. Samples were prepared individually to minimise any on-going metabolism prior to analysis.

2.3.1.2 Urine samples

Individual samples were returned to room temperature (20°C) and centrifuged (Hettich Mikro 120 (C1204) centrifuge, angle rotor A1242) at 11,992g for 5 minutes. 460 µl of urine supernatant was added to 230 µL of phosphate buffer solution. The urine/phosphate buffer mixture was vortexed for 8 seconds before transferring 600 µL to a 5 mm NMR tube.

100 ml of phosphate buffer solution (pH 7.43) contained 2.885 g sodium phosphate dibasic (Na_2HPO_4), 0.525 g sodium phosphate monobasic (NaH_2PO_4), 0.0172 g (1 mM) trimethylsilyl propanoic acid (TSP) and 0.0195 g (3 mM) sodium azide (NaN_3) in 20 ml of D_2O and 80 ml of ribonuclease (RNase) free water. The phosphate buffer was shaken thoroughly until salts dissolved.

2.4 NMR data collection

All ^{1}H-NMR spectra were acquired on a Varian Unity Inova 500 spectrometer (Varian Inc, Palo Alto, California, USA) operating at 499.97 MHz proton frequency, at 20°C.

2.4.1 Carr-Purcel-Meiboom-Gill (CPMG) experiment

The CPMG pulse sequence [RD - 90° - (τ - 180° - τ)$_n$ – acq] was used to obtain metabolic profiles of plasma samples. A relaxation delay (RD) of 2 s, τ 1.5 ms and *n* of 150 was used for data collection. For each spectrum 512 transients were collected into 16,284 pairs of data points with a spectral width of 8,000.00 Hz.

2.4.2 1D Nuclear Overhauser effect spectroscopy (NOESY) experiment

The one-dimensional (1D) NOESY pulse sequence [RD - 90° - t_1 - 90° - t_m - 90° - acq] was used to obtain metabolic profiles for all urine samples. A RD of 2 s, t_m of 1.5ms and t_1 of 3 µs was used for data collection. For each spectrum 512 transients were collected into 16,284 pair of data points with a spectral width of 8,000.00 Hz.

2.4.3 NMR spectral processing

All spectra were processed using ACD Labs software 12.01 (Advanced Chemistry Development, Inc., (ACD/Labs), Toronto, Canada). An exponential line broadening (1 Hz) was applied to each free induction decay (FID) prior to zero filling to 65,536 (plasma) and 131,072 points for urine. The resulting spectra were phased, baseline corrected and referenced. Plasma spectra were referenced to the α-glucose (^{1}H chemical shift 5.23 ppm) and urine to TSP at 0.000 ppm.

2.5 Chemometric analysis of plasma spectra

2.5.1 Binning and dark regions

Prior to binning over a spectral range, several dark regions were created within plasma spectra. Dark regions are areas with an integral set to zero to exclude specific variables from subsequent multivariate analysis. These included water (4.20-5.70 ppm) and glucose (3.18-3.94 ppm) within plasma and water (4.50 – 6.20 ppm) and creatinine (3.034-3.064, 4.043-4.073) within urine spectra. Spectral binning was performed using ACD Labs software 12.01. All spectra were integrated into bins of 0.005 ppm using the intelligent bucketing process with a 50% looseness of width. The intelligent bucketing process limits spectra division between peaks and thus reduces the risk of metabolite loss.

2.5.2 Multivariate analysis

All spectra were mean centred and Pareto-scaled using SIMCA-P+ software version 12.0.1.0 (Umetrics, Umeå, Sweden). PCA was performed to view any clustering or outliers in the scores plots. Loadings plots were used to identify regions of the spectra (bins) responsible for scores plot clustering. PLS-DA was performed to improve distinction of separation between groups of interest and produce models for validation purposes. The quality of each model was assessed by goodness of fit (R^2 **X**) and the ability to predict class membership of a new sample (Q^2 **Y**). Model predictive ability was assessed by permutation testing and leave-one-out cross validation.

3 Results of NMR analysis

3.1 Analysis of plasma

This chapter describes the ^{1}H-NMR spectroscopic analysis of plasma obtained from 44 patients with confirmed pancreaticobiliary cancer and 45 patients with benign pancreaticobiliary disease. Among those with confirmed cancer, one plasma specimen was spoiled during pre-processing (CR4) and ethical considerations and timing prohibited further sample collection. Among the benign cohort three patients (BR5, BR7 and BR13) failed to provide samples. A total of 43 and 42 confirmed cancer and benign patient samples respectively were analysed via ^{1}H-NMR using the CPMG pulse sequence as described in section 2.4.1.

Multivariate analysis was performed as described in section 2.5 in an attempt to identify possible biomarkers for pancreaticobiliary malignancy.

3.1.1 Principle component analysis of the whole plasma dataset

A PCA model for the whole sample cohort produced six PC's. The scores plot of PC 1 versus PC 2 is displayed in Figure 3.1. For visualisation purposes the CR samples are highlighted blue and the BR samples green. The goodness of fit (R^2**X**(cum), the fraction of the sum of squares of all the X-variables that are explained by the model) is 0.83. The predictive ability (Q^2**Y**(cum), the fraction of the total variation of the X-variables that can be predicted by the model) is 0.695. A difference of more than 0.3 between R^2**X** and Q^2**Y** is indicative of a poor model due to either noise or outlying data points. Five CR samples (CR1, CR12, CR13, CR20 and CR 38) lie outside of Hotelling's T^2 confidence interval. Reviewing the medical records of these patients (Table 2-2 and Table 2-4) failed to identify any explanation for the outlying nature of sample CR12. Patient CR20 had however within one-month undergone colorectal surgery (subtotal colectomy) for Duke's B sigmoid and ascending colon carcinoma which may well account for the variation in metabolome. Of interest patients CR1, CR13, and CR38 were all taking levothyroxine for hypothyroidism. Of note patients CR43 and BR21

were also taking levothyroxime however both are clustered with other BR and CR samples.

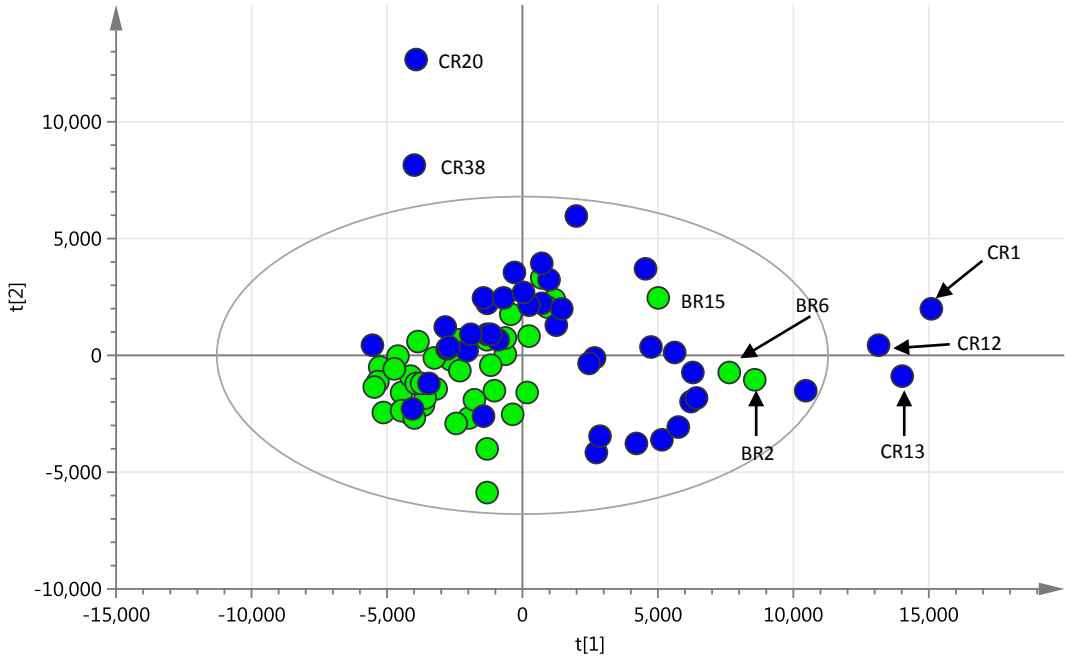

Figure 3.1 PCA scores plot for all 43 cancer (blue) and 42 benign (green) plasma samples showing the first two model components. R^2X = 0.406 and 0.147, and Q^2Y = 0.385 and 0.196 for PC 1 and PC 2, respectively

Through focusing on samples lying within the 95% confidence interval the BR samples appear to be cluster together toward the left hand side of the scores plot. In contrast the CR samples appear to cluster within two discrete areas on the scores plot. CR clusters 1 and 2 are surrounded by black ellipses for ease of viewing alone (Figure 3.2). Three BR samples, BR2, BR6 and potentially BR15, although within the 95% confidence interval appear separate from the main BR cluster. Upon further clinical review, BR6 although simplistically classed as benign is actually pre-malignant with final histology that of IPMN with moderate dysplasia. Histology from patient BR2 identified type 2 autoimmune pancreatitis, a rare form of pancreatitis which may mimic the clinical presentation of pancreatic cancer. Patient BR2 was also of Pakistani ethnicity as opposed to the predominantly Caucasian

patients recruited. Ethnicity is a known confounding factor in metabonomics and will be accounted for in subsequent analysis (138).

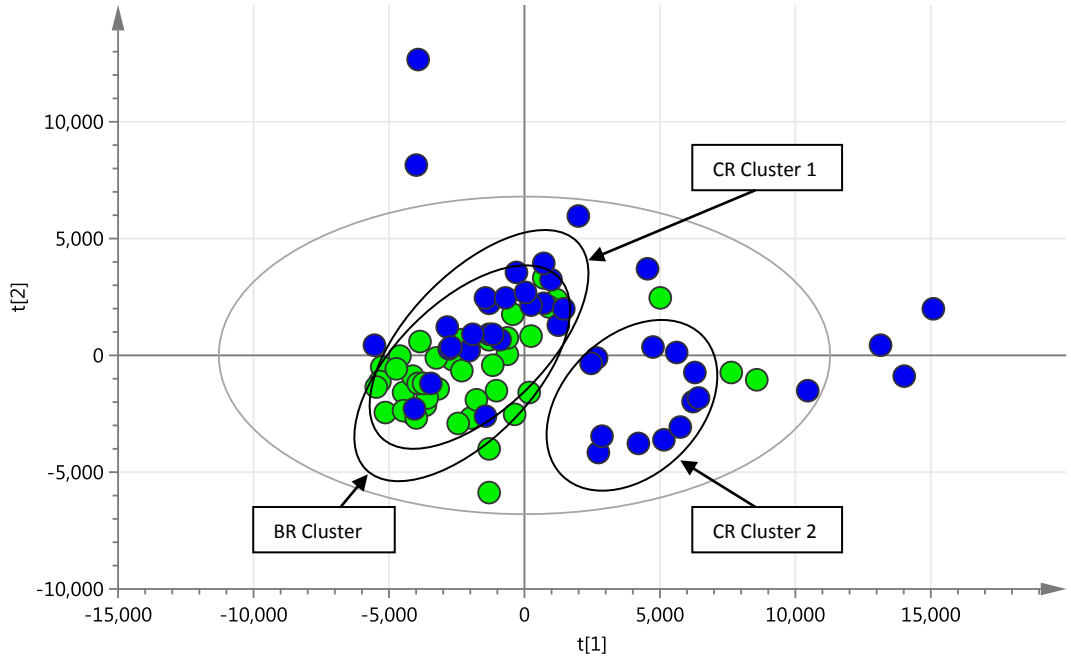

Figure 3.2 PCA scores plot for all 43 cancer and 42 benign plasma samples displaying the first two PC and highlighting clustering of benign (green) and malignant (blue) samples

3.1.2 Principle component analysis for potential confounding variables

In further PCA scores plots there is an attempt to account for true separation between CR and BR samples through testing for potential variables such as sex, ethnicity, pre-operative serum bilirubin, tumour type, resectability, tumour aggressiveness (recurrence and survival data) and the potential effect of various subgroups within the "benign" patient cohort such as gallstones, chronic pancreatitis, autoimmune pancreatitis, benign and pre-malignant pancreatic cystic lesions.

3.1.2.1 Ethnicity as a potential confounding factor

As displayed in Table 2-1 and Table 2-3 the majority of both the BR and CR patients are of Caucasian origin. Among the CR cohort one patient (CR13) is of Pakistani origin along with three patients in the benign cohort (BR2, BR30

and BR32). Sample BR2, as described in section 3.1, although within the 95% confidence interval is separated away from the main BR cohort. CR13 is located outside of the 95% confidence interval. In contrast BR30 and BR32 appear within the main BR cluster on the scores plot. Options for subsequent analysis are to either exclude CR13 and BR2 alone (based on ethnicity and potential levothyroxine use for CR13 and ethnicity and histology for BR2) or to exclude all four patients. Given that the remainder of the study population were of European ethnicity a decision was made to restrict subsequent analysis to the European population given the known confounding factor of ethnicity in metabonomics (138).

3.1.2.2 Benign cohort variation as a confounding factor

As described in section 3.1.1 variation exists with respect to final histological diagnosis within the BR cohort. The majority of samples were collected from patients with gallstones disease (n=31). Other patients were however recruited following pancreatic cystic lesion excision for dysplasia (n=4), benign pancreatic lesion resection (n=6), pancreaticojejunostomy for chronic pancreatitis (n=2), duodenal resection for GIST (n=1) and open excision of retroperitoneal paraganglionoma (Table 2-1 and Table 2-3). This clearly adds potential confounding factors to comparative analysis with a heterogeneous control group. This is demonstrated, as described in section 3.1.1 by the two samples, BR6 and BR2, which are seen to be separate from the main cluster of BR samples. BR2 having type 2 AIP and BR6 IPMN with moderate dysplasia on final histology. To assess for further potential variation within the BR cohort patients were further classified as either having pure benign or pre-malignant histology (Table 2-3). Of those with plasma available for analysis a total of five patients were classified as pre-malignant, four patients with IPMN (BR6, BR8, BR20 and BR33) and one who underwent duodenal resection for GIST (BR33). Interestingly upon constructing a PCA scores plot (Figure 3.3) highlighting benign BR (green), pre-malignant (blue) and cancer CR (red), despite marked heterogeneity in final histology only one sample (BR6), a PPPD for side branch IPMN with low and moderate dysplasia appeared markedly dissimilar to the main cluster of benign samples (green).

Options at this point are to either exclude sample BR6 alone or the remainder of the samples with pre-malignant histology. A decision was made to exclude all pre-malignant samples so as to compare true benign versus malignant patient metabolome. Insufficient patients were recruited into the study for meaningful subgroup analysis of pre-malignant plasma as a separate entity. Despite heterogenicity within the benign cohort a decision was made to include both patients with gallstone disease along with other benign pathology. This was to achieve comparable patient numbers, to reflect clinical practice (as patients rarely present with either gallstones or malignancy) and due to similar distribution on the scores plot.

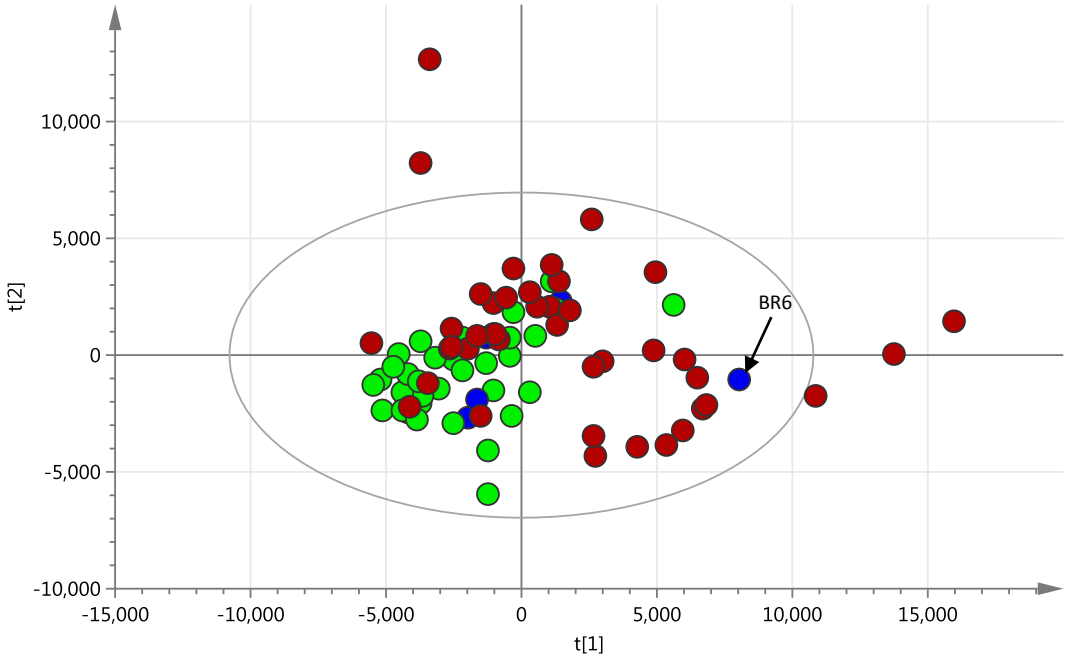

Figure 3.3 PCA scores plot of plasma data for 34 benign (green), 5 pre-malignant (blue) and 42 cancer (red) samples showing the first two PC's among a European based population

The resulting PCA scores plot for benign versus cancer plasma (excluding pre-malignant samples) among a European Population is shown in Figure 3.4. A total of thirty-four benign and forty-two cancer plasma patients were included in each group for analysis. The benign samples remained clustered towards the left hand side of the scores plot (BR cluster) and the cancer

plasmas separated into two clusters (CR cluster 1 and 2). Benign samples BR 1 and BR 15 appeared separated from the main benign cluster. Upon review no obvious past medical, operative or histological explanation could be found to explain this variation.

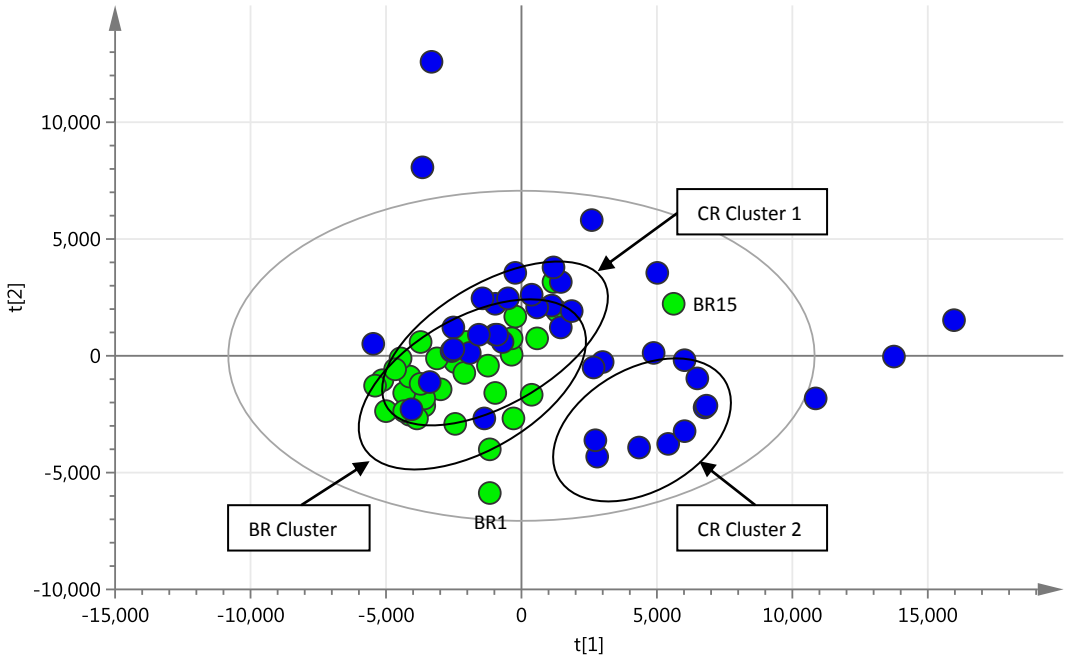

Figure 3.4 PCA scores plot of plasma data for 42 cancer (blue) and 34 benign (green) samples displaying the first two PC's among a European based population with pre-malignant condition exclusion. R^2X = 0.379 and 0.162, and Q^2Y = 0.36 and 0.201 for PC 1 and PC 2, respectively

The goodness of fit (R^2**X**(cum), the fraction of the sum of squares of all the X-variables that are explained by the model) is 0.751. The predictive ability (Q^2**Y**(cum), the fraction of the total variation of the X-variables that can be predicted by the model) is 0.609. A difference of more than 0.3 between R^2**X** and Q^2**Y** is indicative of a poor model due to either noise or outlying data points.

3.1.2.3 Age and sex as confounding factors

Among the European patient cohort patients with benign disease were younger with a mean age of 51.4 versus 65.3 years (Table 2-5). Variability in sex was found to be statistically insignificant (p=0.1657).

The PCA scores plot in Figure 3.5 demonstrates that the clustering of cancer patients between clusters 1 and 2 is not obviously due to sex.

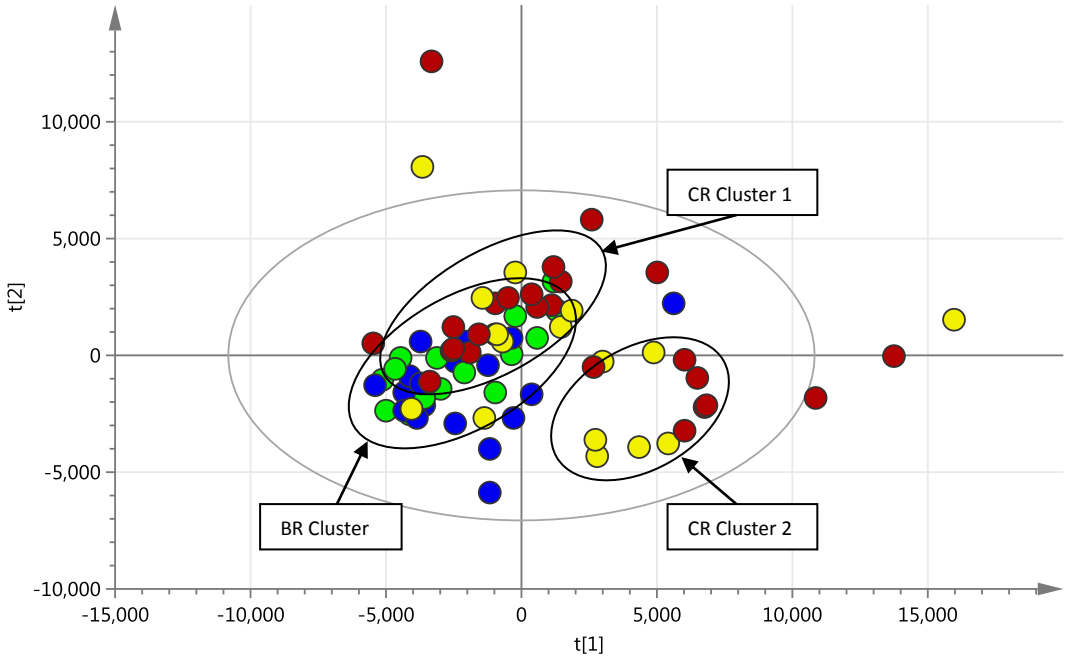

Figure 3.5 PCA scores plot of plasma data for 15 male (green) and 15 female (blue) benign and 26 male (red) and 16 female (yellow) cancer samples among a European based population with pre-malignant condition exclusion. The first two PC's are displayed

3.1.2.4 Cancer subtype and resectability as confounding factors

Cancer resectability does not appear to account for the separation of cancer plasma between clusters 1 and 2 (Figure 3.6). For cancer subtype, patients with both PDA (n=24) or "other" (n=18) cancer subtypes were seen to reside within both CR cluster 1 and 2. There was however a trend towards greater separation among the PDA cohort (Figure 3.7).

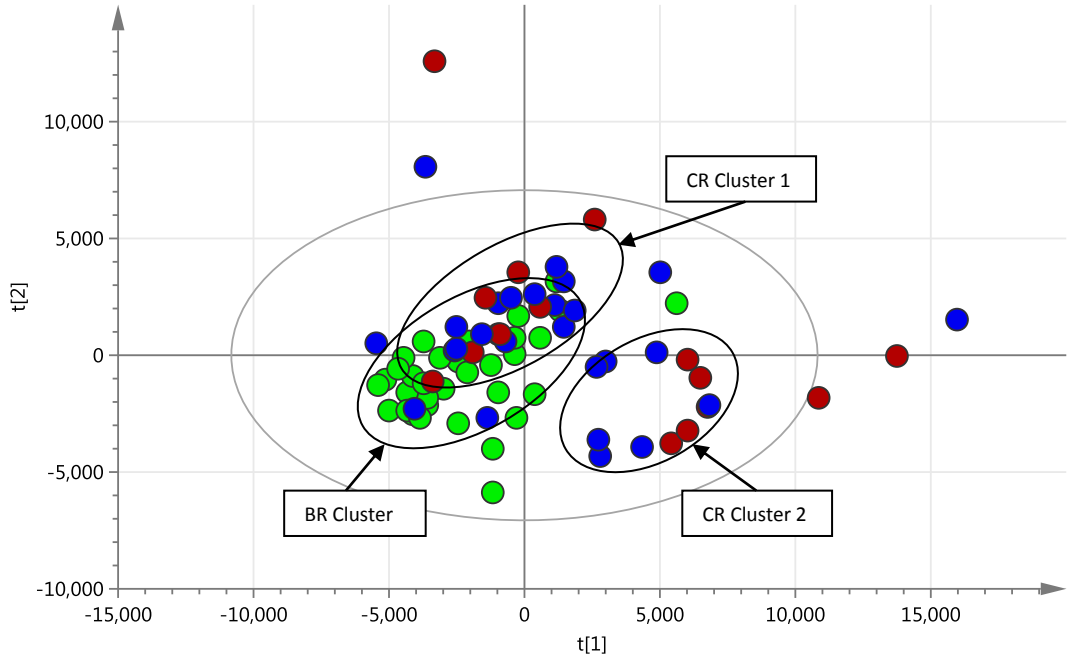

Figure 3.6 PCA scores plot of plasma data for 34 benign (green), 27 resected (blue) and 15 palliative (red) cancer patient samples among a European based population with pre-malignant condition exclusion. The first two PC's are displayed

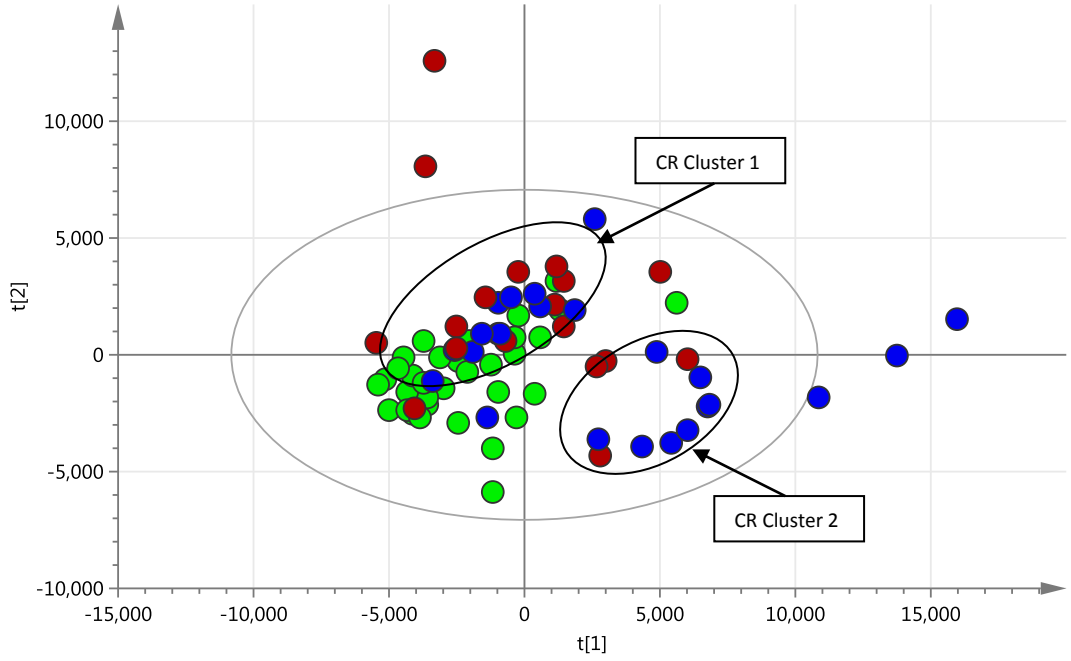

Figure 3.7 PCA scores plot of plasma data for 34 benign (green), 24 PDA (blue) and 18 "other" (red) cancer samples among a European based population with pre-malignant condition exclusion. The first two PC's are displayed

3.1.2.5 Plasma bilirubin as a confounding factor

Patients among the benign cohort had significantly lower plasma bilirubin levels pre-operatively than those with malignant disease (Table 3-1). A PCA scores plot highlighting cancer sample pre-operative bilirubin (micromol/L) variation is displayed in Figure 3.8. All cancer patients with a pre-operative plasma bilirubin greater than 40 micromol/L can be found towards the right-hand side of the scores plot whereas those with pre-operative bilirubin < 20 micromol/L appear to more closely resemble patients within the benign cohort. Figure 3.9 displays a cohort of benign samples and cancer samples with a pre-operative bilirubin < 20 and Figure 3.10 a cohort of benign samples and cancer samples with a pre-operative bilirubin > 40 micromol/L.

Table 3-1 Pre-operative plasma bilirubin (micromol/L) for benign and malignant samples among a European population with pre-malignant condition exclusion

	Cancer (n=42)	Benign (n=34)	P
Pre-op bilirubin (micromol/L) Mean +/- SD	29.1 +/- 32.9	9.3 +/- 4.8	0.0007

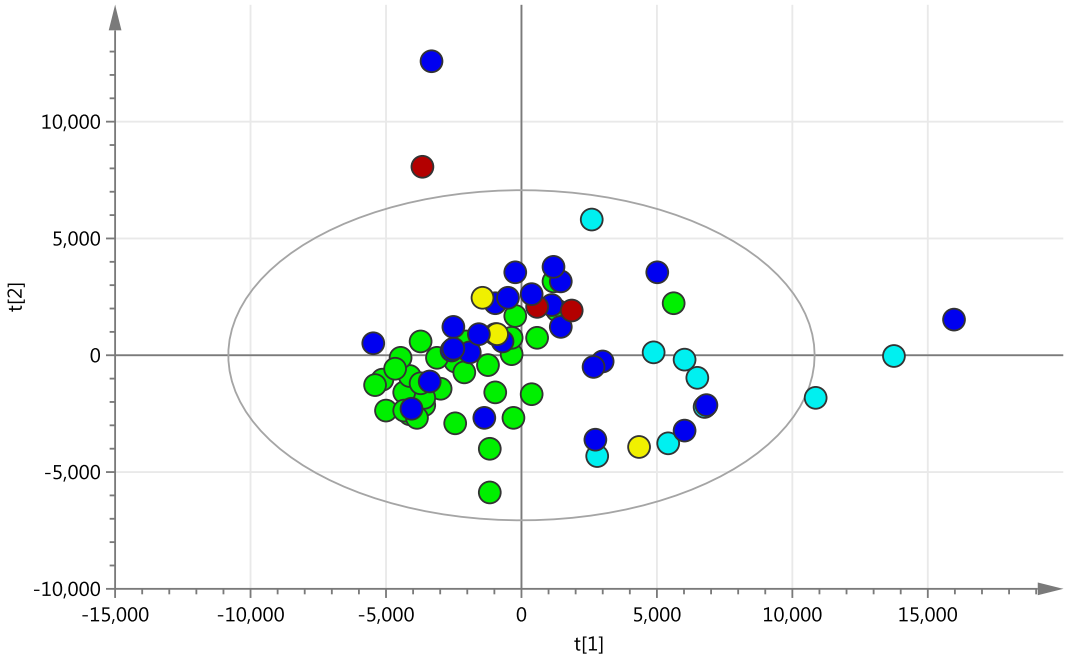

Figure 3.8 PCA scores plot of plasma data for 34 benign (green) patients, cancer samples with pre-operative plasma bilirubin < 20 (dark-blue), bilirubin 20-30 (red), 30-40 (yellow) and > 40 micromol/L (light-blue) among a European based population with pre-malignant condition exclusion

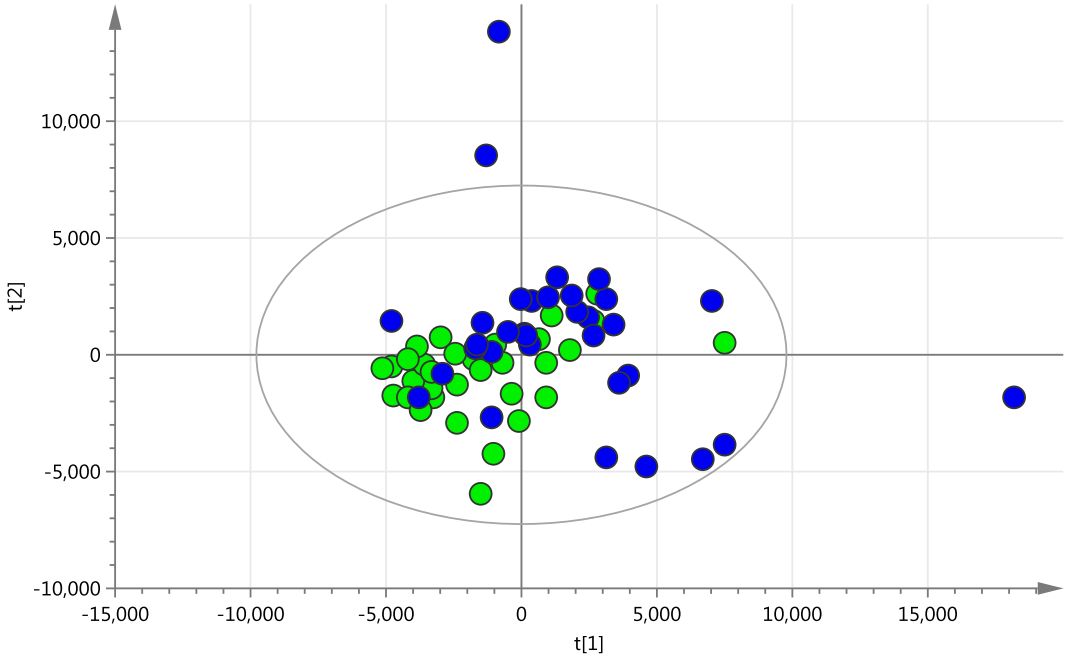

Figure 3.9 PCA scores plot of plasma data for 34 benign (green) and 33 cancer (blue) samples with a pre-operative bilirubin < 40 micromol/L among a European based population with pre-malignant condition exclusion. The first two PC's are displayed. R^2X = 0.379 and 0.162, and Q^2Y = 0.36 and 0.201 for PC 1 and PC 2, respectively

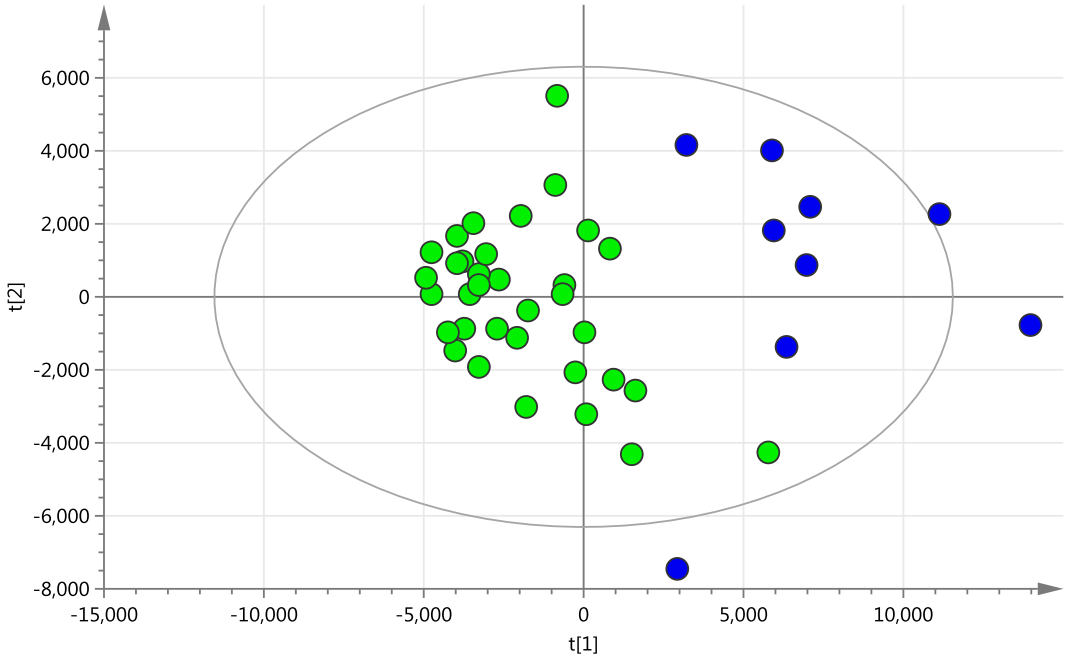

Figure 3.10 PCA scores plot of plasma data for 34 benign (green) and 9 cancer (blue) samples with a pre-operative bilirubin > 40 micromol/L among a European based population with pre-malignant condition exclusion. The first two PC's are displayed. R^2X = 0.335 and 0.184, and Q^2Y = 0.311 and 0.2 for PC 1 and PC 2, respectively

3.1.3 PLS-DA and OPLS plasma analysis

Based on the findings from initial PCA analysis as described in section 3.1.2, subsequent PLS-DA and OPLS analysis will be performed on a sample cohort of benign and cancer plasma among the European population with pre-malignant sample exclusion. Further analysis will also be performed on separate cohorts of European patients with a pre-operative bilirubin of either less-than or greater than 40 micromol/L. The PCA scores plots are displayed in Figure 3.4, Figure 3.9 and Figure 3.10 respectively.

3.1.3.1 PLS-DA analysis of benign and cancer plasma among the European population with pre-malignant sample exclusion

A PLS-DA model for the benign and cancer European plasma cohort with pre-malignant exclusion is displayed in Figure 3.11 by a two-component model.

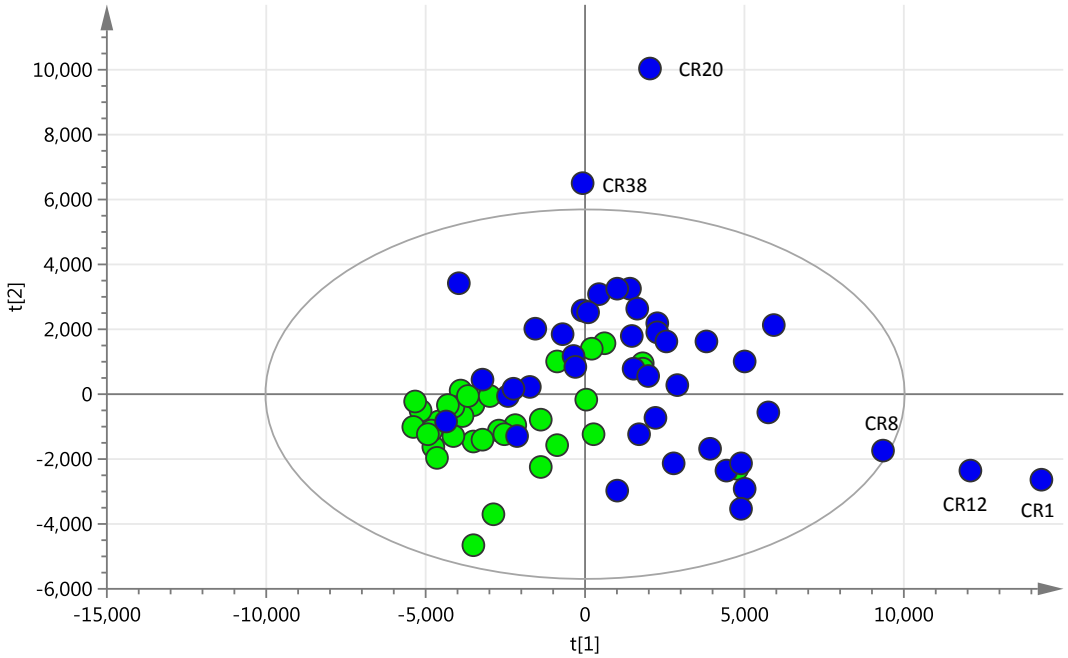

Figure 3.11 PLS-DA scores plot of plasma data for 34 benign and 42 cancer samples in a European population with pre-malignant exclusions. $R^2\boldsymbol{X}$ (cum) = 0.516, $R^2\boldsymbol{Y}$ (cum) = 0.44 and $Q^2\boldsymbol{Y}$ (cum) = 0.319

Five samples (CR1, CR8, CR12, CR20 and CR38) in Figure 3.11 lie at or beyond Hotelling's T^2 confidence interval. As described in section 3.1.1 patient numbers CR1 and CR38 were both taking prescription medication (levothyroxine) for hypothyroidism. Patient CR20 had within one month undergone subtotal colectomy for ascending and sigmoid colon adenocarcinoma. No obvious factor within the past medical or prescription history for patients CR8 and CR12 could be identified to account for significant metabolic variance. Following exclusion of extreme outliers (CR1, CR8, CR12, CR20 and CR38) a repeat PLS-DA model is generated and displayed in Figure 3.12 by a single component model.

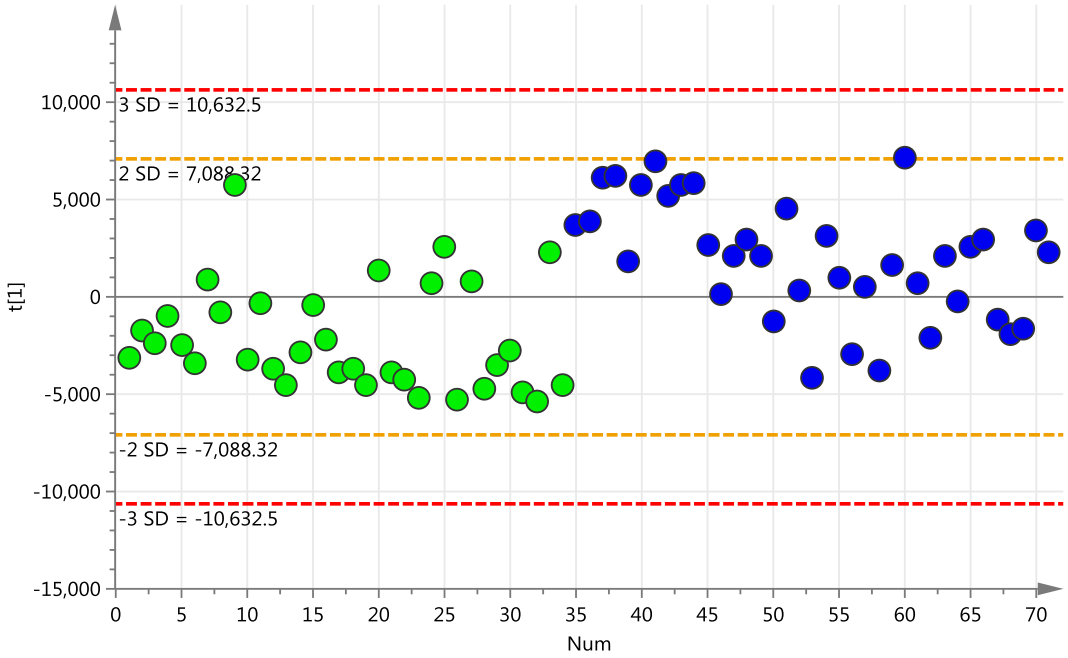

Figure 3.12 PLS-DA single component model for benign n=34 (green) and cancer n=37 (blue) plasma samples among a European population with pre-malignant and extreme outlier exclusion. $R^2\boldsymbol{X}$ (cum) = 0.318, $R^2\boldsymbol{Y}$ (cum) = 0.354 and $Q^2\boldsymbol{Y}$ (cum) = 0.308

3.1.3.1.1 Permutation testing of PLS-DA model

Permutation testing demonstrates how a model in which the Y-variables (class membership) are randomised compare to the original PLS-DA model. For the PLS-DA model displayed in Figure 3.12, all permutated $\boldsymbol{R^2}$ and $\boldsymbol{Q^2}$ values to the left were lower than the original points to the right, and lower than the original values. This indicates model validity with respect to both benign (Figure 3.13) and cancer sample (Figure 3.14) permutation.

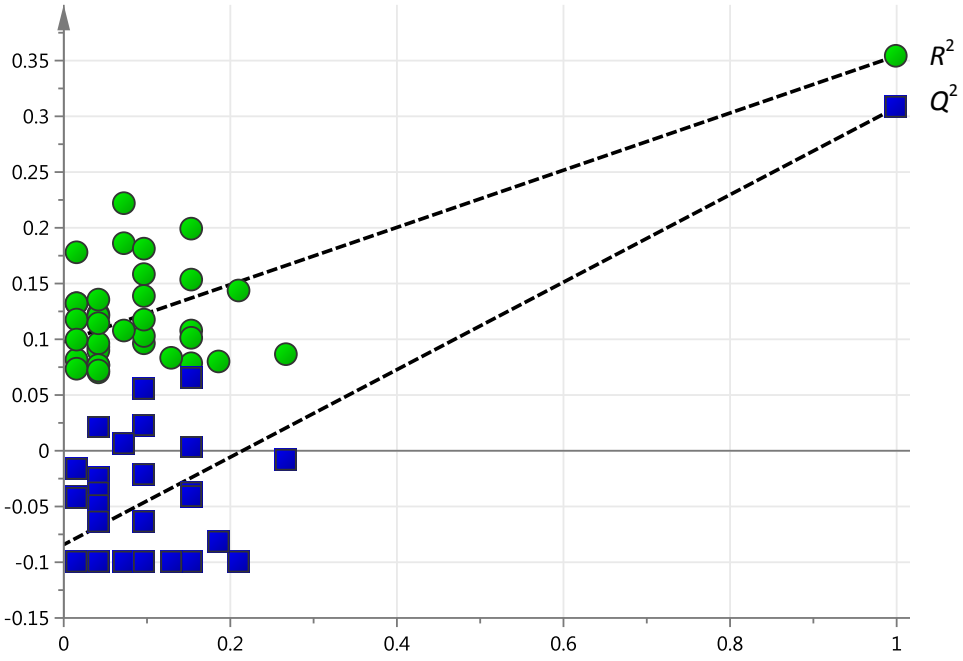

Figure 3.13 Permutation testing plot for the PLS-DA model shown in Figure 3.12. Thirty-six permutations selected with reference to thirty-four benign samples within the model. $R^2\boldsymbol{Y}$ = 0.0977 and $Q^2\boldsymbol{Y}$ -0.0842

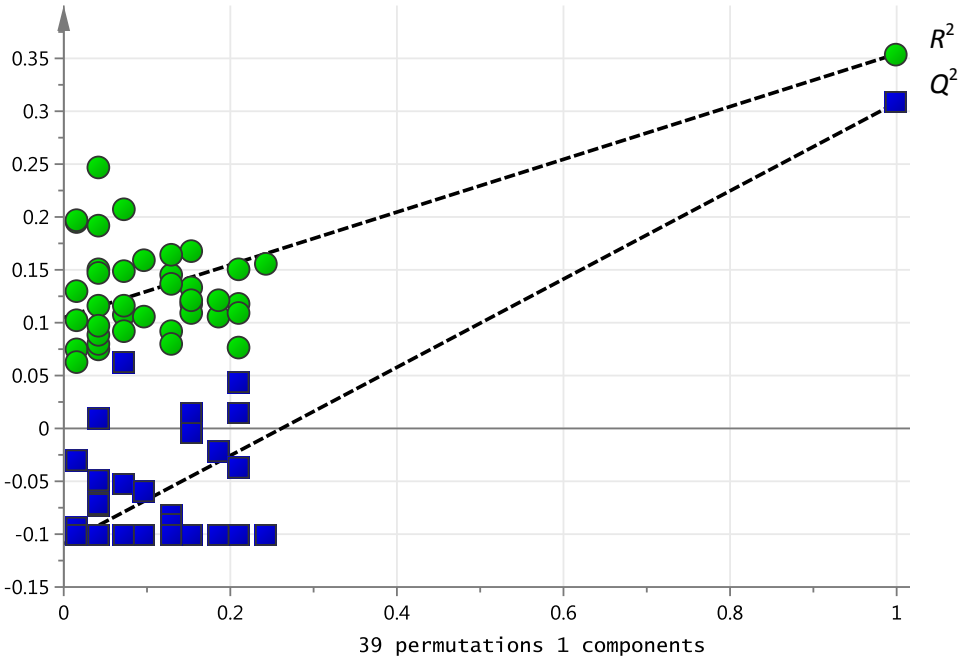

Figure 3.14 Permutation testing plot for the PLS-DA model shown in Figure 3.12. Thirty-nine permutations selected with reference to thirty-seven cancer samples within the model. $R^2\boldsymbol{Y}$ = 0.105 and $Q^2\boldsymbol{Y}$ -0.109

3.1.3.1.2 Cross validation of PLS-DA models

Cross validation of PLS-DA or OPLS models may be achieved in a number of ways, including by testing if new additional samples are correctly classified. As no further appropriate samples were available it was decided to internally validate by sequentially removing one-third of cancer and benign data sets - building a model on the remaining two-thirds of samples and testing how the “removed” one-third are classified when re-introduced (Table 3-2). The resulting model sensitivity, specificity and predictive values are displayed (Table 3-3).

Table 3-2 Cross validation of the model displayed in Figure 3.12

*A sample was regarded as belonging to a class with a Y predicted value of >0.50. Green indicates correct and red incorrect classification respectively

Sample	Y Predicted*	
	Benign	Cancer
BR1	0.91	0.09
BR3	0.74	0.26
BR4	0.67	0.33
BR9	0.61	0.39
BR10	0.49	0.51
BR11	0.63	0.37
BR12	0.30	0.70
BR14	0.53	0.47
BR15	-0.10	1.10
BR16	0.68	0.32
BR17	0.38	0.62
CR2	0.37	0.63
CR3	0.19	0.81
CR5	0.15	0.85
CR6	0.55	0.45
CR7	0.18	0.82
CR9	0.06	0.94
CR10	0.23	0.77
CR11	0.26	0.74
CR14	0.20	0.80
CR15	0.09	0.91
CR16	0.44	0.56
CR17	0.38	0.62

Sample	Y Predicted*	
	Benign	Cancer
BR18	0.85	0.15
BR19	0.88	0.12
BR21	0.72	0.28
BR22	0.58	0.42
BR23	0.73	0.27
BR24	0.83	0.17
BR25	0.84	0.16
BR26	0.92	0.08
BR27	0.44	0.56
BR29	0.85	0.15
BR31	0.90	0.10
CR18	0.46	0.54
CR19	0.30	0.70
CR21	0.43	0.57
CR22	0.67	0.33
CR23	0.28	0.72
CR24	0.53	0.47
CR25	0.87	0.13
CR26	0.32	0.68
CR27	0.51	0.49
CR29	0.80	0.20
CR28	0.57	0.43
CR30	0.88	0.12

Sample	Y Predicted*	
	Benign	Cancer
BR34	0.93	0.07
BR35	0.42	0.58
BR36	0.31	0.69
BR37	0.93	0.07
BR38	0.45	0.55
BR39	0.87	0.13
BR40	0.78	0.22
BR41	0.74	0.26
BR42	0.92	0.08
BR43	0.92	0.08
BR44	0.33	0.67
BR45	0.85	0.15
CR31	0.40	0.60
CR32	-0.03	1.03
CR33	0.46	0.54
CR34	0.71	0.29
CR35	0.37	0.63
CR36	0.55	0.45
CR37	0.34	0.66
CR39	0.32	0.68
CR40	0.64	0.36
CR41	0.68	0.32
CR42	0.62	0.38
CR43	0.25	0.75
CR44	0.34	0.66

Table 3-3 Model sensitivity, specificity, positive and negative predictive values

Test	Cancer		
	Present (37)	**Absent (34)**	**Total**
Positive	True positive (24)	False positive (9)	**33**
Negative	False negative (13)	True negative (25)	**38**

	%	**95% C.I**
Sensitivity	64.9	47.5-79.8
Specificity	73.5	55.64-87.1
Positive predictive value	72.7	54.5-86.7
Negative predictive value	65.8	48.6-80.3

3.1.3.1.3 OPLS loading plot and metabolite identification

A loading plot generated through OPLS from the model shown in Figure 3.12 is displayed in Figure 3.15. The loadings pq plot is a superimposition of the p plot and the q plot, for the first component of the OPLS model. The loading vector p1 corresponds to the co-variances between the X-variables and the score vector t1, whereas the loading vector q1 expresses the importance of the variables in approximating Y variation correlated to X, in the first component. The X-axis on the loadings plot corresponds with the X-axis (in ppm) on the NMR spectrum. Variation in the loadings plot can therefore be directly related to the NMR spectrum and hence metabolite identification. Distribution above zero on the Y-axis in Figure 3.15 represents up-regulation of signal among cancer samples and vice versa.

Due to the use of the intelligent bucketing process for spectral binning (to minimize metabolite loss) the resultant OPLS loadings plots fail to align perfectly with the original NMR spectra.

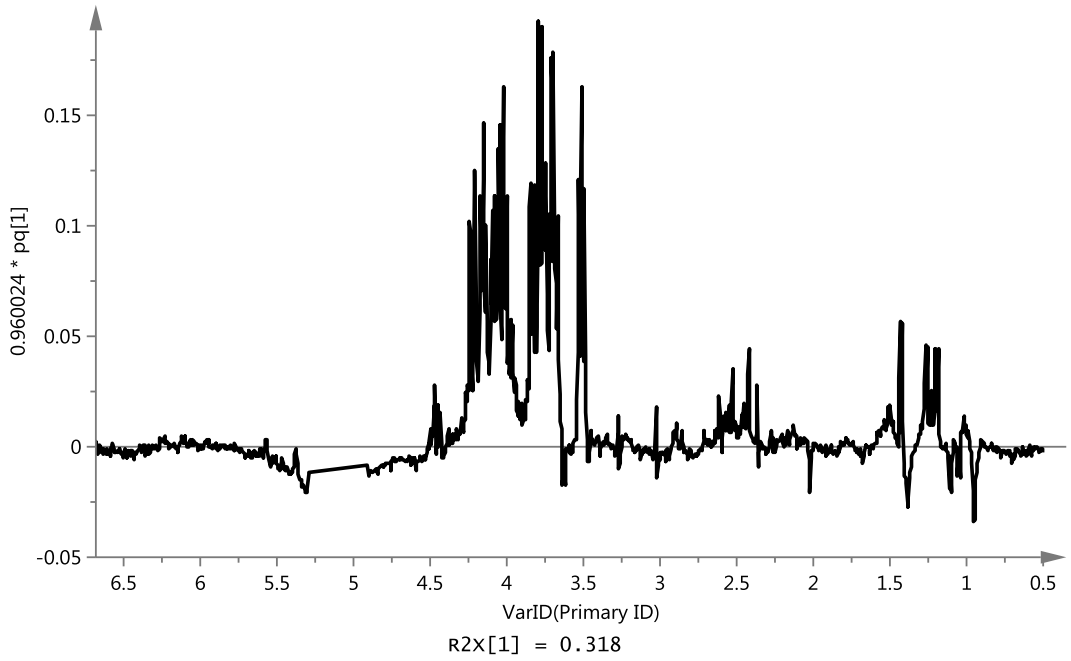

Figure 3.15 OPLS Loadings plot for the model displayed in Figure 3.12

The loadings plot shown in Figure 3.15 is magnified to highlight chemical shift range 0.9 – 2.1 ppm (Figure 3.16). The corresponding region on a cancer NMR spectrum (CR16 as an example) is magnified in Figure 3.17. Example metabolites are highlighted in both Figure 3.16 and Figure 3.17 through reference to a NMR database of plasma metabolite chemical shifts and peak multiplicity (139).

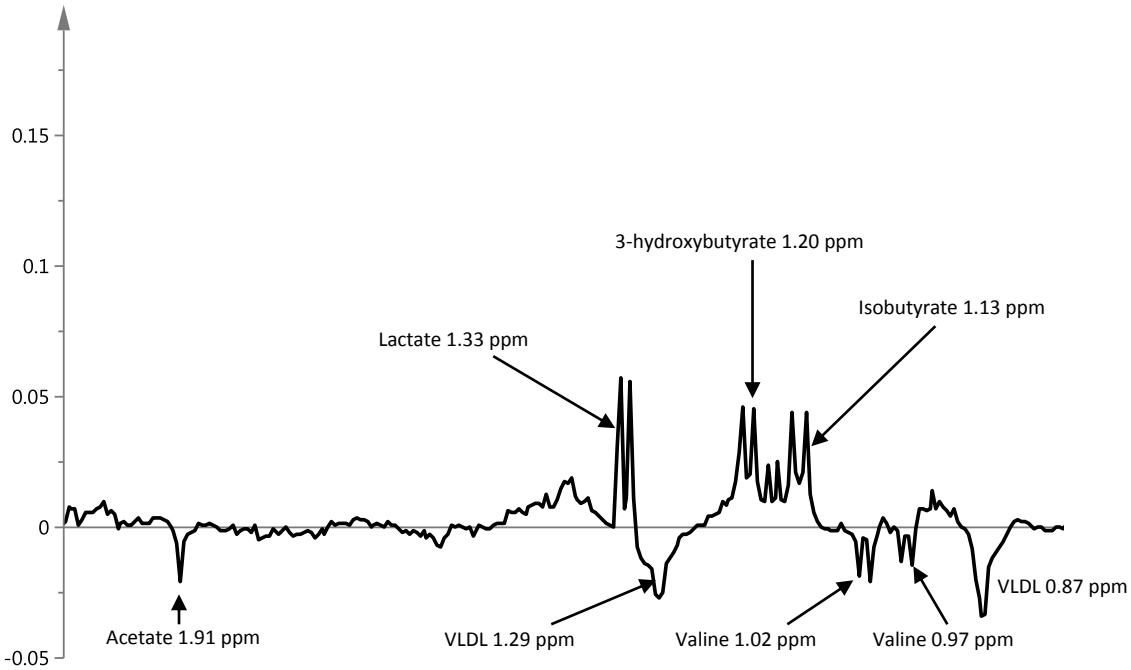

Figure 3.16 OPLS Loadings plot for the model displayed in Figure 3.12 with expansion of 0.8–2.0 ppm chemical shift

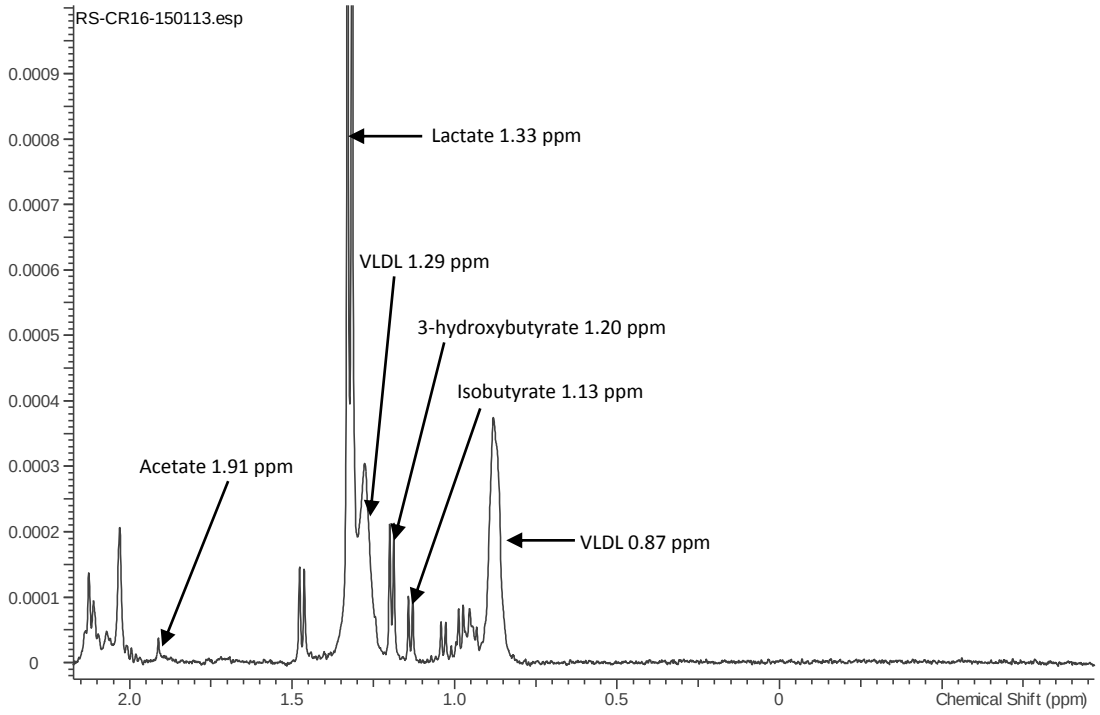

Figure 3.17 NMR spectrum for sample CR16 with expansion of chemical shift 0-2.5 ppm

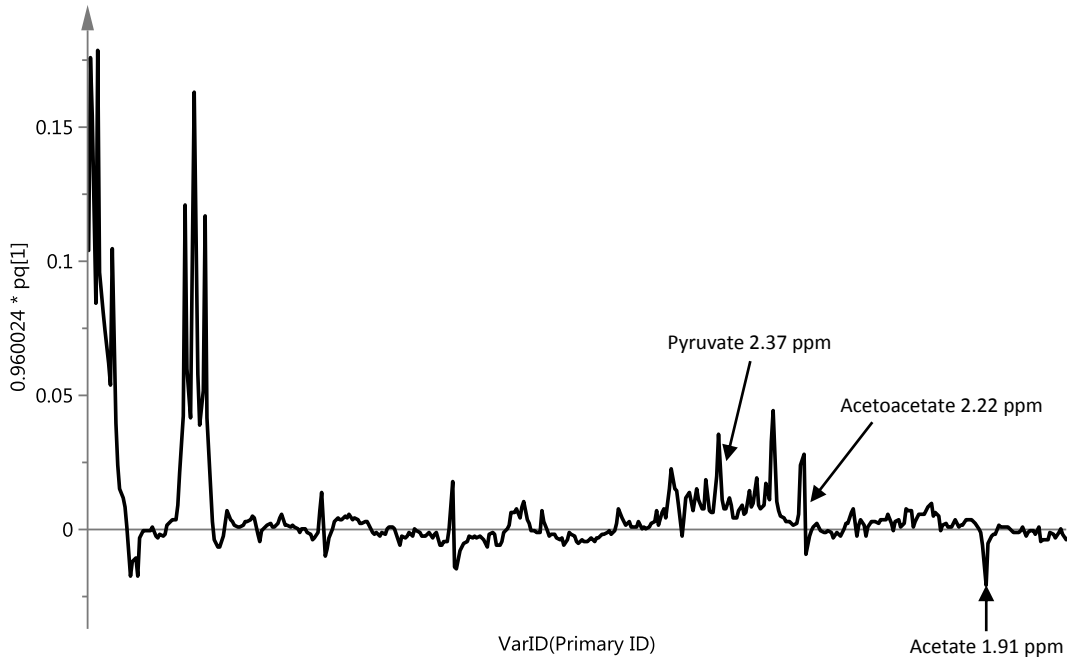

Figure 3.18 OPLS Loadings plot for the model displayed in Figure 3.12 with expansion of 1.8–3.0 ppm chemical shift

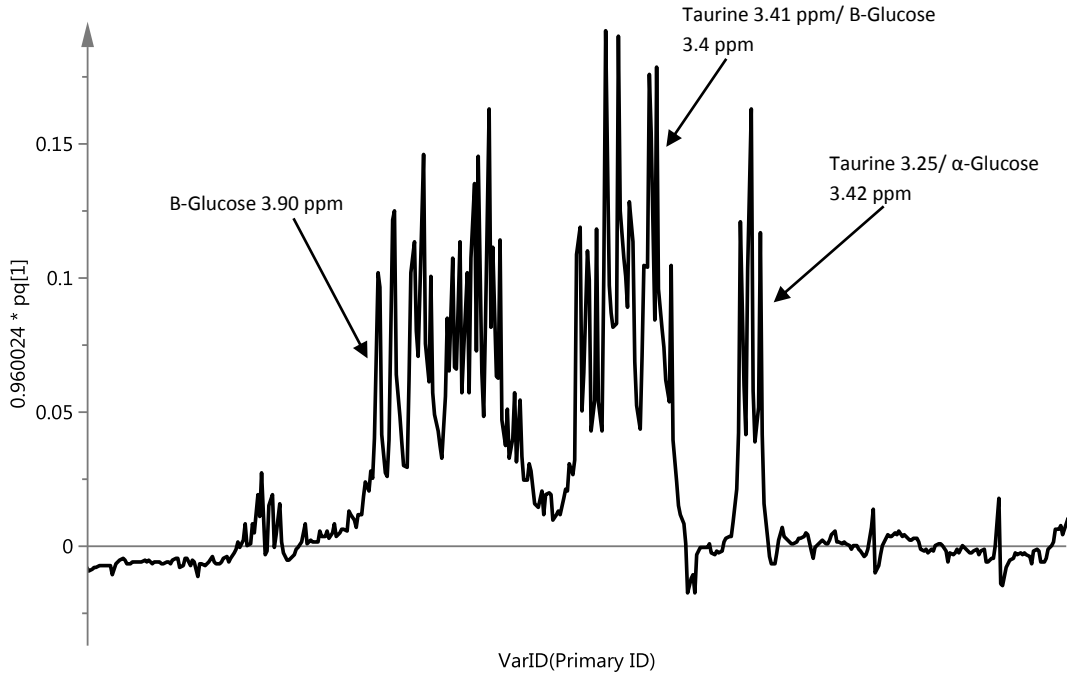

Figure 3.19 OPLS Loadings plot for the model displayed in Figure 3.12 with expansion of 3.0–4.5 ppm chemical shift

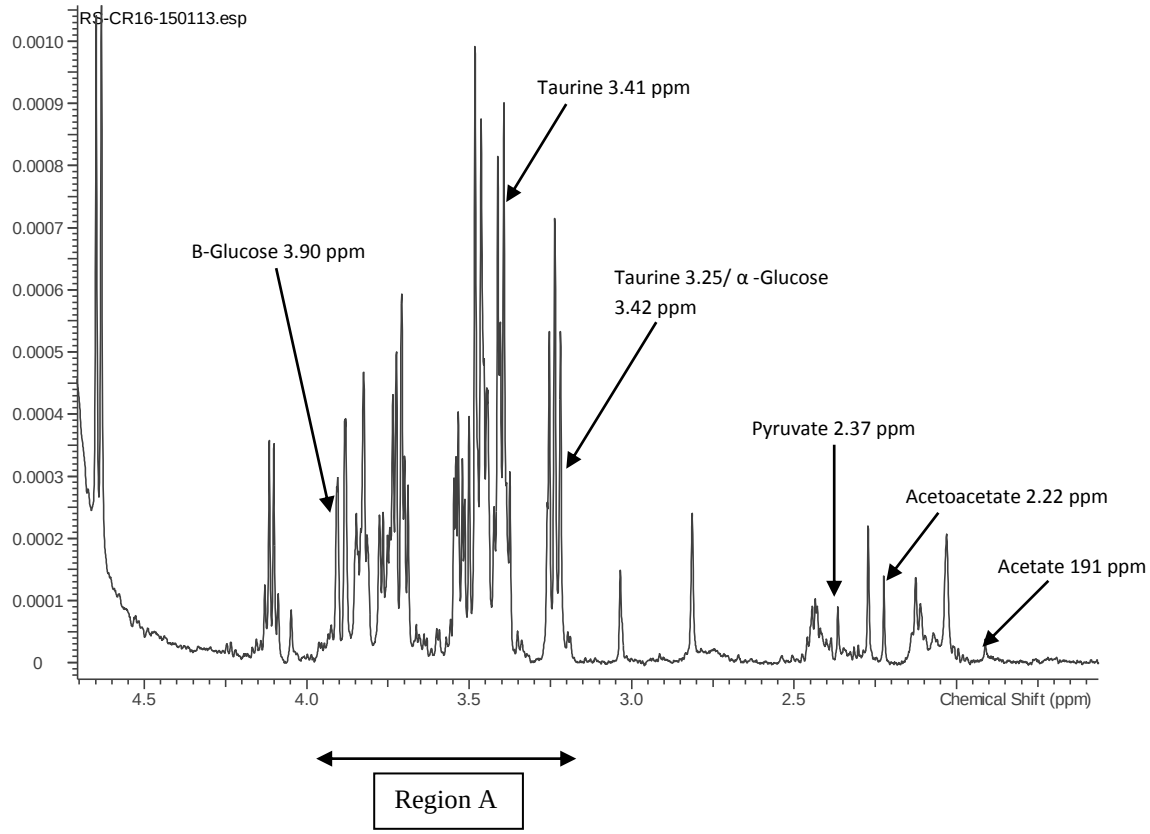

Figure 3.20 NMR spectrum for sample CR16 with expansion of chemical shift 1.9-4.5 ppm

A summary of the assigned metabolites across the entire spectral range is displayed within Table 3-3. Region A as displayed in Figure 3.20 is particularly complex making it difficult to unambiguously assign metabolites by 1D-NMR.

Table 3-4 Metabolite identification for model displayed in Figure 3.12

H chemical shift (ppm)	Multiplicity	Metabolite	Metabolite change in cancer
0.87	t	Lipid (mainly VLDL)	↓
0.97	d	Valine	↓
1.02	d	Valine	↓
1.13	d	Isobutyrate	↑
1.2	d	3-hydroxybutyrate	↑
1.29	m	Lipid (mainly VLDL)	↓
1.33	d	Lactate	↑
1.91	m	Acetate	↓
2.22	s	Acetoacetate	↑
2.37	s	Pyruvate	↑
3.25/3.42	t	Taurine/ α-glucose*	↑
3.40/3.41	t	β-glucose/Taurine*	↑
3.9	dd	β-glucose	↑

*2D-NMR required to confirm metabolite assignment

3.1.3.2 PLS-DA analysis of benign and cancer plasma with a pre-operative bilirubin less than 40 micromol/L

A PLS-DA model for benign and cancer plasma samples among a European population with a pre-operative bilirubin of less-than 40 micromol/L with pre-malignant exclusion is displayed below (Figure 3.21). One sample (CR1) was excluded as an extreme outlier. The PLS-DA model generated was explained by a single component (Figure 3.22).

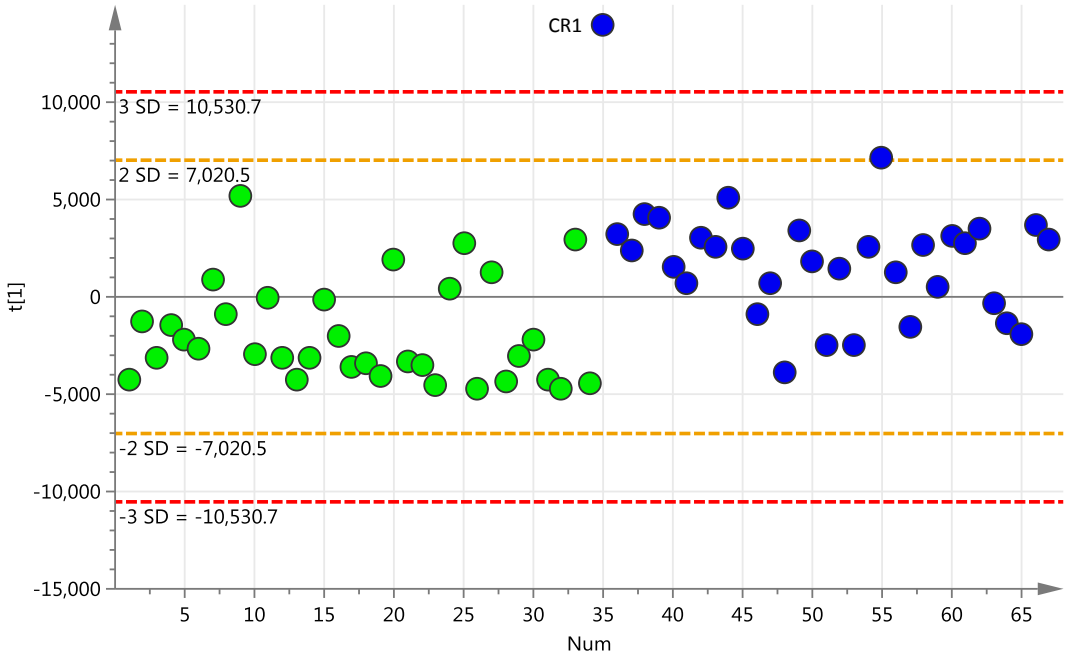

Figure 3.21 PLS-DA scores plot of plasma data for 34 benign (green) and 33 cancer (blue) samples with a pre-operative bilirubin of less-than 40 micromol/L with pre-malignant exclusions. $R^2\boldsymbol{X}$ (cum) = 0.311, $R^2\boldsymbol{Y}$ (cum) = 0.32 and $Q^2\boldsymbol{Y}$ (cum) = 0.265

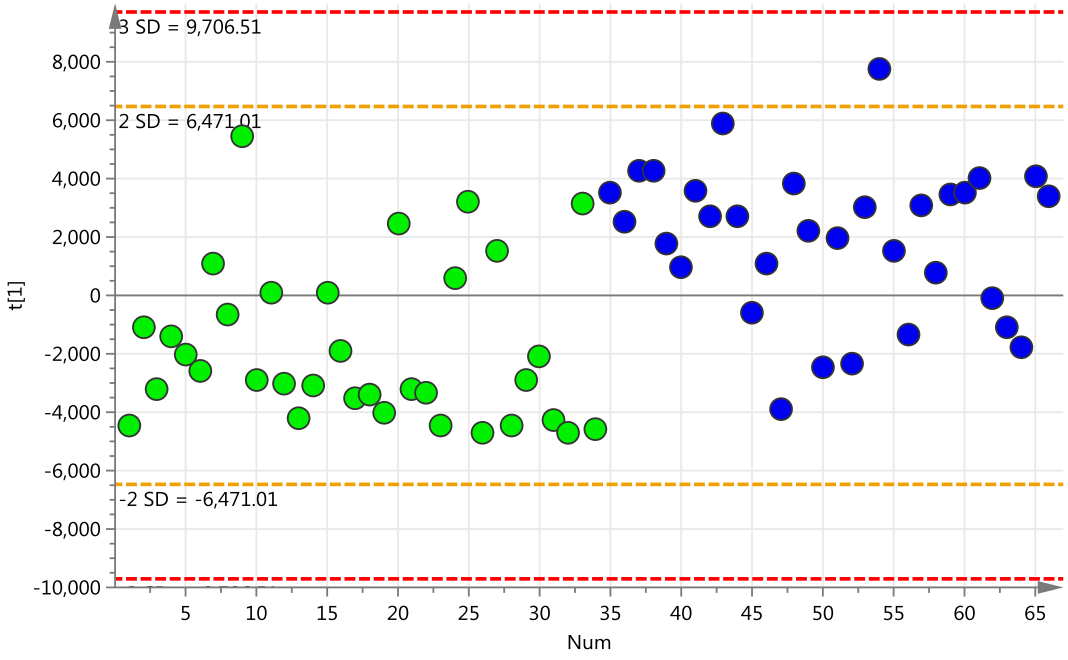

Figure 3.22 PLS-DA scores plot of plasma data for 34 benign (green) and 32 cancer (blue) samples with a pre-operative bilirubin of less than 40 micromol/L with exclusions. R^2X (cum) = 0.272, R^2Y (cum) = 0.348 and Q^2Y (cum) = 0.302

3.1.3.2.1 Permutation testing of PLS-DA model

Validity of the model displayed in Figure 3.22 was demonstrated through permutation of both benign (Figure 3.23) and cancer (Figure 3.24) sample class membership.

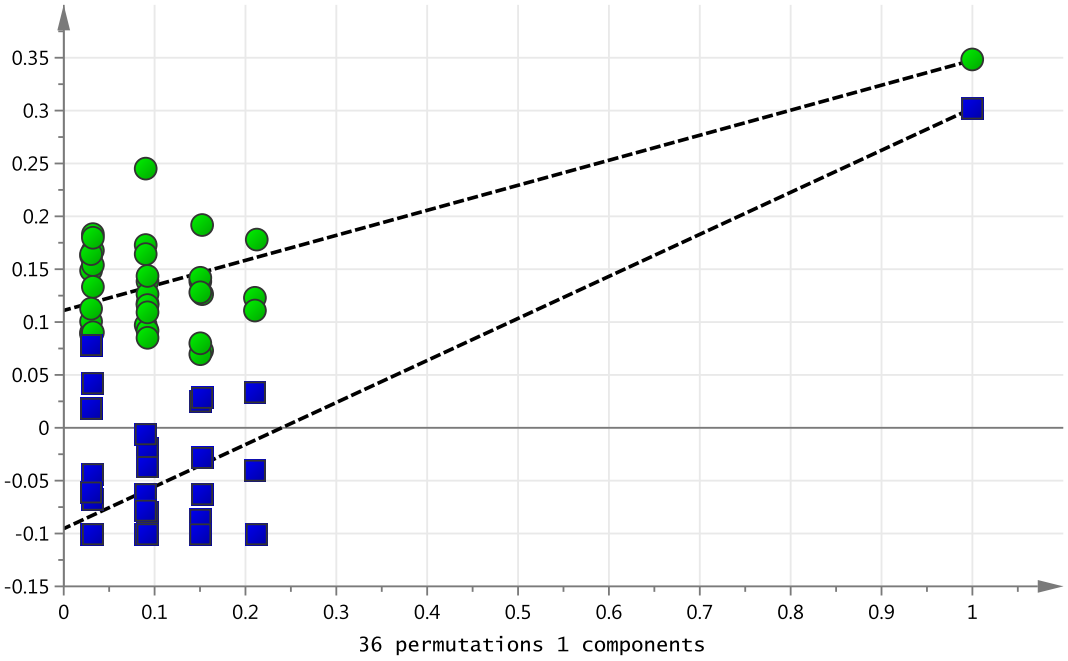

Figure 3.23 Permutation testing plot for the PLS-DA model shown in Figure 3.22. Thirty-six permutations selected with reference to thirty-four benign samples within the model. $R^2\boldsymbol{Y}$ = 0.111 and $Q^2\boldsymbol{Y}$ = -0.0954

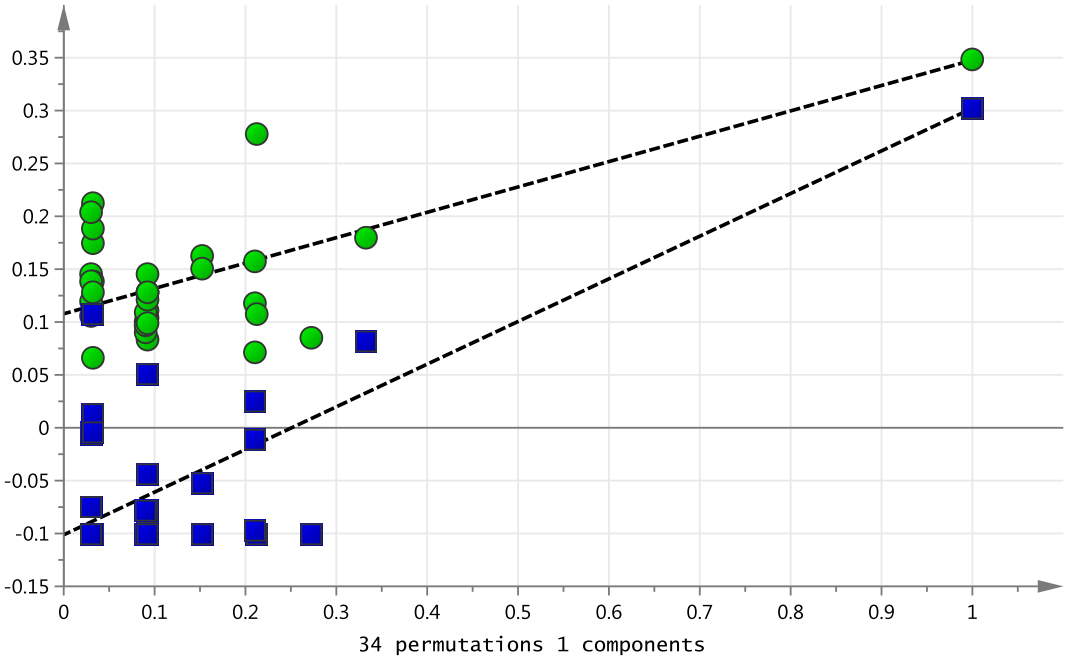

Figure 3.24 Permutation testing plot for the PLS-DA model shown in Figure 3.22. Thirty-four permutations selected with reference to thirty-two cancer samples within the model. $R^2\boldsymbol{Y}$ = 0.108 and $Q^2\boldsymbol{Y}$ = -0.111

3.1.3.2.2 Cross section model validation

Cross validation of the model is performed (Table 3-5) as described in section 3.1.3.1.2 for the PLS-DA model displayed in Figure 3.22. The resultant sensitivity and specificity of the model is displayed in Table 3-6.

Table 3-5 Cross validation of the model displayed in Figure 3.22

*A sample was regarded as belonging to a class with a Y predicted value of >0.50. Green indicates correct and red incorrect classification respectively. Class allocation was not possible for sample BR10 (blue)

	Y predict	
Sample	Benign	Cancer
BR1	0.95	0.05
BR3	0.8	0.2
BR4	0.7	0.3
BR9	0.64	0.36
BR10	0.5	0.5
BR11	0.66	0.34
BR12	0.33	0.67
BR14	0.55	0.45
BR15	-0.01	1.01
BR16	0.7	0.3
BR17	0.41	0.59
BR18	0.74	0.26
CR10	0.27	0.73
CR2	0.42	0.58
CR7	0.28	0.72
CR14	0.26	0.74
CR16	0.48	0.52
CR17	0.39	0.61
CR18	0.16	0.84
CR19	0.34	0.66
CR20	-0.1	1.1
CR21	0.25	0.75
CR22	0.55	0.45

	Y predict	
Sample	Benign	Cancer
BR19	0.97	0.03
BR21	0.82	0.18
BR22	0.55	0.45
BR23	0.75	0.25
BR24	0.9	0.1
BR25	0.9	0.1
BR26	0.98	0.02
BR27	0.3	0.7
BR29	0.87	0.13
BR31	0.87	0.13
BR34	1.03	-0.03
CR24	0.45	0.55
CR25	0.95	0.05
CR26	0.21	0.79
CR27	0.4	0.6
CR28	0.39	0.61
CR29	0.85	0.15
CR30	0.81	0.19
CR31	0.31	0.69
CR32	-0.2	1.2
CR33	0.47	0.53
CR34	0.69	0.31

	Y predict	
Sample	Benign	Cancer
BR35	0.41	0.59
BR36	0.22	0.78
BR37	0.95	0.05
BR38	0.34	0.66
BR39	0.92	0.08
BR40	0.78	0.22
BR41	0.72	0.28
BR42	0.95	0.05
BR43	0.94	0.06
BR44	0.19	0.81
BR45	0.93	0.07
CR35	0.23	0.77
CR36	0.44	0.56
CR37	0.22	0.78
CR38	0.27	0.73
CR39	0.19	0.81
CR40	0.54	0.46
CR41	0.64	0.36
CR42	0.67	0.33
CR43	0.14	0.86
CR44	0.19	0.81

Table 3-6 Model sensitivity, specificity, positive and negative predictive values

Test	Cancer		Total
	Present (32)	Absent (34)	
Positive	True positive (24)	False positive (8)	32
Negative	False negative (8)	True negative (25)	33

	%	95% C.I
Sensitivity	75	56.6-88.5
Specificity	75.8	57.7-88.9
Positive predictive value	75	56.6-88.5
Negative predictive value	75.8	57.7-88.9

3.1.3.2.3 OPLS loading plot and metabolite allocation

A loadings plot generated through OPLS form the model shown in Figure 3.22 is displayed below (Figure 3.25). Metabolite allocation for this model is described in section 3.1.4.

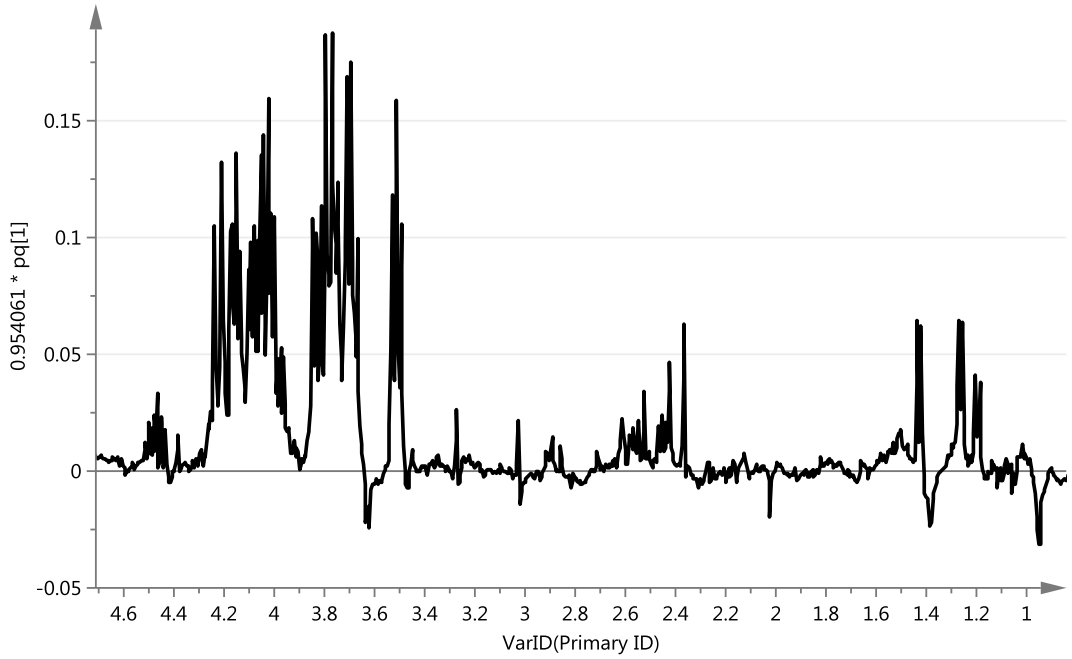

Figure 3.25 OPLS loadings plot for the model displayed in Figure 3.22

3.1.3.3 PLS-DA analysis of benign and cancer plasma with a pre-operative bilirubin greater than 40 micromol/L

Nine patients with a bilirubin greater than 40 micromol/L were included into a cohort of nine patients with benign disease selected at random. The resulting PLS-DA scores plot was explained by a single component (Figure 3.26).

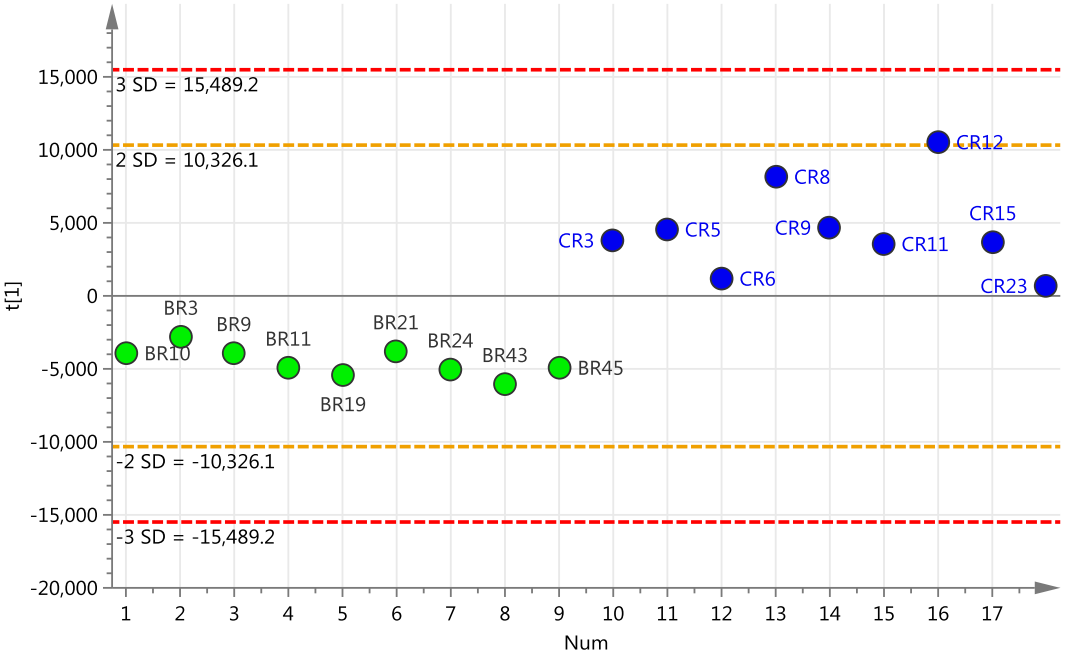

Figure 3.26 PLS-DA scores plot of plasma data for 9 benign (green) and 9 cancer (blue) samples with a pre-operative bilirubin of greater than 40 micromol/L with exclusions. $R^2\boldsymbol{X}$ (cum) = 0.498, $R^2\boldsymbol{Y}$ (cum) = 0.814 and $Q^2\boldsymbol{Y}$ (cum) = 0.763

3.1.3.3.1 Permutation testing of PLS-DA model

With an R^2Y intercept values of 0.223 and 0.18 and Q^2Yvalues of -0.172 and -0.213 with respect to benign and cancer specimens respectively the model is valid.

3.1.3.3.2 Cross section model validation

Cross validation of the model is performed (Table 3-7) as described in section 3.1.3.1.2 for the PLS-DA model displayed in Figure 3.26. The resultant sensitivity and specificity of the model is displayed in Table 3-8.

Table 3-7 Cross validation of the model displayed in Figure 3.26

*A sample was regarded as belonging to a class with a Y predicted value of >0.50. Green indicates correct and red incorrect classification respectively

Sample	Y predicted*	
	Benign	Cancer
BR3	0.75	0.25
BR9	0.80	0.20
BR10	0.79	0.21
CR3	0.26	0.74
CR5	0.18	0.82
CR6	0.44	0.56

Sample	Y predicted*	
	Benign	Cancer
BR11	0.88	0.12
BR19	0.96	0.04
BR21	0.79	0.21
CR8	-0.21	1.21
CR9	0.09	0.91
CR11	0.20	0.80

Sample	Y predicted*	
	Benign	Cancer
BR24	1.00	0.00
BR43	1.09	-0.09
BR45	0.95	0.05
CR12	-0.47	1.47
CR15	0.20	0.80
CR23	0.60	0.40

Table 3-8 Model sensitivity, specificity, positive and negative predictive values

Test	Cancer		Total
	Present (9)	Absent (9)	
Positive	True positive (8)	False positive (0)	8
Negative	False negative (1)	True negative (9)	10

	%	95% C.I
Sensitivity	88.9	51.7-98.1
Specificity	100	66.2-100
Positive predictive value	100	62.9-100
Negative predictive value	90	55.5-98.3

3.1.3.3.3 OPLS loadings plot and metabolite allocation

A loadings plot generated through OPLS form the model shown in Figure 3.26 is displayed below (Figure 3.27). Metabolite allocation for this model is described in section 3.1.4.

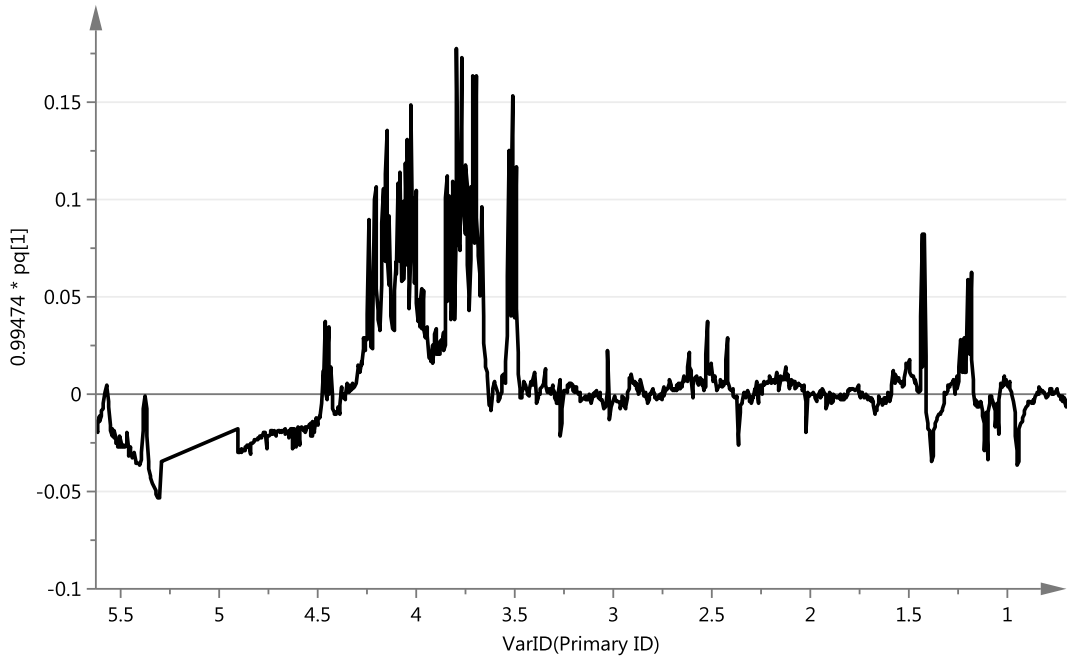

Figure 3.27 OPLS loadings plot for the model displayed in Figure 3.26

3.1.4 Metabolite comparison for overall benign and malignant samples (with exclusions) versus cancer patients with or without pre-operative jaundice

OPLS loading plots for overall benign and malignant samples, patients with a pre-operative bilirubin less than or equal to 40 micromol/L and patients with bilirubin greater than 40 micromol/L are displayed (Figure 3.28). Two regions of spectral variation (representing metabolite change) are highlighted (region one and two). Expansion of regions one and two along with metabolite assignment is displayed within Figure 3.29 and Figure 3.30, respectively.

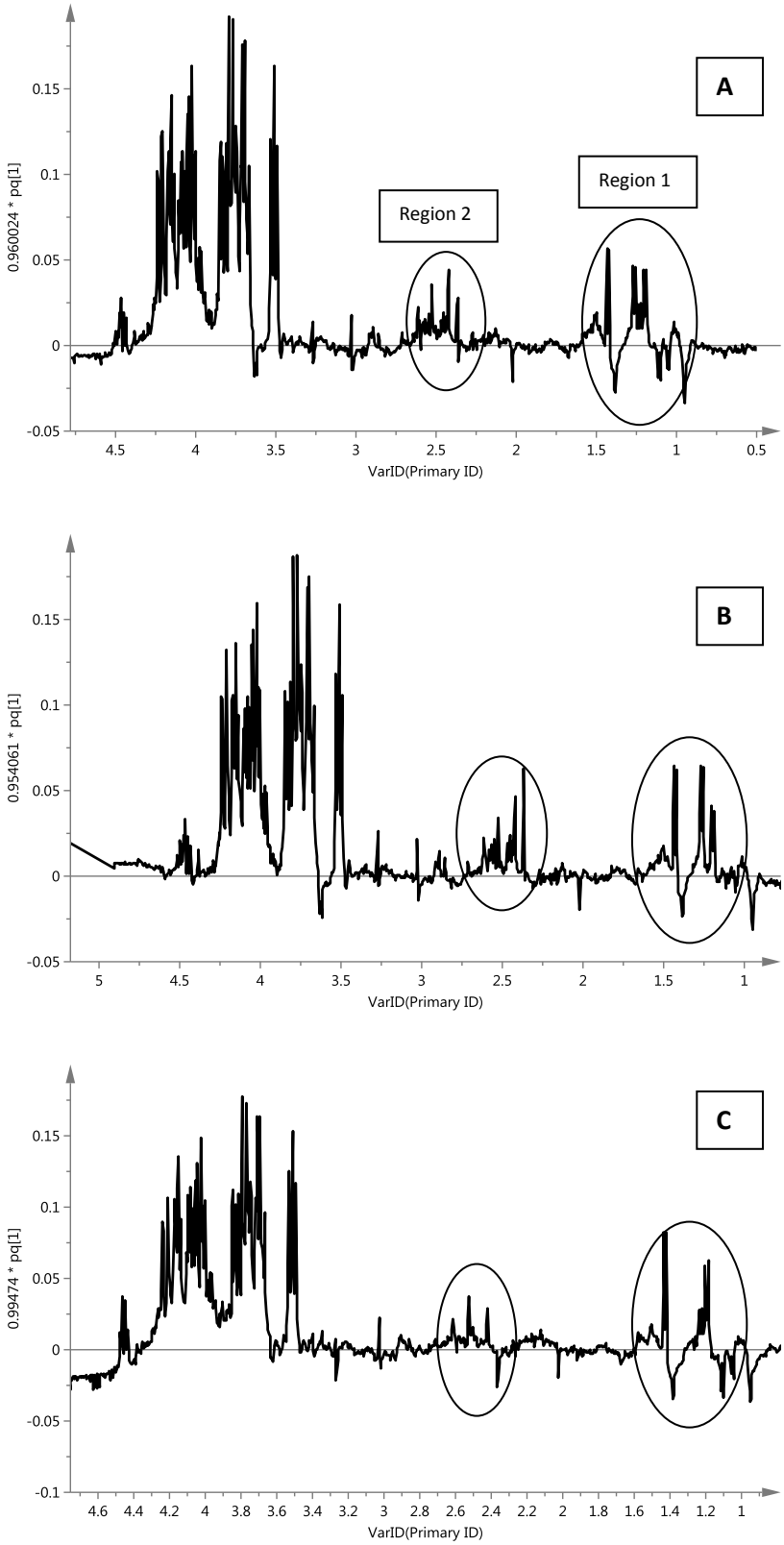

Figure 3.28 OPLS loading plots for overall benign and cancer samples (A), jaundiced (B) and non-jaundiced patients (C)

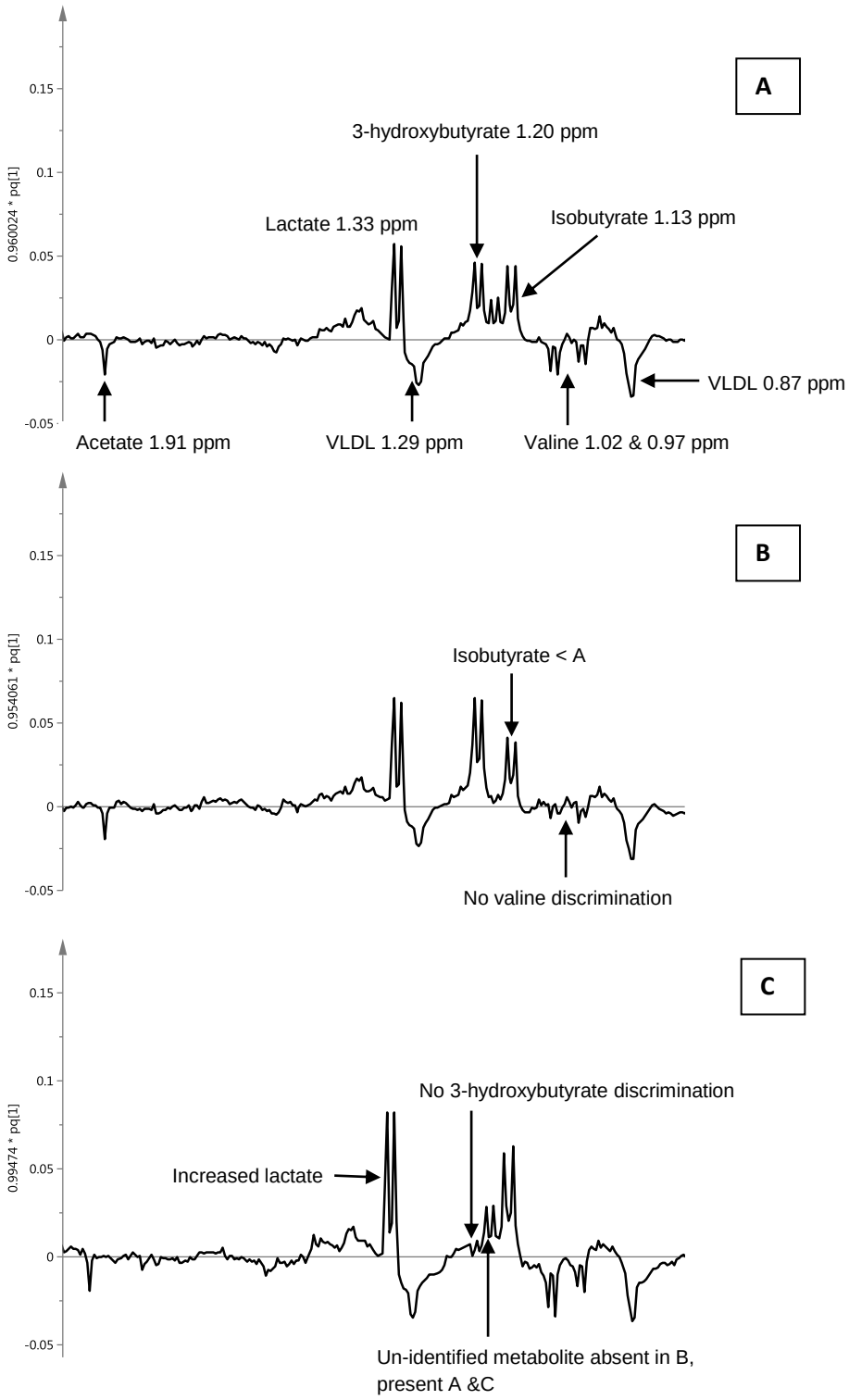

Figure 3.29 OPLS loading plots for overall benign and cancer samples (A), jaundiced (B) and non-jaundiced patients (C) with expansion of chemical shift 0.8-2.0 ppm (Region 1 Figure 3.28)

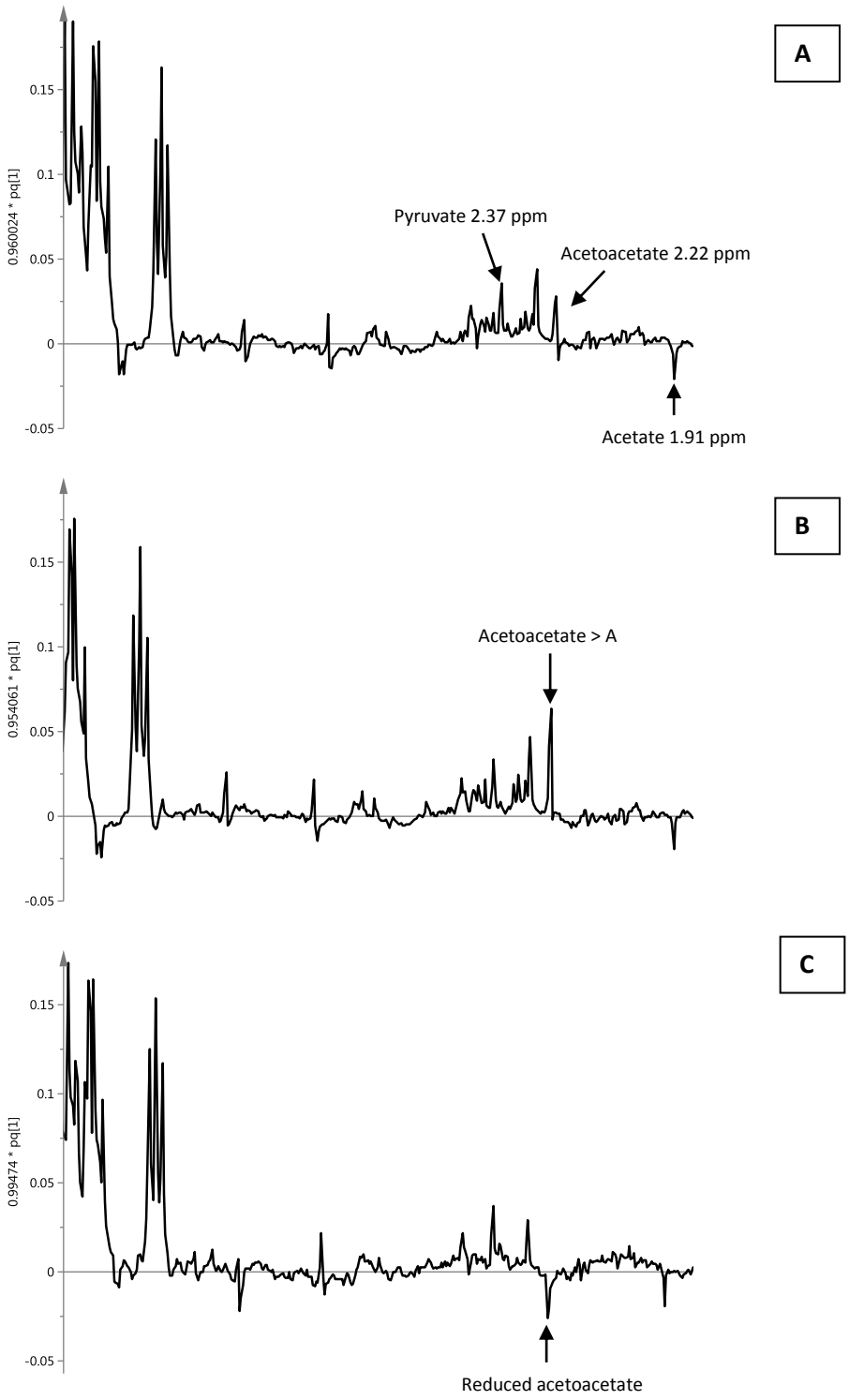

Figure 3.30 OPLS loading plots for overall benign and cancer samples (A), jaundiced (B) and non-jaundiced patients (C) with expansion of chemical shift 1.8-3.0 ppm (Region 2 Figure 3.28)

Metabolite allocation in table format is displayed in Table 3-9.

Table 3-9 Metabolite assignment for overall benign and cancer samples (with exclusions) and patients with or without pre-operative jaundice

H chemical shift (ppm)	Multiplicity	Metabolite	Metabolite change in cancer	Bil < 40micromol/L	Bil > 40 micromol/L
0.87	t	Lipid (mainly VLDL)	↓	↓	↓
0.97	d	Valine	↓	↓	NC
1.02	d	Valine	↓	↓	NC
1.13	d	Isobutyrate	↑	↑	↑
1.2	d	3-hydroxybutyrate	↑	NC	↑
1.29	m	Lipid (mainly VLDL)	↓	↓	↓
1.33	d	Lactate	↑	↑	↑
1.91	m	Acetate	↓	↓	↓
2.22	s	Acetoacetate	↑	↓	↑
2.37	s	Pyruvate	↑	↑	↑
3.25/3.42	t/dd	Taurine/β-glucose*	↑	↑	↑
3.40/3.41	t	β-glucose/Taurine*	↑	↑	↑
3.9	dd	β-glucose	↑	↑	↑

3.2 Analysis of Urine

This chapter describes the ^{1}H-NMR spectroscopic analysis of urine obtained from 44 patients with confirmed pancreaticobiliary cancer and 45 patients with benign pancreaticobiliary disease. Among those with benign disease, three patients (BU3, BU4 and BU40) failed to provide a urine sample. A total of 44 and 42 confirmed cancer and benign patient samples respectively were subsequently profiled via ^{1}H-NMR using the 1D NOESY pulse sequence.

Multivariate analysis was performed as described in section 2.5 in an attempt to identify potential novel biomarkers for pancreaticobiliary malignancy.

3.2.1 Principle component analysis of the whole urine cohort

A PCA model for the whole dataset produced six PC's (Figure 3.31). An $R^2\boldsymbol{X}$(cum) and $Q^2\boldsymbol{Y}$(cum) of 0.808 and 0.405 respectively is indicative of a poor model.

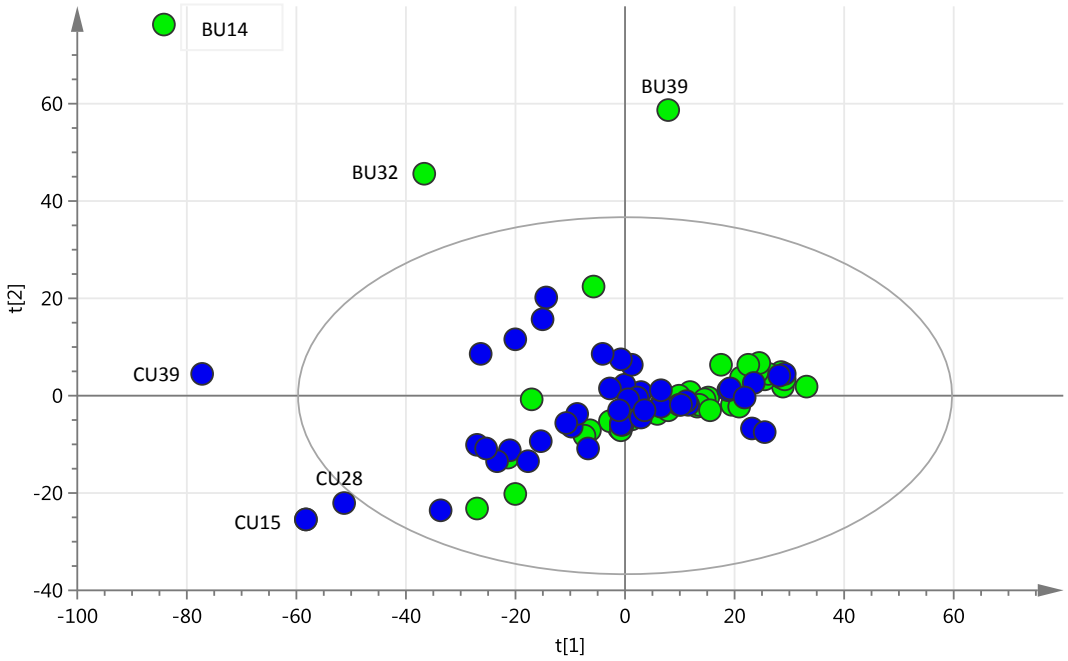

Figure 3.31 PCA scores plot of urine data for all 43 cancer (blue) and 42 benign (green) samples

A repeat PCA scores plot with outlier sample exclusion (BU14, BU32, BU39, CU18, CU28 and CU39) failed to show clear separation with a $R^2\boldsymbol{X}$(cum) and $Q^2\boldsymbol{Y}$(cum) of 0.744 and 0.389 respectively (Figure 3.32).

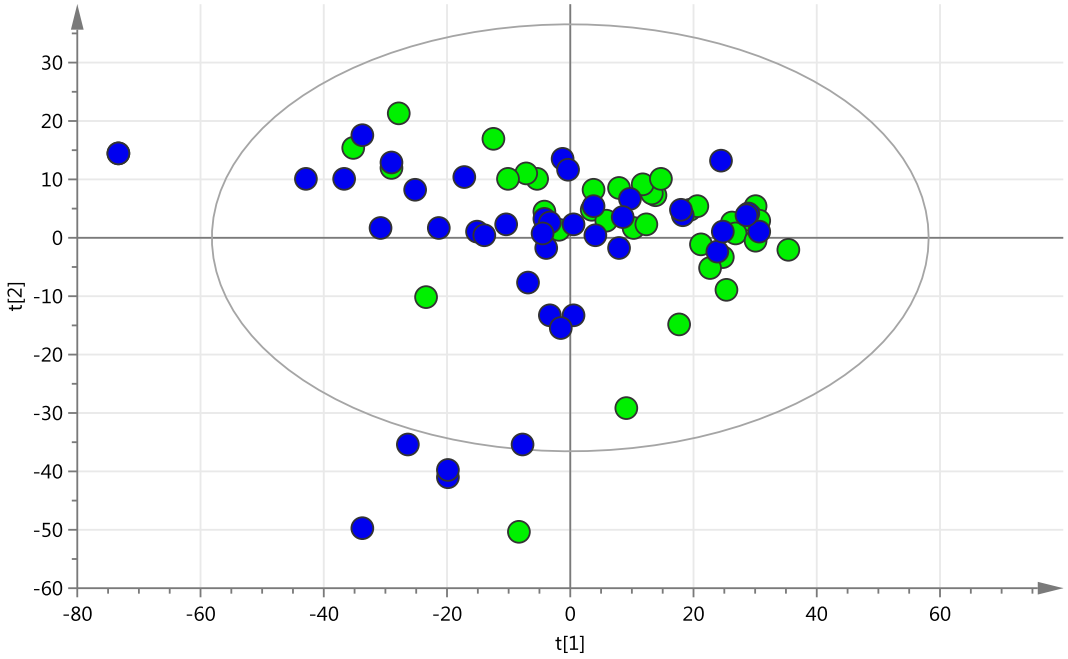

Figure 3.32 PCA scores plot of urine data for 40 cancer (blue) and 39 benign samples

3.2.2 Principle component analysis for potential confounding variables

In further PCA models separation between samples according to variables such as sex, ethnicity, pre-operative serum bilirubin, tumour type and tumour respectability is tested.

3.2.2.1 Ethnicity as a potential confounding factor

Unlike plasma chemometric analysis as described in section 3.1.2.1, PCA modelling failed to highlight any separation of data according to ethnicity with an $R^2\boldsymbol{X}$ (cum) and ($Q^2\boldsymbol{Y}$(cum) of 0.816 and 0.164 respectively (Figure 3.33). Samples BU14 and BU18 appeared as extreme outliers (Figure 3.33). Patient BU16 underwent laparoscopic cholecystectomy on a semi-urgent basis due to acute pancreatitis. Upon further review patient BU18 was found

to have an extensive prescription drug history. Further multivariate analysis of urine will include patients of all ethnicity.

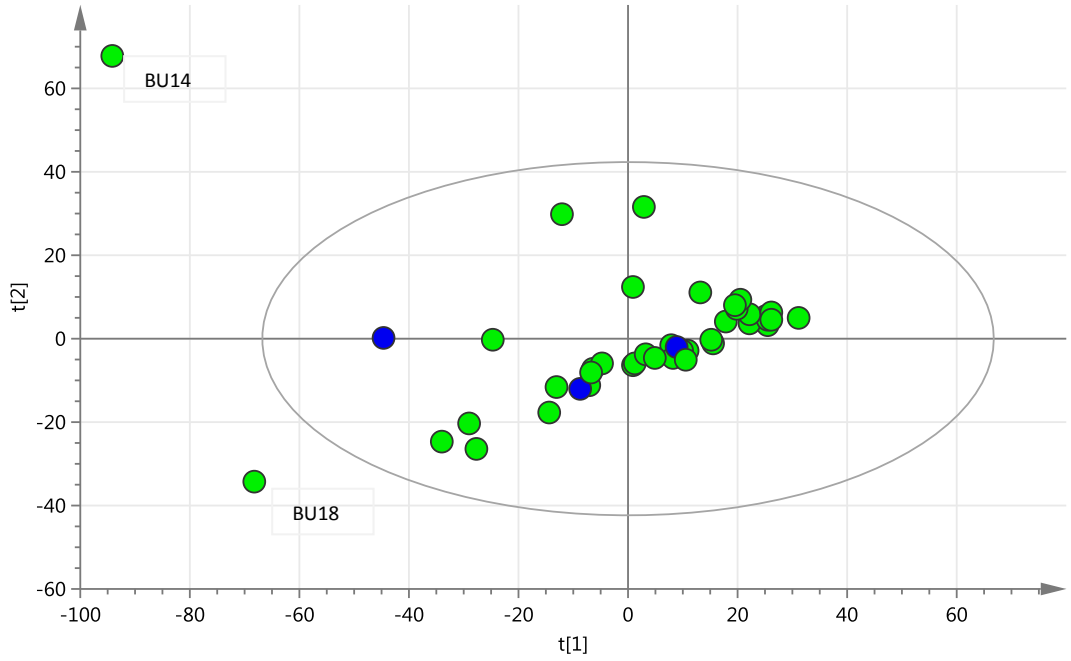

Figure 3.33 PCA scores plot of urine data for benign European (green) and non-European (blue) urine

3.2.2.2 Benign cohort disease status as a confounding factor

PCA modelling failed to identify any clear separation between benign or pre-malignant benign samples with an $R^2\boldsymbol{X}$ (cum) and $Q^2\boldsymbol{Y}$(cum) of 0.816 and 0.164 respectively (Figure 3.34). For subsequent urine meta-analysis both benign and pre-malignant urine samples will be considered as a single cohort. BU14 and BU16 are obvious outliers for the potential reasons described in section 3.2.2.1.

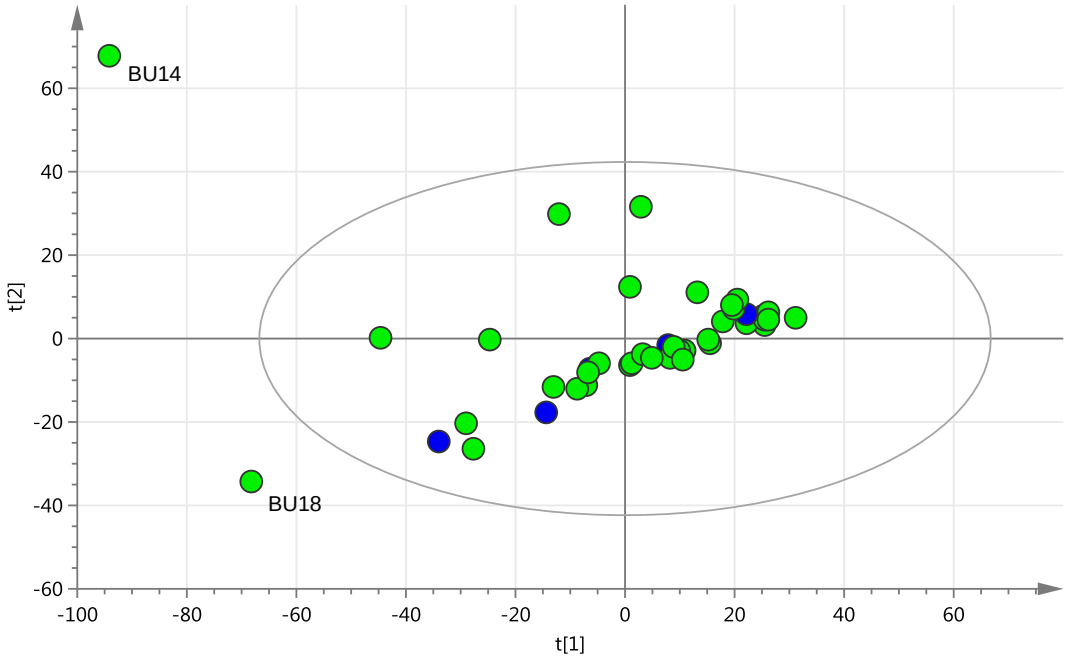

Figure 3.34 PCA scores plot of urine data for benign (green) and pre-malignant (blue) samples

3.2.2.3 Age and sex as confounding factors

The patients among the benign urine cohort are younger with a mean age of 52.1 versus 65.9 years. No statistically significant sex variation was apparent (Table 3-12).

3.2.2.4 Age and sex variation for 44 cancer and 34 benign patients

	Cancer (n=44)	Benign (n=42)	Statistical Significance (p)
Age (yrs) mean +/- SD	65.9 +/- 9.7	52.1 +/- 17.7	0.0001
Sex M/F	27/17	19/23	0.1942

PCA modelling failed to highlight any obvious clustering of data within the benign cohort according to sex with an $R^2\boldsymbol{X}$ (cum) and $Q^2\boldsymbol{Y}$(cum) of 0.816 and 0.164 respectfully (Figure 3.35).

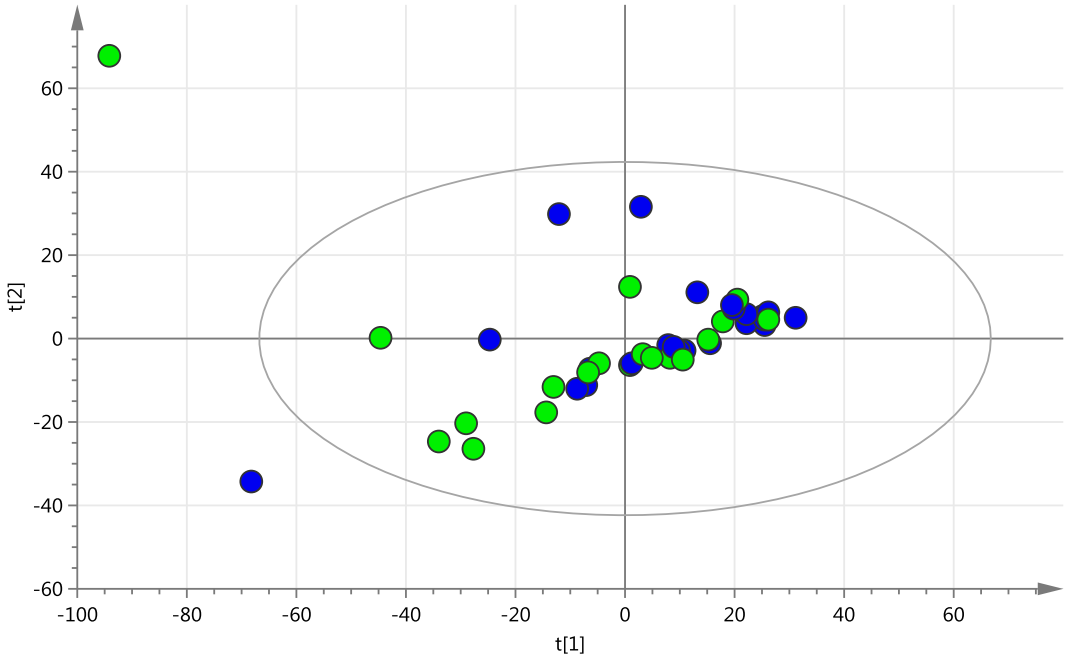

Figure 3.35 PCA scores plot of data for male (green) and female (blue) benign samples

Potential clustering according to sex was identified among the cancer cohort on PCA modelling. As displayed in Figure 3.36, male samples appear predominantly within cluster 1 and females within cluster 2. With an $R^2\boldsymbol{X}$ (cum) and $Q^2\boldsymbol{Y}$(cum) of 0.697 and 0.365 respectively, the model does not reach validity due to either noise or outlying data points. A repeat PCA model, following exclusion of outlying samples (CR15, CR26, CR28 and CR39) failed to display any clustering according to sex among the cancer cohort with an $R^2\boldsymbol{X}$(cum) and $Q^2\boldsymbol{Y}$(cum) of 0.715 and 0.326 respectively (Figure 3.37).

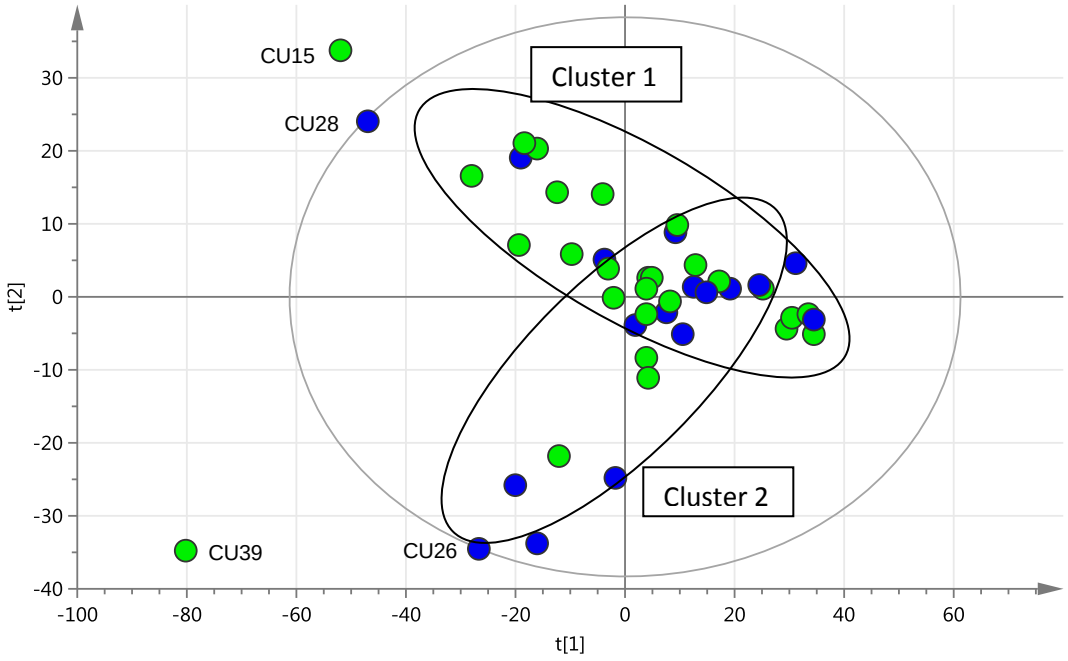

Figure 3.36 PCA scores plot of urine data for male (green) and female (blue) cancer samples

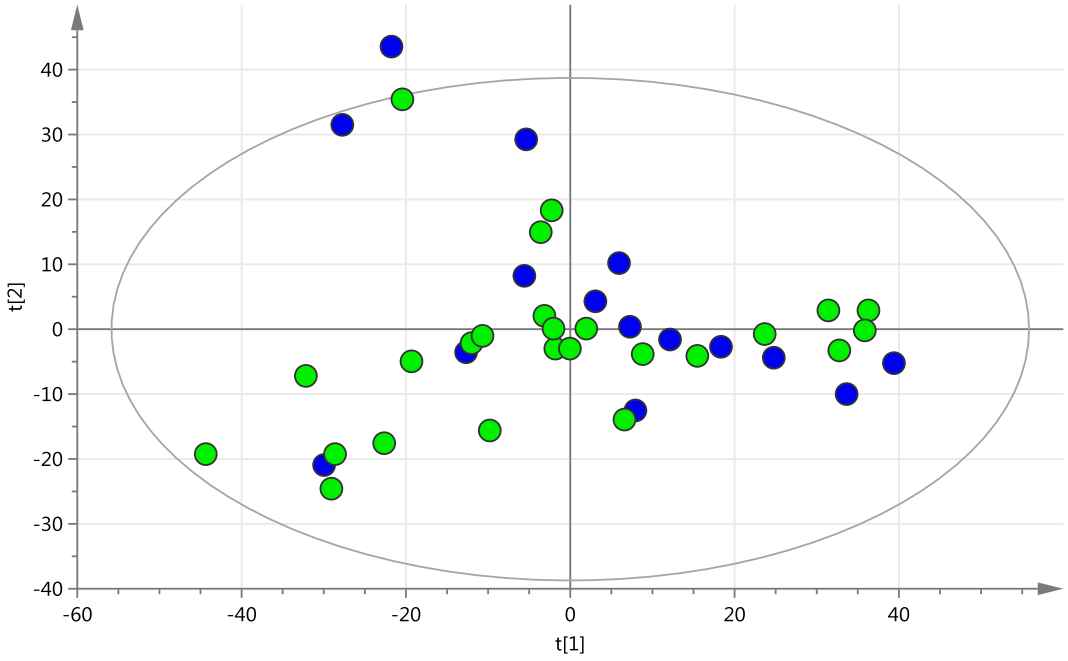

Figure 3.37 PCA scores plot of urine data for male (green) and female (blue) cancer samples with exclusions

3.2.2.5 Cancer subtype and resectability

PCA modelling failed to identify any clustering of data according to cancer type (PDA versus other malignancy) with an $R^2\boldsymbol{X}$ (cum) and $Q^2\boldsymbol{Y}$ (cum) of 0.808 and 0.405 respectively (Figure 3.38). A repeat model excluding outlying samples (BU14, BU32, BU39, CU15, CU28 and CU39) also failed to identify obvious clustering of data with an $R^2\boldsymbol{X}$ (cum) and $Q^2\boldsymbol{Y}$ (cum) of 0.756 and 0.362 respectively.

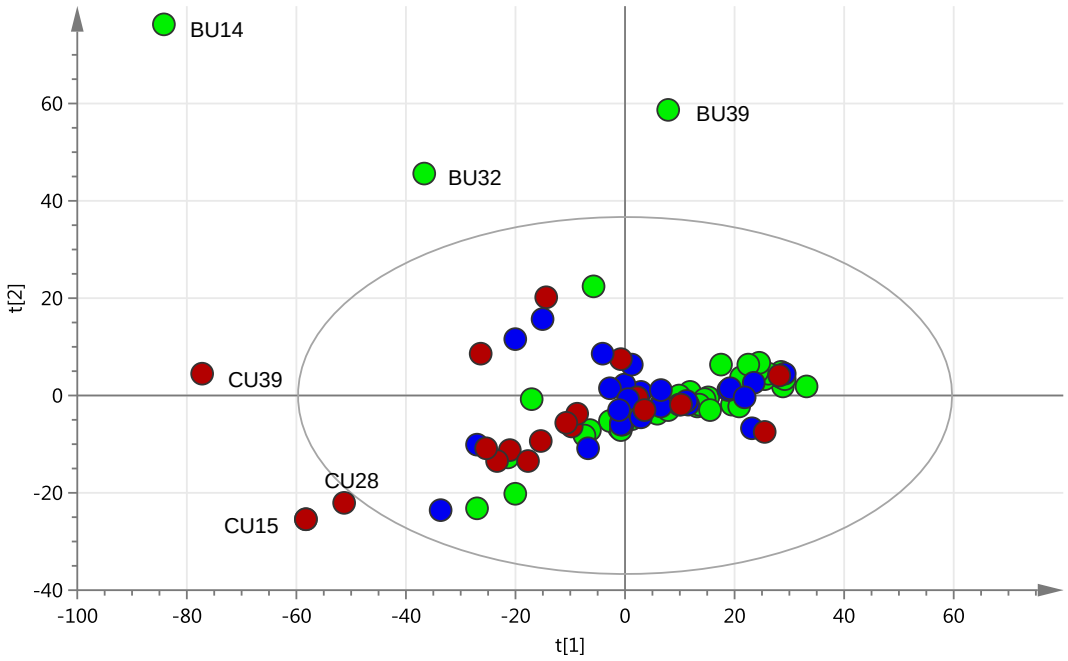

Figure 3.38 PCA scores plot of urine data for benign (green), PDA (blue) and other CU (red)

PCA modelling failed to identify any obvious clustering of data according to cancer resection status (resected or palliative operation) with an $R^2\boldsymbol{X}$ (cum) and $Q^2\boldsymbol{Y}$(cum) of 0.808 and 0.405 respectively (Figure 3.39). A repeat PCA model following extreme outlying sample exclusion also failed to identify any clustering of data with an $R^2\boldsymbol{X}$ (cum) and $Q^2\boldsymbol{Y}$(cum) of 0.756 and 0.362 respectively (Figure 3.40). Similarly a PCA scores plot for resectable versus irresectable cancer urine (without benign cohort inclusion) also failed to reveal any obvious clustering of data with an $R^2\boldsymbol{X}$ (cum) and $Q^2\boldsymbol{Y}$(cum) of 0.697 and 0.365 respectively (Figure 3.41).

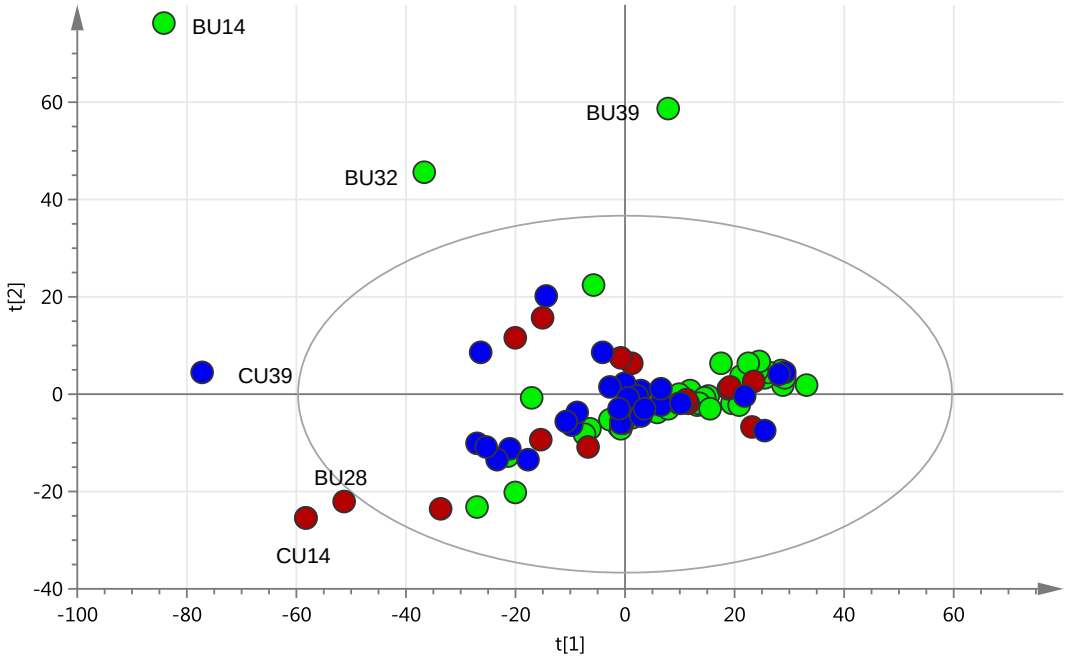

Figure 3.39 PCA scores plot of urine data for benign (green), resectable (blue) and palliative cancer samples. Outlying samples labelled

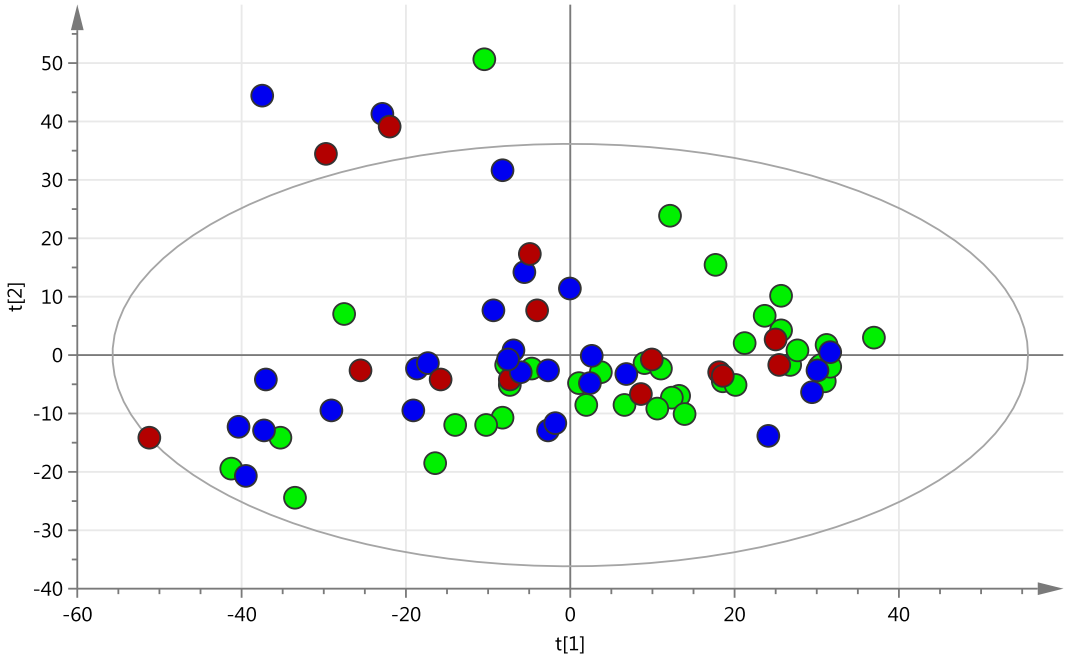

Figure 3.40 PCA scores plot of urine data for benign (green), resectable (blue) and palliative cancer samples with outlier exclusions

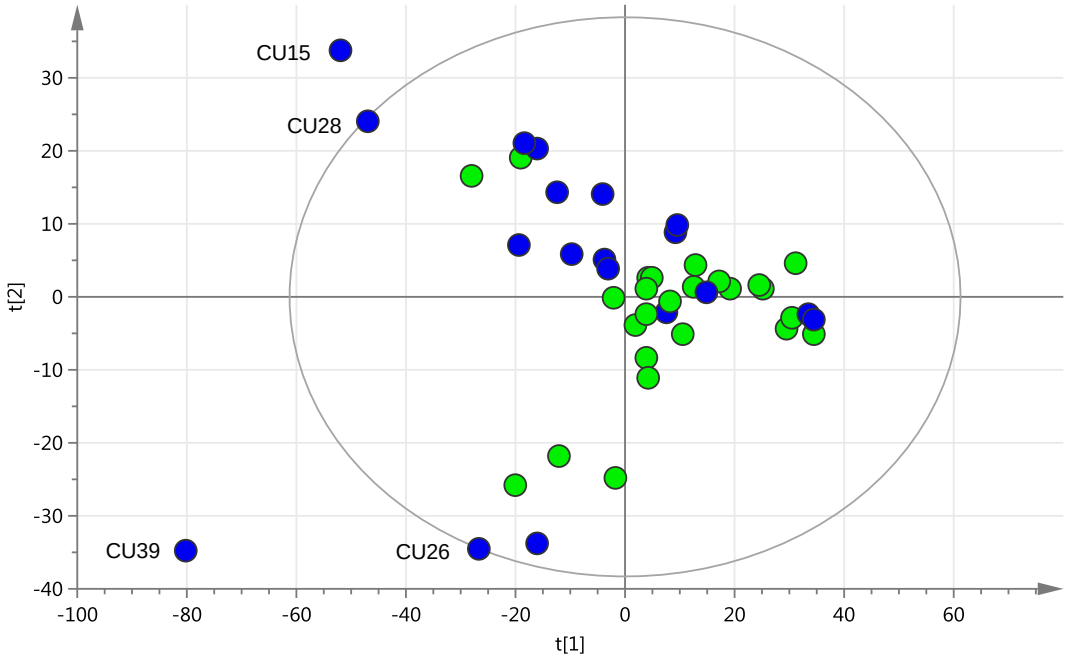

Figure 3.41 PCA scores plot of urine data for resectable (green) and palliative (blue) cancer samples

3.2.2.6 Pre-operative bilirubin variation

PCA modelling of benign samples, cancer samples with a pre-operative bilirubin of less than 40 or greater than 40 micromol/L failed to reveal any obvious clustering of data with an $R^2\boldsymbol{X}$ (cum) and $Q^2\boldsymbol{Y}$(cum) of 0.808 and 0.405 respectively (Figure 3.42). PCA modelling of cancer urine with a pre-operative bilirubin less than or equal to 40 versus greater than 40 micromol/L (excluding benign cohort) also failed to highlight any separation with an $R^2\boldsymbol{X}$ (cum) and $Q^2\boldsymbol{Y}$ (cum) of 0.697 and 0.365 respectively (Figure 3.43).

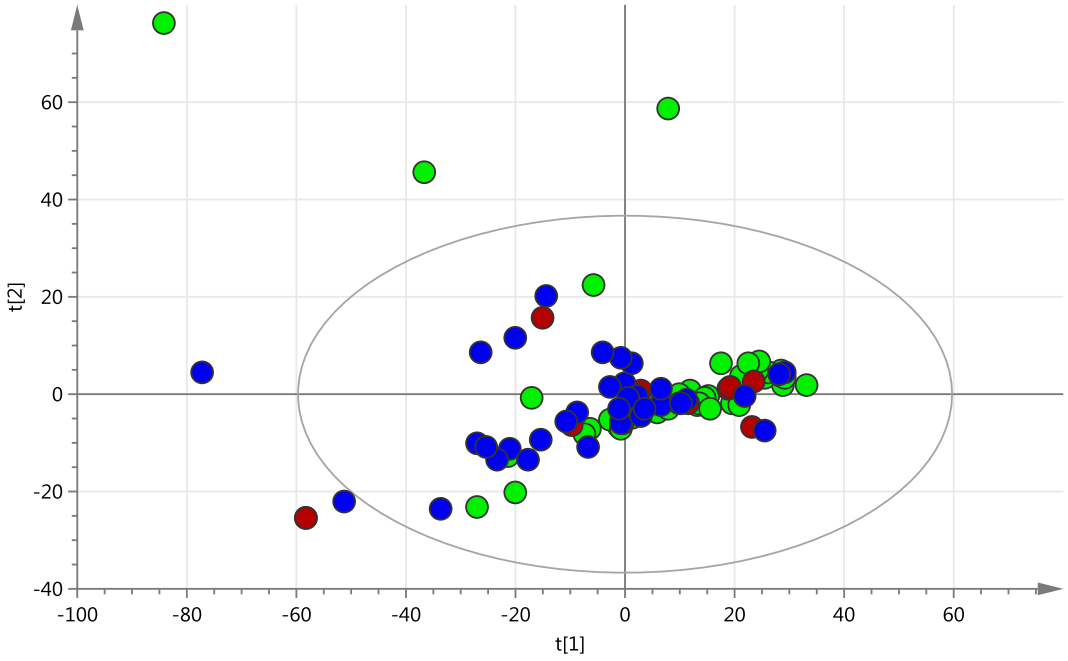

Figure 3.42 PCA scores plot of urine data for benign (green), cancer samples with pre-operative bilirubin < 40 (blue) and > 40 micromol/L (red)

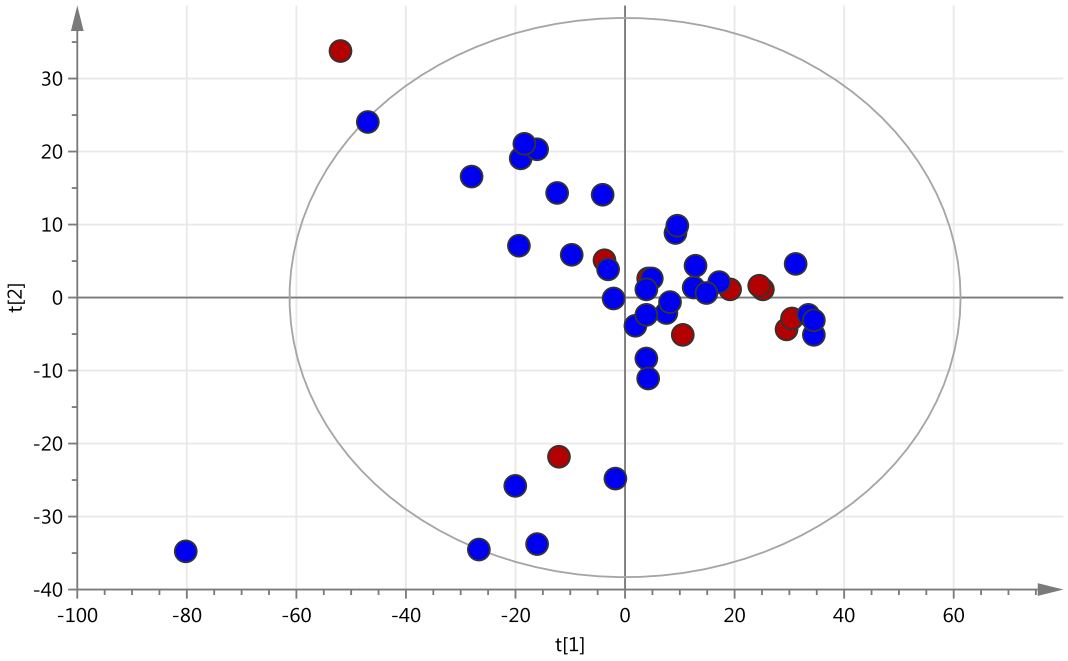

Figure 3.43 PCA scores plot of urine data for cancer samples with pre-operative bilirubin < 40 (blue) and > 40 micromol/L (red)

PCA modelling for benign urine samples and cancer patients with a pre-operative bilirubin less than or equal to 40 micromol/L failed to identify any obvious separation with a $R^2\boldsymbol{X}$ (cum) and $Q^2\boldsymbol{Y}$(cum) of 0.789 and 0.402 respectfully (Figure 3.44). A similar scores plot is displayed for urine samples with benign disease and cancer patients with a pre-operative bilirubin greater than 40 micromol/L with a resulting $R^2\boldsymbol{X}$ (cum) and $Q^2\boldsymbol{Y}$(cum) of 0.779 and 0.269 respectively (Figure 3.45). Given the large discrepancy in the number of benign patients (n=42) versus the number of cancer patients with a pre-operative bilirubin greater than 40 micromol/L (n=10) a separate model with random benign sample reduction is shown (Figure 3.46). $R^2\boldsymbol{X}$ (cum) and $Q^2\boldsymbol{Y}$ (cum) for this model is 0.726 and 0.313 respectfully.

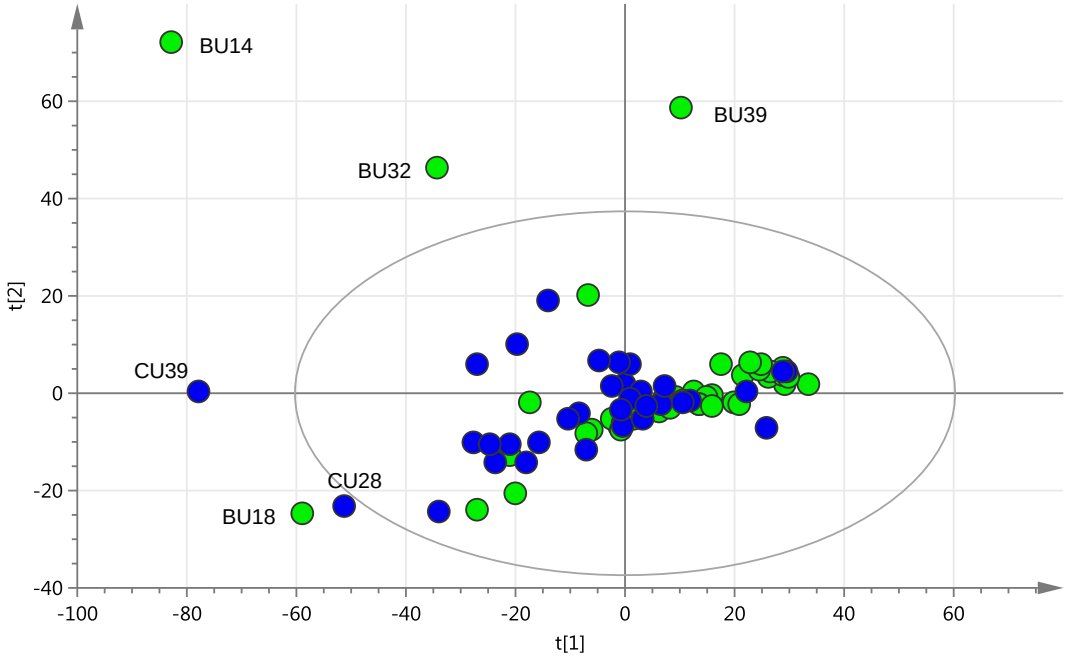

Figure 3.44 PCA scores plot of urine data for urine samples with benign disease (green) and those with a pre-operative bilirubin of less than or equal to 40 micromol/L (blue)

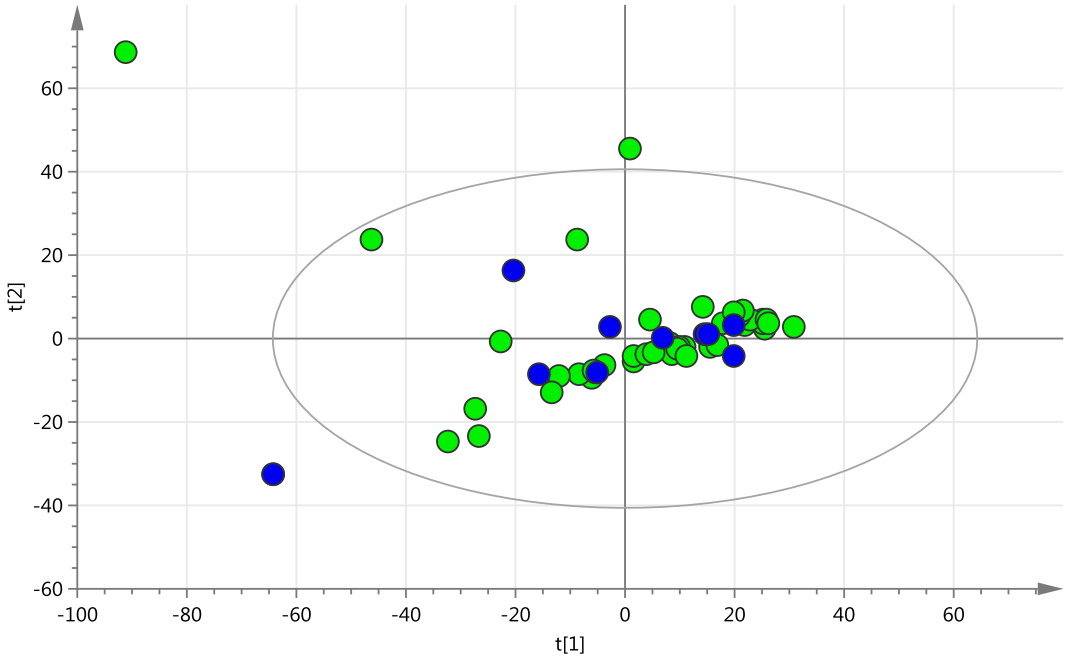

Figure 3.45 PCA scores plot of urine data for urine samples with benign disease (green) and those with a pre-operative bilirubin of greater than 40 micromol/L (blue)

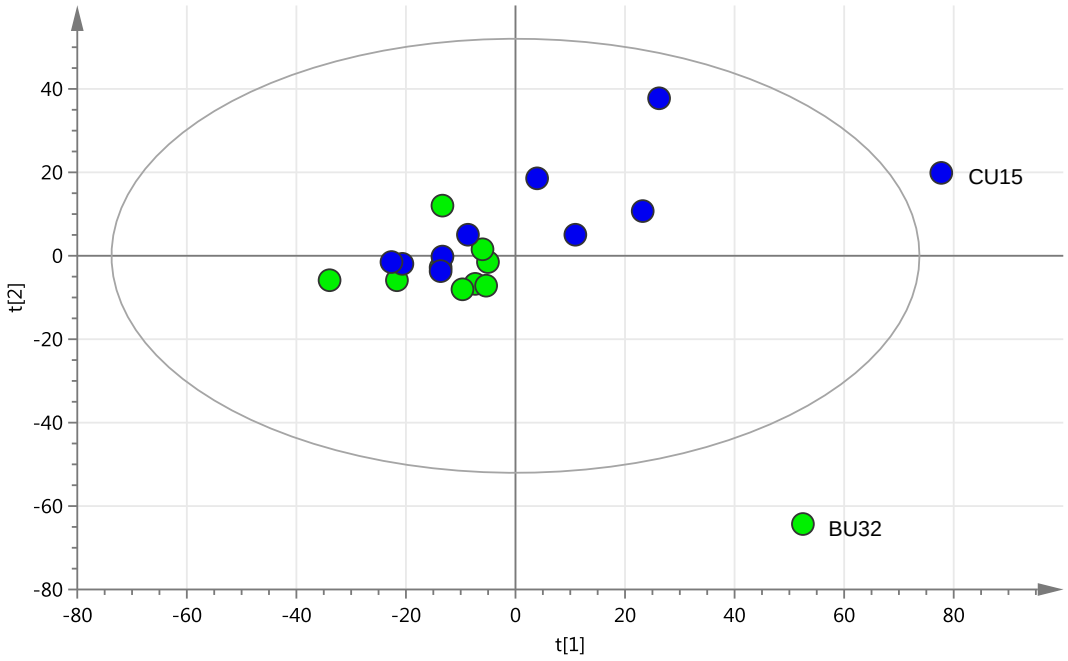

Figure 3.46 PCA scores plot of urine data for urine samples with benign disease (green) and those with a pre-operative bilirubin of greater than 40 micromol/L (blue) with random benign reductions

3.2.3 PLS & OPLS-DA analysis of urine samples

Unsupervised PCA analysis for urine, as described in section 3.2.2 failed to identify any potential clustering of either benign or malignant samples within principle components one and two. Supervised PLS and OPLS-DA analysis has thus been applied to selected cohorts to assess for any separation not revealed by PCA analysis alone.

3.2.3.1 PLS analysis for overall urine data for benign and malignant samples

It was not possible to generate a PLS-DA model using the cohort adopted for the PCA in either Figure 3.31 or Figure 3.32.

3.2.3.2 PLS-DA analysis according to cancer subtype

A PLS model comparing PDA urine versus “other” cancer urine within the cancer cohort is displayed below (Figure 3.47).

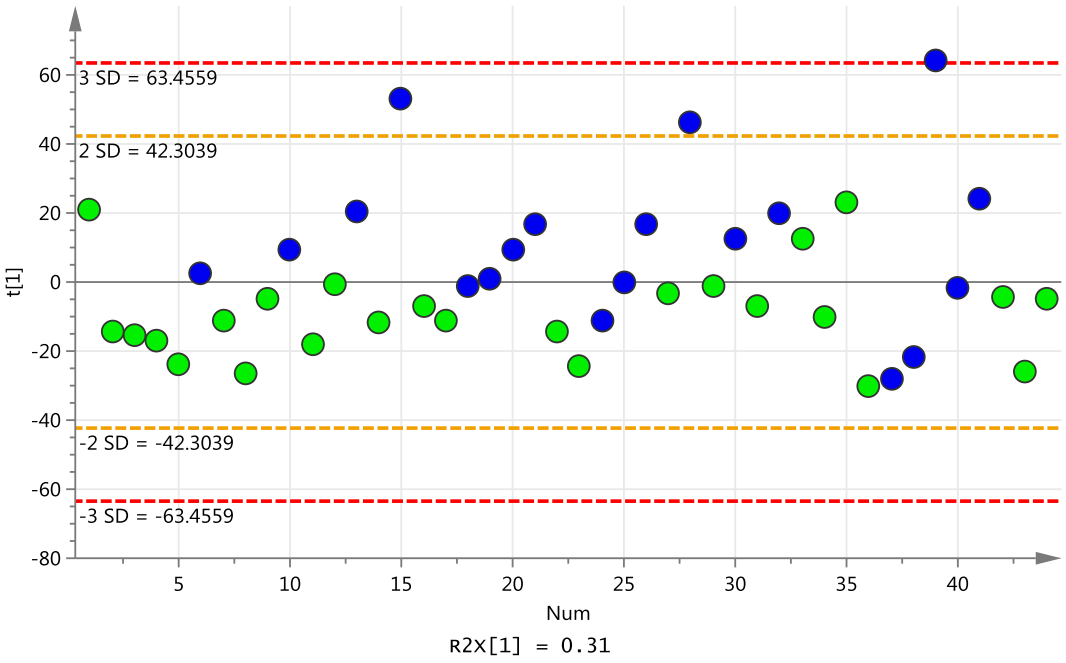

Figure 3.47 PLS-DA scores plot of urine data for 25 PDA (green) and 19 “other” cancer samples (blue). $R^2\boldsymbol{X}$ (cum) = 0.31, $R^2\boldsymbol{Y}$ (cum) = 0.257 and $Q^2\boldsymbol{Y}$ (cum) = 0.169

The model displayed in Figure 3.47 is poorly fitted and not valid as several of the permutated $\boldsymbol{R^2}Y$ values (to the left) lie at or above that of the original model with respect to both PDA (Figure 3.48) and “other” cancer (Figure 3.49) class permutation.

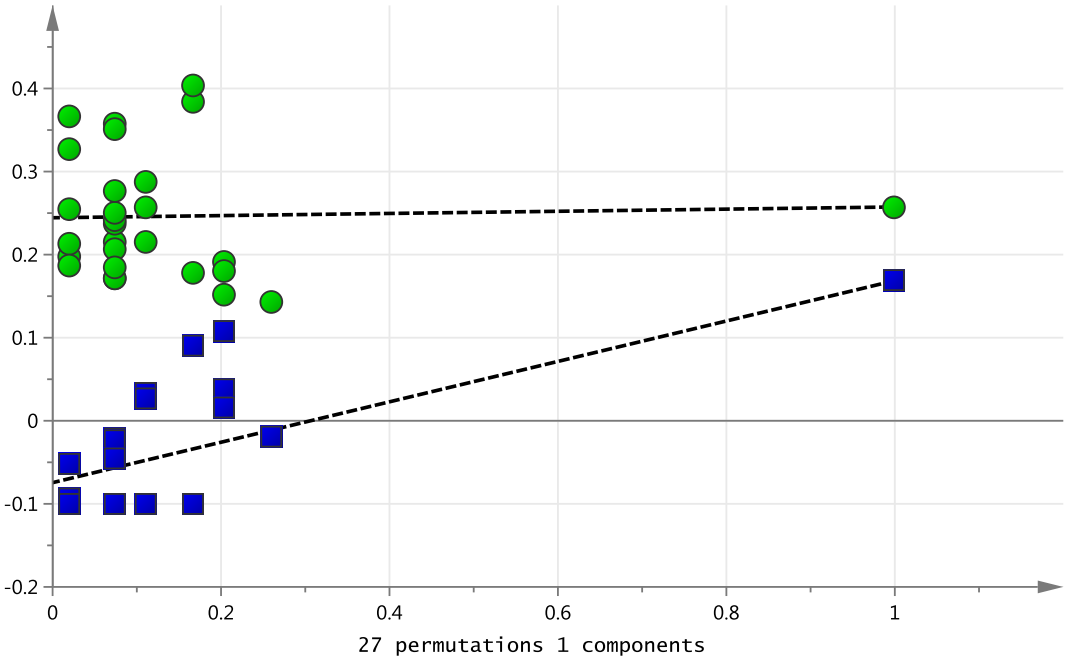

Figure 3.48 Permutation testing plot for the PLS-DA model shown in Figure 3.47. Twenty-seven permutations selected with reference to twenty-five PDA samples within the model. $R^2\boldsymbol{Y}$ = 0.244 and $Q^2\boldsymbol{Y}$ -0.0745

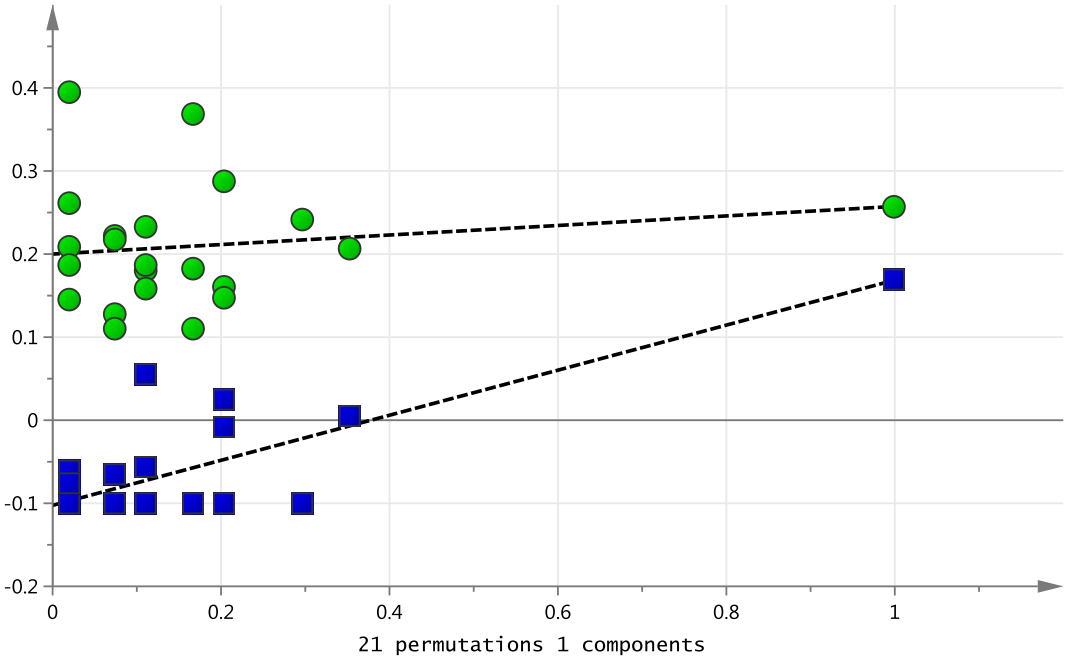

Figure 3.49 Permutation testing plot for the PLS-DA model shown in Figure 3.47. Twenty-one permutations selected with reference to nineteen "other cancer" samples within the model. $R^2\boldsymbol{Y}$ = 0.207 and $Q^2\boldsymbol{Y}$ -0.0999

The results of an attempt at cross validation through removing a third of samples in turn from both the PDA and "other" cancer cohort within the PLS-DA model is displayed within Table 3-10. The resultant sensitivity and specificity of the model is displayed (Table 3-11).

Table 3-10 Cross validation of the model displayed in Figure 3.47

Sample	Y predict	
	PDA	Other CU
CU1-PDA	-0.11	1.11
CU2-PDA	0.98	0.02
CU3-PDA	0.96	0.04
CU4-PDA	0.50	0.50
CU5-PDA	0.96	0.04
CU7-PDA	1.24	-0.24
CU8-PDA	0.87	0.13
CU9-PDA	0.71	0.29
CU11-PDA	0.51	0.49
CU6	0.76	0.24
CU10	0.90	0.10
CU13	0.24	0.76
CU15	-0.04	1.04
CU18	0.89	0.11
CU19	0.81	0.19
CU20	0.61	0.39

Sample	Y predict	
	PDA	Other CU
CU12-PDA	0.49	0.51
CU14-PDA	0.63	0.37
CU16-PDA	0.60	0.40
CU17-PDA	0.67	0.33
CU22-PDA	0.71	0.29
CU23-PDA	0.83	0.17
CU27-PDA	0.59	0.41
CU29-PDA	0.57	0.43
CU21	0.39	0.61
CU24	0.69	0.31
CU25	0.57	0.43
CU26	0.41	0.59
CU28	0.02	0.98
CU30	0.45	0.55

Sample	Y predict	
	PDA	Other CU
CU31-PDA	0.62	0.38
CU33-PDA	0.23	0.77
CU34-PDA	0.69	0.31
CU35-PDA	0.09	0.91
CU36-PDA	1.10	-0.10
CU42-PDA	0.56	0.44
CU43-PDA	0.96	0.04
CU44-PDA	0.58	0.42
CU32	0.27	0.73
CU37	1.07	-0.07
CU38	1.16	-0.16
CU39	-0.78	1.78
CU40	0.62	0.38
CU41	0.15	0.85

*A sample was regarded as belonging to a class with a Y predicted value of >0.50. Green indicates correct and red incorrect classification respectively. Class membership could not be allocated for sample CU4

Table 3-11 Model sensitivity, specificity, positive and negative predictive values

Test	PDA	
	Present (32)	**Absent* (38)**
Positive	True positive (20)	False positive (10)
Negative	False negative (4)	True negative (9)

	%	**95% C.I**
Sensitivity	83.3	62.6-95.2
Specificity	47.4	24.5-71.1
Positive predictive value	66.7	47.2-82.3
Negative predictive value	69.2	38.6-90.7

3.2.3.3 PLS-DA for cancer resectability

It was not possible to generate a PLS-DA model using the cohort adopted for the PCA in Figure 3.41.

3.2.3.4 PLS-DA for bilirubin variation

A PLS-DA model for benign urine and cancer urine with a pre-operative bilirubin less than or equal to 40micromol/L was not possible without excluding the outlying samples (BU14, BU18, BU32, BU39, CU28 and CU39) which lie beyond the 95% confidence interval in the PCA model (Figure 3.44). The resulting PLS-DA model is expressed in a single component (Figure 3.50).

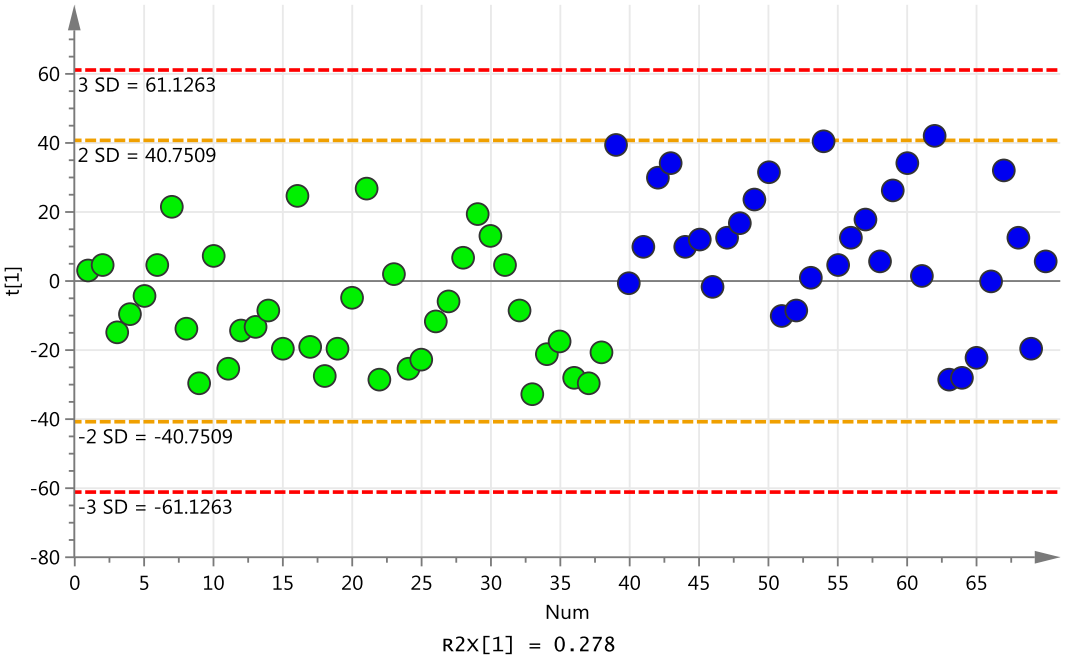

Figure 3.50 PLS-DA scores plot of urine data for 38 benign (green) and 32 cancer samples with a bilirubin less than or equal to 40micromol/L (blue) with outlier exclusions. $R^2\boldsymbol{X}$ (cum) = 0.278, $R^2\boldsymbol{Y}$ (cum) = 0.229 and $Q^2\boldsymbol{Y}$ (cum) = 0.131

Validity of the model was not confirmed through permutation testing with several permuted $\boldsymbol{R^2}Y$ values to the left lying above that of the original model with respect to both benign (Figure 3.51) and cancer (Figure 3.52) class permutation.

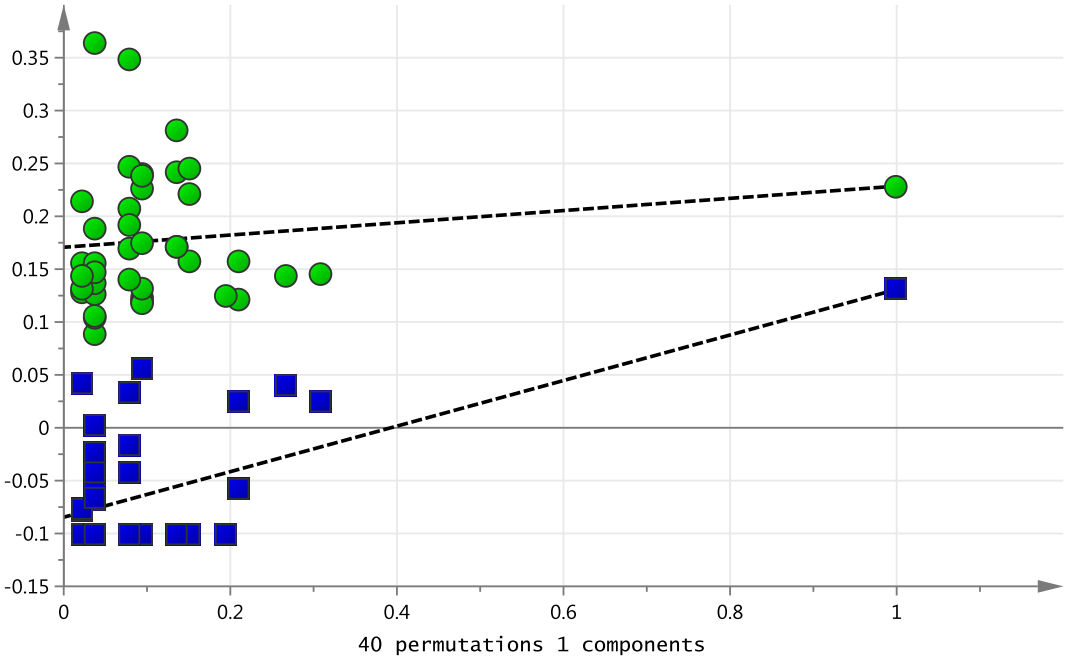

Figure 3.51 Permutation testing plot for the PLS-DA model shown in Figure 3.50. Forty permutations selected with reference to thirty-eight benign samples within the model. $R^2\boldsymbol{Y}$ = 0.171 and $Q^2\boldsymbol{Y}$ -0.0846

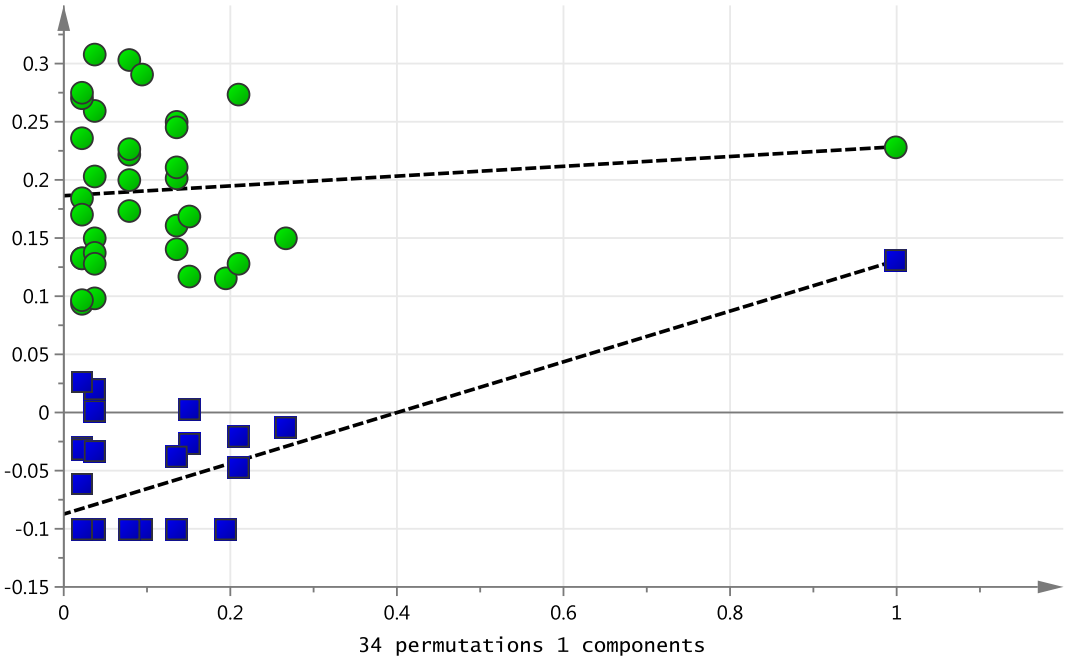

Figure 3.52 Permutation testing plot for the PLS-DA model shown in Figure 3.50. Thirty-four permutations selected with reference to thirty-two cancer samples within the model. $R^2\boldsymbol{Y}$ = 0.186 and $Q^2\boldsymbol{Y}$ -0.0874

Cross validation of the model in Figure 3.50 performed as described previously produced the results displayed in Table 3-12. The resultant model sensitivity and specificity is shown (Table 3-13).

Table 3-12 Cross validation of the model displayed in Figure 3.50

*A sample was regarded as belonging to a class with a Y predicted value of >0.50. Green indicates correct and red incorrect classification respectively

Sample	Y predicted	
	Benign	Cancer
BU1	0.47	0.53
BU2	0.45	0.55
BU5	0.67	0.33
BU6	0.61	0.39
BU7	0.56	0.44
BU8	0.48	0.52
BU9	0.24	0.76
BU10	0.65	0.35
BU11	0.83	0.17
BU12	0.36	0.64
BU13	0.78	0.22
BU15	0.71	0.29
BU16	0.64	0.36
CU1	0.12	0.88
CU2	0.59	0.41
CU7	0.55	0.45
CU10	0.26	0.74
CU13	0.16	0.84
CU14	0.55	0.45
CU16	0.49	0.51
CU17	0.60	0.40
CU18	0.08	0.92
CU19	0.41	0.59
CU20	0.35	0.65

Sample	Y predicted	
	Benign	Cancer
BU17	-1.99	2.99
BU19	0.75	0.25
BU20	0.89	0.11
BU21	0.79	0.21
BU22	0.80	0.20
BU23	0.73	0.27
BU24	0.80	0.20
BU25	0.56	0.44
BU26	0.81	0.19
BU27	0.64	0.36
BU28	0.77	0.23
BU29	0.74	0.26
BU30	0.72	0.28
CU21	0.35	0.65
CU22	0.90	0.10
CU24	0.86	0.14
CU25	0.80	0.20
CU26	0.02	0.98
CU27	0.60	0.40
CU29	0.58	0.42
CU30	0.38	0.62
CU31	0.57	0.43
CU32	0.36	0.64
CU33	0.00	1.00

Sample	Y predicted	
	Benign	Cancer
BU31	0.72	0.28
BU33	0.66	0.34
BU34	0.72	0.28
BU35	-0.21	1.21
BU36	0.58	0.42
BU37	0.76	0.24
BU38	1.11	-0.11
BU41	0.61	0.39
BU42	1.00	0.00
BU43	0.92	0.08
BU44	1.02	-0.02
BU45	0.21	0.79
CU34	0.58	0.42
CU35	0.47	0.53
CU36	0.96	0.04
CU37	0.96	0.04
CU38	1.60	-0.60
CU40	0.61	0.39
CU41	0.56	0.44
CU42	0.34	0.66
CU43	0.60	0.40
CU44	0.49	0.51

Table 3-13 Model sensitivity, specificity, positive and negative predictive values

Test	Cancer		Total
	Present (32)	**Absent (38)**	
Positive	True positive (15)	False positive (8)	23
Negative	False negative (17)	True negative (30)	36

	%	**95% C.I**
Sensitivity	46.9	29.1-65.2
Specificity	78.9	62.7-90.4
Positive predictive value	65.2	42.7-83.6
Negative predictive value	63.8	48.5-77.3

A PLS-DA model for benign urine and cancer urine with a pre-operative bilirubin of greater than 40micromol/L is displayed (Figure 3.53).

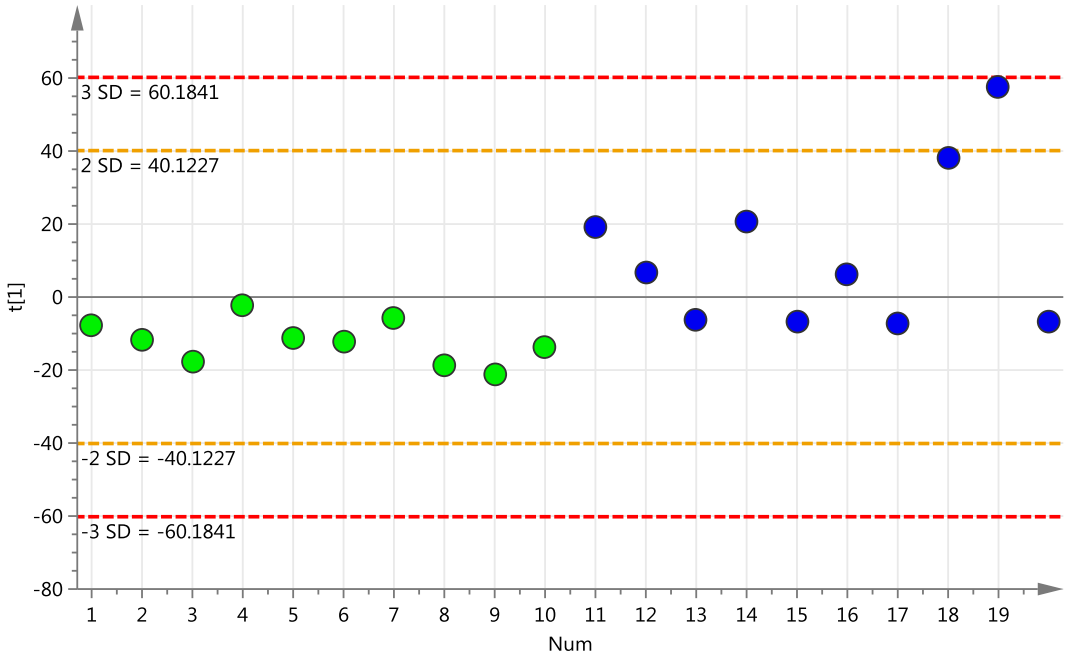

Figure 3.53 PLS-DA scores plot of urine data for 10 benign (green) and 10 cancer samples with a bilirubin greater than 40micromol/L (blue). $R^2\boldsymbol{X}$ (cum) = 0.301, $R^2\boldsymbol{Y}$ (cum) = 0.388 and $Q^2\boldsymbol{Y}$ (cum) = 0.227

Validity of the model displayed in Figure 3.53 was not confirmed through permutation testing with several permuted $\boldsymbol{R^2}Y$ and $\boldsymbol{Q^2}Y$ values lying above that of the original model for benign (Figure 3.54) and cancer (Figure 3.55) class permutation.

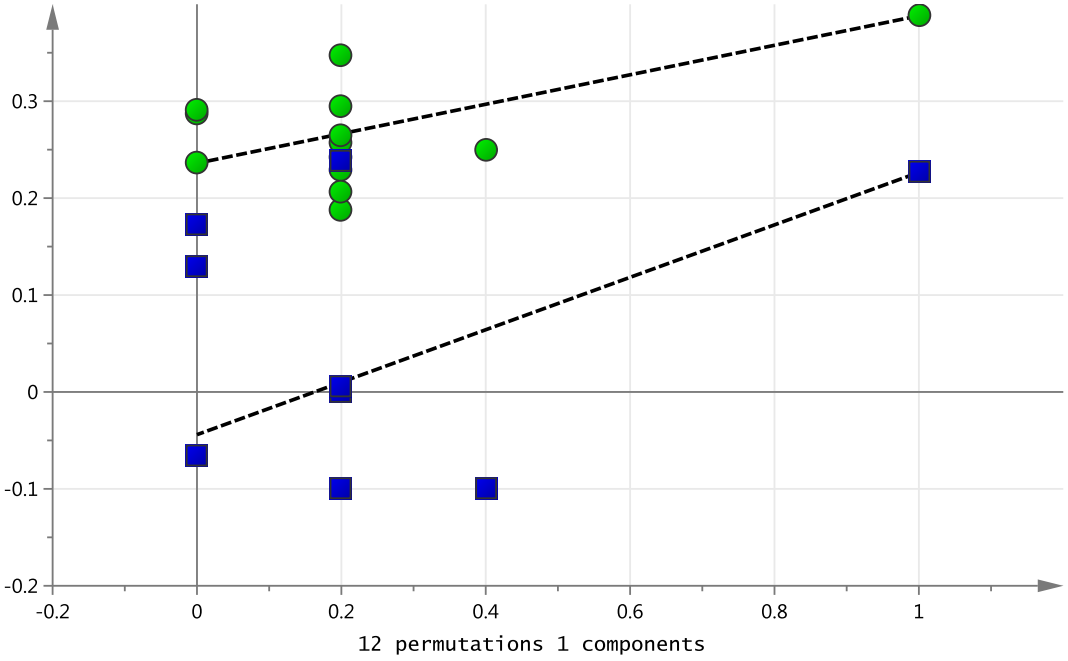

Figure 3.54 Permutation testing plot for the PLS-DA model shown in Figure 3.53. Twelve permutations selected with reference to ten benign samples within the model. $R^2\boldsymbol{Y}$ = 0.236 and $Q^2\boldsymbol{Y}$ -0.0441

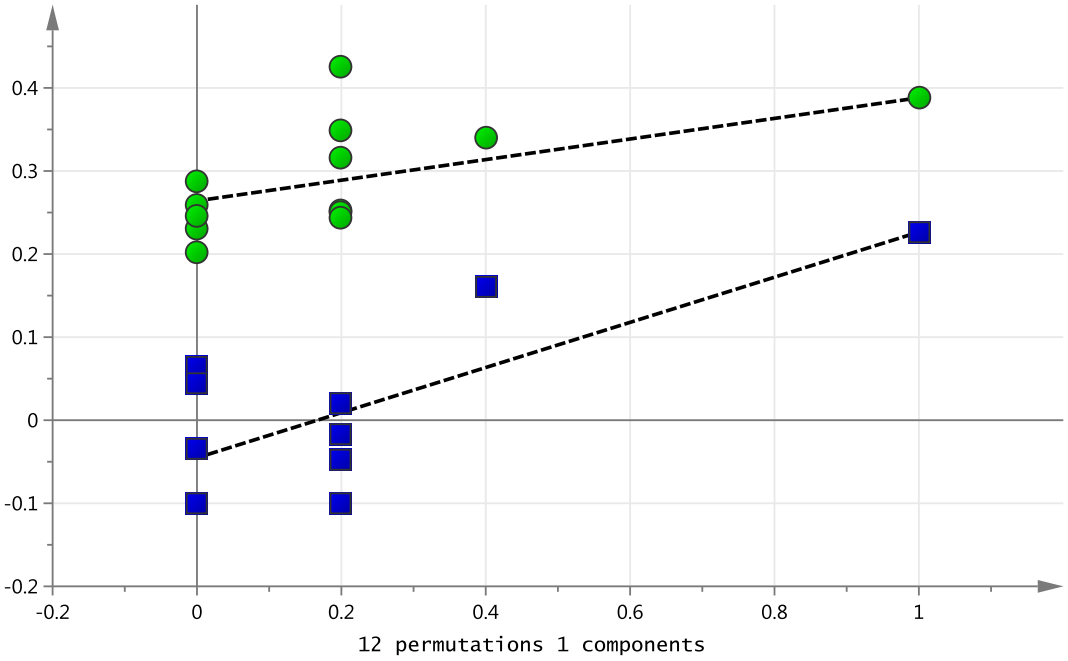

Figure 3.55 Permutation testing plot for the PLS-DA model shown in Figure 3.53. Twelve permutations selected with reference to ten cancer samples within the model. $R^2\boldsymbol{Y}$ = 0.264 and $Q^2\boldsymbol{Y}$ = -0.0453

Cross validation of the model was performed as before and the results displayed in Table 3-14 and Table 3-15.

Table 3-14 Cross validation of the model displayed in Figure 3.53

*A sample was regarded as belonging to a class with a Y predicted value of >0.50. Green indicates correct and red incorrect classification respectively

	Y predict*	
Sample	**Benign**	**Cancer**
BU6	0.57	0.43
BU10	0.67	0.33
BU13	0.74	0.26
BU15	0.47	0.53
CU3	0.23	0.77
CU4	0.13	0.87
CU5	0.60	0.40
CU6	0.24	0.76

	Y predict*	
Sample	**Benign**	**Cancer**
BU16	0.69	0.31
BU21	0.71	0.29
BU30	0.59	0.41
CU8	0.66	0.34
CU9	0.41	0.59
CU11	0.70	0.30

	Y predict*	
Sample	**Benign**	**Cancer**
BU32	-0.80	1.80
BU38	0.84	0.16
BU42	1.06	-0.06
CU12	-0.07	1.07
CU15	-2.55	3.55
CU23	0.60	0.40

Table 3-15 Model sensitivity, specificity, positive and negative predictive values

Test	PDA		Total
	Present (32)	Absent (38)	
Positive	True positive (6)	False positive (1)	7
Negative	False negative (4)	True negative (8)	12

	%	95% C.I
Sensitivity	60	26.4-87.6
Specificity	88.9	51.7-98.2
Positive predictive value	85.7	42.2-97.6
Negative predictive value	66.7	34.9-89.9

4 Discussion

For the purpose of this study eighty-five plasma samples were analysed by ^{1}H-NMR using the CPMG pulse sequence. Initial PCA analysis was suggestive of plasma metabolome separation between benign and malignant disease cohorts (Figure 3.1). Several outlying samples (CR1, CR12, CR13, CR20 and CR38) were identified. Patients CR1, CR13 and CR38 were all prescribed levothyroxine for hypothyroidism. This may be a confounding factor for metabonomic variation among these patients. No clear discrimination in the metabolome was apparent with regards to sex (Figure 3.5). In addition no clear metabonomic variation between benign and pre-malignant disease cohorts was identified (Figure 3.3). This is a potential direct result of insufficient patient recruitment within the pre-malignant arm of the study. The current study was not designed to evaluate the metabolome of patients with known pre-malignant disease. As such only five patients with pre-malignant disease were included within the study (diagnosed only following pancreatic resection for presumed malignancy). Among the five patients with pre-malignant disease, one patient (BR6) laid outside of the 95% confidence interval as an outlying sample. In order to evaluate the

metabolome of patients with benign versus malignant disease alone patients with pre-malignant disease were excluded from further analysis.

Pre-operative plasma bilirubin accounted for significant variability in the metabolome among cancer patients. A cut off bilirubin of 40 micromol/L appeared to be significant (Figure 3.8). It is important to emphasise that although the pre-operative plasma bilirubin was statistically dissimilar between benign and cancer cohorts respectively (p 0.0007), the mean bilirubin among the cancer cohort was only 29.1 versus 9.3 micromol/L among benign patients. From a clinical standpoint a bilirubin of 29 micromol/L is only mildly elevated and is at the lower end of that which would be detectable upon clinical examination. Many patients included in the resection arm of the study had already undergone ERCP and biliary stent insertion. This implies that pre-operative samples were collected from patients with a clearing plasma bilirubin. If plasma samples had been collected prior to ERCP and stenting, the effect of bilirubin on the metabolome may have been more pronounced. It is therefore imperative that future studies document and account for plasma bilirubin in chemometric analysis.

PLS-DA analysis was subsequently performed on a cohort of 34 benign and 42 cancer samples (Figure 3.11). Validity was assessed through both permutation testing and cross validation techniques (section 3.1.3). Model sensitivity, specificity, positive and negative predictive values of 64.9, 73.5, 72.7 and 65.8% respectively were generated through internal cross validation techniques (Table 3-3). Similar PLS-DA models were generated for benign and malignant plasma samples among a European population with a pre-operative plasma bilirubin of less than 40 micromol/L (Figure 3.22) and greater than 40 micromol/L (Figure 3.26).

OPLS loading plots facilitated metabolite identification (Table 3-9). Up-regulated metabolites in a state of malignancy included isobutyrate, 3-hydroxybutyrate, lactate, acetoacetate, pyruvate, taurine and β-glucose. Down-regulated metabolites included VLDL, valine and acetate. Upon

subgroup analysis valine appeared suppressed only among non-jaundiced cancer patients. 3-hydroxybutyrate appeared up-regulated only among cancer patients with a bilirubin > 40 micromol/L. Acetoacetate was up-regulated among cancer patients with a bilirubin > 40 but suppressed in those with a bilirubin < 40 micromol/L (Table 3-9).

To date one publication compares the ^{1}H-NMR pancreatic cancer plasma metabolome to that of healthy volunteers or those with chronic pancreatitis (130). The study is relatively small in size and included only nineteen pancreatic cancer patients, twenty patients with chronic pancreatitis and twenty healthy volunteers. VLDL, valine, 3-hydroxybutyrate and lactate were metabolites identified in common with this current study. In contrast to the current study Zhang *et al.* described up-regulation of VLDL along with suppression of LDL and HDL among pancreatic cancer subjects (when compared against healthy controls). VLDL was also interestingly found to be up-regulated among patients with chronic pancreatitis when compared against healthy controls. This discrepancy with the current study may be accounted for by the recruitment of patients with benign pancreaticobiliary disease (including chronic pancreatitis) to comprise the control arm of the current study (rather than healthy volunteers). In common with the current study down-regulation of valine among cancer patients was described. In addition I however emphasise that down regulation was only apparent among patients with a plasma bilirubin < 40 micromol/L in the current study.

The described up-regulation of taurine along with glucose among pancreaticobiliary cancer patients was similarly described by Tesiram *et al.* who evaluated the serum metabolome of patients with pancreatic cancer in comparison to that of healthy volunteers (133). Up-regulation of glucose was similarly described by Bathe *et al.* who studied the serum metabolome pancreatic cancer along with that of patients with benign biliary disease (131).

In contrast to the current study OuYang *et al.* along with Zhang *et al.* described down regulation of 3-Hydroxybutyrate among pancreatic cancer subjects (130, 132). In the current study 3-hydroxybutyrate appeared up-

regulated among patients with a pre-operative bilirubin > 40 micromol/L although unchanged among those with a pre-operative bilirubin of < 40 micromol/L.

^{1}H-NMR analysis of urine using the 1D-NOESY pulse sequence failed to provide statistically significant evidence of urinary metabonomic variation between benign and malignant pancreaticobiliary disease cohorts (section 3.2). These findings most likely reflect the known environmental influence over the urinary metabolome. Lenz *et al.* in 2003 performed H-NMR metabonomic analysis on both plasma and urine obtained from a group of 12 healthy male subjects on two separate intervals 14-days apart. Despite standardization of diet there was considerable inter-subject (but not intra-subject) variability of the urinary metabolome (140).

In conclusion I report the feasibility of differentiating the plasma metabonomic profile of patients with pancreaticobiliary malignancy from those with benign disease. Subgroup analysis in the current study failed to identify significant separation according to cancer cell type, resectability or pre-malignant disease status. I believe this is likely a reflection of patient recruitment rather than a true finding. Pre-operative plasma bilirubin however appeared to have a significant effect on the metabolome.

To facilitate biomarker discovery for pancreatic cancer future NMR studies require to carefully consider and describe techniques to counteract the confounding effect of bilirubin on the plasma metabolome. One potential solution is to describe varying metabonomic profiles for a neoplastic state based upon the presenting bilirubin range. This clearly adds complexity to future clinical biomarker utilization. A second option is to identify and zero fill a section of the metabolome directly confounded by plasma bilirubin prior to chemometric analysis. A third option is to deploy the novel biomarker following pre-operative biliary drainage procedures or among a non-jaundiced patient cohort. This clearly delays and limits clinical utility of the biomarker and would preclude its use in surgical units who perform pancreatic resections in patients with an obstructed biliary system.

Larger studies are clearly required to further define the plasma metabolome of pancreaticobiliary malignancy. Although it may be possible to define histological cell type and predict resectability, chemosenstivity or indeed prognosis I believe the first step is to aid in the differentiation of a malignant pancreaticobiliary process from that of benign disease. A combination strategy along with other modalities such as clinical imaging and currently available biomarkers such as CA 19-9 would most likely be used. To verify the described plasma findings and to facilitate biomarker development it would be useful in future studies to evaluate the metabolome at diagnosis, following resection and at the time of potential disease recurrence. Any future metabonomic profile of disease or novel biomarker model would also require validation with an external dataset.

5 References

1. Pancreatic cancer survival is lowest of all common cancers: Office for National Statistics. Available from: http://www.ons.gov.uk/ons/rel/vsob1/cancer-statistics-registrations--england--series-mb1-/no--42--2011/sty-pancreatic-cancer.html.
2. Cashman R, Gerhardt P, Goldberg I, Handy V, Levin M. The probability of developing cancer. Journal of the National Cancer Institute. 1956;17:155-73.
3. Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, et al. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11: International Agency for Research on Cancer; 2013. Available from: http://globocan.iarc.fr.
4. Singh S, Longmire W, Reber H. Surgical Palliation for pancreatic cancer. The UCLA experience. Annals of Surgery. 1990;212:132-9.
5. Key Sites Study: Pancreas Report 2000: Northern and Yorkshire Cancer Registry and Information Service (NYCRIS). Available from: http://www.nycris.nhs.uk/research/archive/key_sites/.
6. Richter A, Niedergethmann M, Sturm J. Long-term results of partial pancreaticoduodenectomy for ductal adenocarcinoma of the pancreatic head: 25-year experience. World Journal of Surgery. 2003;27(3):324-9.
7. Amikura K, Kobari M, Matsuno S. The time of occurrence of liver metastasis in carcinoma of the pancreas. International Journal of Pancreatology. 1995;17(2):139-46.
8. Kayahara M, Nagakawa T, Ueno K. An evaluation of radical resection for pancreatic cancer based on the mode of recurrence as determined by autopsy and diagnostic imaging. Cancer. 1993;72(7):2118-23.

9. Karlson B, Ekbom A, Lindgren P, Kallskog V, Rastad J. Abdominal US for diagnosis of pancreatic tumor: prospective cohort analysis. Radiology. 1999;213:107-11.
10. Callery M, Chang K, Fishman E, Talaminti M, William Traverso L, Linehan D. Pre-treatment assessment of resectable and borderline resectable pancreatic cancer: expert consensus statement Annals of Surgical Oncology. 2009;16:1727-33.
11. Farma J, Santillan A, Melis M, Walters J, Belinc D, Chen D, et al. PET/CT fusion scan enhances CT staging in patients with pancreatic neoplasms. Annals of Surgical Oncology. Annals of Surgical Oncology. 2008;15:2465-71.
12. DiMango E, Buxton J, Regan P, Hattery R, Wilson D, Suarez J, et al. Ultrasonic endoscope. Lancet. 1980:629-31.
13. Mahmoudi N, Enns R, Amar J, Al Ali J, Lam E, Telford J. Biliary Brush Cytology: factors associated with positive yields on biliary brush cytology. World Journal of Gastroenterology. 2008;14:569-73.
14. Savides T, Donohue M, Hunt G, Al-Haddad M, Aslanian H, Ben-Menachem T, et al. EUS-guided FNA diagnostic yield of malignancy in solid pancreratic masses: a benchmark for quality performance measurement. Gastrointestinal Endoscopy. 2007;66:277-82.
15. Seidel G, Zahurak M, Lacobuzio-Donahue C, Sohn T, Adsay N, Yeo C, et al. Almost all infiltrating colloid carcimomas of the pancreas and peri-ampullary region arise from in situ papillary neoplasms: a study of 39 cases. American Journal of Surgical Pathology. 2002;26:56-63.
16. Adsay N, Pierson C, Sarkar F, Abrams J, Weaver D, Conlon K, et al. Colloid (mucinous non-cystic) carcinoma of the pancreas. . American Journal of Surgical Pathology. 2001;25:26-42.
17. Kardon D, Thompsom L, Przgodzki R, Heffess C. Adenosquamous Carcinoma of the Pancreas: A Clinicopatholgic Series of 25 cases Modern Pathology. 2001;14(5):443-51.
18. Paner G, Thompsom K, Reyes C. Hepatid carcinoma of the pancreas. Cancer. 2000;88(7):1582-9.
19. Goggins M, Offerhaus G, Hilgers W, Griffin C, Shekher M, Tang D, et al. Pancreatic adenocarcinomas with DNA replecation errors (RER+) are associated with wild-type K-ras and characteristic histopathology, poor differentiation, a syncytial growth pattern, and pushing borders suggest RER+. American Journal of Pathology. 1998;69:38-43.
20. Wilentz R, Goggins M, Redston M, Marcus V, Adsay N, Sohn T, et al. Genetic, immunohistochemical, and clinical features of medullary carcinoma of the pancreas: a newly described and characterized entity. American Journal of Pathology. 2000;156:1641-51.
21. Hruban R, Fukushima N. Pancreatic adenocarcinoma: update on the surgical pathology of carcinomas of ductal origin and PanINs. Modern Pathology. 2007;20(S61-70).
22. Erkan M, Hausmann S, Michalski C, Fingerle A, Dobritz M, Kleeff J, et al. The role of stroma in pancreatic cancer: diagnostic and therapeutic implications. Nature Reviews Gastroenterology & Hepatology. 2012;9:454-67.

23. Olive K, Jacobetz M, Davidson C, Gopinathan A, McIntyre D, Honess D, et al. Inhibition of Hedgehog signaling enhances delivery of chemotherapy in a mouse model of pancreatic cancer. Science. 2009;324:1457-61.
24. Neoptolemos J, Stocken D, Friess H, Bassi C, Dunn J, Hickey H, et al. A randomized trial of chemoradiotherapy and chemotherapy after resection of pancreatic cancer. New England Journal of Medicine. 2004;350:1200-10.
25. Gillen S, Schuster T, Meyer zum Büschenfelde C, Friess H, Kleeff J. Preoperative/Neoadjuvant Therapy in Pancreatic Cancer: A Systematic Review and Meta-analysis of Response and Resection Percentages. PLOS Medicine. 2010.
26. Lim KH, Chung E, Khan A, Cao D, Linehan D, Ben-Josef E, et al. Neoadjuvant Therapy of Pancreatic Cancer: The Emerging Paradigm? The Oncologist 2012;17:192-200.
27. Conroy T, Desseigne F, Ychou M, Bouche O, Guimbaud R, Becouaram Y, et al. FOLFIRINOX versus gemcitabine for metastatic pancreatic cancer. New England Journal of Medicine. 2011;364:1817-25.
28. Yachida S, Jones S, Bozic I, Antal T, Leary R, Fu B, et al. Distant metastasis occurs late during the genetic evolution of pancreatic cancer. Nature. 2010;467:1114-7.
29. Biankin A, Waddell N, Kassahn K, Gingras M, Muthuswamy L, Johns A, et al. Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. Nature. 2012;491:399-405.
30. Ansari D, Chen B, Dong L, Zhou MT, Andersson R. Pancreatic cancer: translational research aspects and clinical implications. World Journal of Gastroenterology. 2012;18:1417-24.
31. Shi G, Direnzo D, Qu C, Barney D, Miley D, Konieczny S. Maintenance of acinar cell organization is critical to preventing Kras-induced acinar-ductal metaplasia. Oncogene. 2013;32(15):1950-8.
32. Baumgart M, Werther M, Bockholt A, Scheurer M, Ruschoff J, Dietmaier W. Genomic instability at both the base pair level and the chromosomal level is detectable in earliest PanIN lesions in tissues of chronic pancreatitis. Pancreas. 2010;39:1093-103.
33. Force USPST. Screening for family and intimate partner violence: recommendation statement. Annals of Internal Medicine. 2004;140:382-6.
34. Canto MI, Harinck F, Hruban RH, Offerhaus GJ, Poley JW, Kamel I, et al. International Cancer of the Pancreas Screening (CAPS) Consortium summit on the management of patients with increased risk for familial pancreatic cancer. Gut. 2013;62:339-47.
35. Lowenfels A, Maisonneuve P. Epidemiology and prevention of pancreatic cancer. Japanese Journal of Clinical Oncology. 2004;34:238-44.
36. Boffetta P, Hecht S, Gray N, Gupta P, Straif K. Smokeless tobacco and cancer. The Lancet Oncology. 2008;9:667-75.
37. Li D, Morris J, Liu J, Hassan M, Day R, Bondy M, et al. Body mass index and risk, age of onset, and survival in patients with pancreatic cancer. JAMA. 2009;301:2553-62.
38. Ansary-Moghaddam A, Huxley R, Barzi F, Lawes C, Ohkubo T, Fang X, et al. The effect of modifiable risk factors on pancreatic cancer mortality in

populations of the Asia-Pacific region. . Cancer Epidemiology, Biomarkers & Prevention. 2006;15:2435-40.
39. Chari S, Leibson C, Rabe K, Ransom J, de Andrade M, Petersen G. Probability of pancreatic cancer following diabetes: a population-based study. Gastroenterology. 2005;129:504-11.
40. Thiebaut A, Jiao L, Silverman D, Cross A, Thompson F, Subar A, et al. Dietary fatty acids and pancreatic cancer in the NIH-AARP diet and health study. Journal of the National Cancer Institute. 2009;101:1001-11.
41. Marsh G, Lucas L, Youk A, Schall L. Mortality patterns among workers exposed to acrylamide: 1994 follow up. Occupational and Enviromental Medicine. 1999;56:181-90.
42. Ojajarvi I, Partanen T, Ahlbom A, Boffetta P, Hakulinen T, Jourenkova N, et al. Occupational exposures and pancreatic cancer: a meta-analysis. Occupational and Enviromental Medicine. 2000;57:316-24.
43. Blair A, Stewart P, Tolbert P, Grauman D, Moran F, Vaught J, et al. Cancer and other causes of death among a cohort of dry cleaners. British Journal of Industrial Medicine. 1990;47:162-8.
44. Bracci P, Wang F, Hassan M, Gupta S, Li D, Holly E. Pancreatitis and pancreatic cancer in two large pooled case-control studies. Cancer Causes and Control. 2009;2(9):1723-31.
45. Shi C, Hruban R, Klein A. Familial pancreatic cancer. Archives of Pathology and Laboratory Medicine. 2009;133(3):365-74.
46. Permuth-Wey J, Egan K. Family history is a significant risk factor for pancreatic cancer: results from a systematic review and meta-analysis. Familial Cancer 2009;8(2):109-17.
47. Gruber SB, Entius MM, Petersen GM, Laken SJ, Longo PA, Boyer R, et al. Pathogenesis of adenocarcinoma in Peutz-Jeghers syndrome. Cancer research. 1998;58:5267-70.
48. van Lier M, Wagner A, Mathus-Vliegen E, Kuipers E, Steyerberg E, van Leerdam M. High cancer risk in Peutz-Jeghers syndrome: a systematic review and surveillance recommendations. American Journal of Gastroenterology. 2010;105:1258-64.
49. Lynch H, Fusaro R, Sandberg A, Bixenman H, Johnsen L, Lynch J, et al. Chromosome instability and the FAMMM syndrome. Cancer Genetics and Cytogenetics. 1993;71:27-39.
50. Haluska F, Hodi F. Molecular genetics of familial cutaneous melanoma. Journal of Clinical Oncology. 1998;16:670-82.
51. Goldstein A, Chan M, Harland M, Hayward N, Demenais Demenais F, Bishop D, et al. Features associated with germline CDKN2A mutations: a GenoMEL study of melanoma-prone families from three continents. Journal of Medical Genetics. 2007;44:99-106.
52. Lynch H, Fusaro R, Lynch J, Brand R. Pancreatic cancer and the FAMMM syndrome. Familial Cancer. 2008;7:103-12.
53. Brose M, Rebbeck T, Calzone K, Stopfer J, Nathanson K, Weber B. Cancer risk estimates for BRCA1 mutation carriers identified in a risk evaluation program. Journal of the National Cancer Institute. 2002;94:1365-72.

54. Hartge P, Struewing J, Wacholder S, Brody L, Tucker M. The prevalence of common BRCA1 and BRCA2 mutations among Ashkenazi Jews. The American Journal of Human Genetics. 1999;64:963-70.
55. Geary J, Sasieni P, Houlston R, Izatt L, Eeles R, Payne SJ, et al. Gene-related cancer spectrum in families with hereditary non-polyposis colorectal cancer (HNPCC). Familial Cancer. 2008;7:163-72.
56. Kastrinos F, Mukherjee B, Tayob N, Wang F, Sparr J, Raymond V, et al. Risk of pancreatic cancer in families with Lynch syndrome. JAMA. 2009;302:1790-5.
57. Lowenfels A, Maisonneuve P, Whitcomb D. Risk factors for cancer in hereditary pancreatitis. International Hereditary Pancreatitis Study Group. Medical Clinics of North America. 2000;84:565-73.
58. Lowenfels A, Maisonneuve P, DiMagno E, Elitsur Y, Gates L, Perrault J, et al. Hereditary pancreatitis and the risk of pancreatic cancer. International Hereditary Pancreatitis Study Group. Journal of the National Cancer Institute. 1997;89:442-6.
59. Schneider R, Slater E, Sina M, Habbe N, Fendrich V, Matthäi E, et al. German national case collection for familial pancreatic cancer (FaPaCa): ten years experience. Familial Cancer. 2011;10:323-30.
60. McFaul C, Greenhalf W, Earl J, Howes N, Neoptolemos J, Kress R, et al. Anticipation in familial pancreatic cancer. Gut. 2006;55:252-8.
61. de Jong K, Nio C, Hermans J, Dijkgraaf M, Gouma D, van Eijck C, et al. High prevalence of pancreatic cysts detected by screening magnetic resonance imaging examinations. Clinical Gastroenterology and Hepatology. 2010;8(9):806-11.
62. Spinelli K, Fromwiller T, Daniel R. Cystic pancreatic neoplasms: observe or operate Annals of Surgery. 2004;239(5):651-9.
63. Robinson S, Scott J, Oppong K, White S. What to do for the incidental pancreatic cystic lesion? Surgical Oncology. 2014;23:1-9.
64. Correa-Gallego C, Ferrone C, Thayer S, Wargo J, Warshaw A, Fernandez-Del Castillo C. Incidental pancreatic cysts: do we really know what we are watching? Pancreatology. 2010;10:144-50.
65. Yoon W, Brugge W. Pancreatic cystic neoplasms: diagnosis and management. Gastroenterology Clinics of North America. 2012;41:103-18.
66. Farrell J, Fernandez-del Castillo C. Pancreatic cystic neoplasms: management and unanswered questions. Gastroenterology. 2013;144:1303-15.
67. van der Waaij L, van Dullemen H, Porte R. Cyst fluid analysis in the differential diagnosis of pancreatic cystic lesions: a pooled analysis. Gastrointestinal Endoscopy. 2005;62:383-9.
68. Strobel O, Z'Graggen K, Schmitz-Winnenthal F, Friess H, Kappeler A, Zimmermann A. Risk of malignancy in serous cystic neoplasms of the pancreas. Digestion. 2003;68:24-33.
69. Sun H, Kim S, Kim M, Lee J, Han J, Choi B. CT imaging spectrum of pancreatic serous tumors: based on new pathologic classification. European Journal of Radiology. 2010;75:e45-e55.

70. Crippa S, Salvia R, Warshaw A, Dominguez I, Bassi C, Falconi M. Mucinous cystic neoplasm of the pancreas is not an aggressive entity: lessons from 163 resected patients. Annals of Surgery. 2008;247:571-9.
71. Brugge W. Cystic pancreatic lesions: can we diagnose them accurately what to look for? FNA marker molecular analysis resection, surveillance, or endoscopic treatment? Endoscopy. 2006;38 (Suppl. 1)(S40-S47).
72. Cizginer S, Turner B, Bilge A, Karaca C, Pitman M, Brugge W. Cyst fluid carcinoembryonic antigen is an accurate diagnostic marker of pancreatic mucinous cysts. Pancreas. 2011;40:1024-8.
73. Kucera S, Centeno B, Springett G, Malafa M, Chen Y, Weber J. Cyst fluid carcinoembryonic antigen level is not predictive of invasive cancer in patients with intraductal papillary mucinous neoplasm of the pancreas. Journal of the Pancreas. 2012;13:409-13.
74. Tanaka M, Fernandez-del Castillo C, Adsay V, Chari S, Falconi M, Jang J. International consensus guidelines 2012 for the management of IPMN and MCN of the pancreas. Pancreatology. 2012;12:183-97.
75. Brugge W, Lauwers G, Sahani D, Fernandez-del Castillo C, Warshaw A. Cystic neoplasms of the pancreas. New England Journal of Medicine. 2004;351:1218-26.
76. Waters J, Schmidt C, Pinchot J, White P, Cummings O, HA Pea. CT vs MRCP: optimal classification of IPMN type and extent. Journal of gastrointestinal surgery : official journal of the Society for Surgery of the Alimentary Tract. 2008;12:101-9.
77. Crippa S, Fernandez-Del Castillo C, Salvia R, Finkelstein D, Bassi C, Dominguez I. Mucin-producing neoplasms of the pancreas: an analysis of distinguishing clinical and epidemiologic characteristics. Clinical Gastroenterology and Hepatology. 2010;8:213-9.
78. Winter J, Yeo C, Broady J. Diagnostic, Prognostic, and Predictive Biomarkers in Pancreatic Cancer. Journal of Surgical Oncology 2013;107:15-22.
79. Gold P, Freedman S. Demonstration of tumor-specific antigens in human colonic carcinomata by immunological tolerance and absorption techniques. Journal of Experimental Medicine. 1965;121:439-62.
80. Carpelan-Holmstrom M, Louhimo J, Stenman U, Alfthan H, Haglund C. CEA, CA 19–9 and CA 72–4 improve the diagnostic accuracy in gastrointestinal cancers. Anticancer Research. 2002;22(2311-2316).
81. Koprowski H, Steplewski Z, Mitchell K, Herlyn M, Herlyn D, Fuhrer P. Colorectal carcinoma antigens detected by hybridoma antibodies. Somatic Cell Genetics. 1979;5(6):957-71.
82. Vestergaard E, Hein H, Meyer H, Grunnet N, Jorgensen J, Wolf H, et al. Reference values and biological variaton for tumor marker CA 19-9 in serum for different Lewis and secretor genotypes and evaluation of secretor and Lewis genotyping in a Caucasian population. Clinical Chemistry. 1999;45(1):54-61.
83. Kim J, Lee K, Lee J, Paik S, Rhee J, Choi K. Clinical usefullness of carbohydrate antigen 19-9 as a screening test for pancreatic cancer in an

asymptomatic population. Journal of Gastroenterology and Hepatology. 2004;19:182-6.
84. Goonetilleke K, Siriwardena A. Systematic review of carbohydrate antigen (CA 19-9) as a biochemical marker in the diagnosis of pancreatic cancer. European Journal of Surgical Oncology. 2007;33(3):266-70.
85. Kim Y, Kim H, Park J, Park D, Cho Y, Sohn C, et al. Can pre-operative CA 19-9 and CEA levels predict resectability of patients with pancreatic adenocarcinoma. Journal of Gastroenterology and Hepatology. 2009;24:1869-75.
86. Schlieman M, Ho H, Bold R. Utility of tumour markers in determining resectability of pancreatic cancer. Archives of Surgery. 2003;138:951-5.
87. Berger A, Meszoely I, Ross E, Watson J, Hoffmann J. Undetectable preoperative levels of serum CA 19-9 correlate with improved survival for patients with resectable pancreatic adenocarcinoma. Annals of Surgical Oncology. 2004;11:644-9.
88. Smith R, Bosonnet L, Ghaneh P, Raraty M, Sutton R, Campbell F, et al. Pre-operative CA19-9 levels and lymph node ratio are independent predictors of survival in patients with resected pancreatic ductal adenocarcinoma. Digestive Surgery. 2008;25(3):226-32.
89. Kondo N, Murakami Y, Uemura K, Hayashidani Y, Sudo T, Hashimoto Y, et al. Prognostic impact of perioperative serum CA 19-9 levels in patients with resectable pancreatic cancer. Annals of Surgical Oncology. 2010;17(9):2321-9.
90. Koopmann J, Rosenzweig C, Zhang Z, Canto M, Brown D, Hunter M, et al. Serum markers in patients with resectable pancreatic adenocarcinoma: macrophage inhibitory cytokine 1 versus CA19-9. Clinical Cancer Research. 2006;12:442-6.
91. Duffy M, Sturgeon C, Lamerz R, Haglund C, Holubec V, Klapdor R, et al. Tumor markers in pancreatic cancer: a European Group on Tumor Markers (EGTM) status report. Annals of Oncology. 2011;21(441-447).
92. Duraker N, Hot S, Polat Y, Hobek A, Gencler N, Urhan N. CEA, CA 19-9 and CA 125 in the differential diagnosis of benign and malignant pancreatic diseases with or without jaundice. European Journal of Surgical Oncology. 2007;95:142-7.
93. Marrelli D, Caruso S, Pedrazzani C, Neri A, Fernandes E, Marini M, et al. CA 19-9 serum levels in obstructive jaundice: clinical value in benign and malignant conditions. American Journal of Surgery. 2009;198(3):333-9.
94. Hong S, Park J, Hruban R, Goggins M. Molecular signatures of pancreatic cancer. Archives of Pathology and Laboratory Medicine. 2011;135:716-27.
95. Lohr M, Kloppel G, Maisonneuve P, Lowenfels A, Luttges J. Frequency of K-ras mutations in pancreatic intraductal neoplasias associated with pancreatic ductal adenocarcinoma and chronic pancreatitis: a meta-analysis. Neoplasia. 2005;7:17-23.
96. Herreros-Villanueva M, Rodrigo M, Claver M, Munz P, Lastra E, Garcia-Giron C, et al. KRAS, BRAF, EGFR and HER2 gene status in a Spanish population of colorectal cancer. Molecular Biology Reports. 2011;38:1315-20.

97. Rozenblum E, Schutte M, Goggins M, Hahn S, Panzer S, Zahurak M, et al. Tumor-suppressive pathways in pancreatic carcinoma. Cancer research. 1997;57:1731-4.
98. Hustinx S, Leoni L, Yeo C, Brown P, Goggins M, Kern S, et al. Concordant loss of MTAP and p16/CDKN2A expression in pancreatic intraepithelial neoplasia: evidence of homozygous deletion in a noninvasive precursor lesion. Modern Pathology. 2005;18:959-63.
99. Hruban R, Goggins M, Parsons J, Kern S. Progression model for pancreatic cancer. Clinical Cancer Research. 2000;6:2969-72.
100. G G, Neal C P, CJ P, WP S, Berry D. Molecular prognostic markers in pancreatic cancer: a systematic review. European Journal of Cancer. 2005;41:2213-36.
101. Sturm P, Hruban R, Ramsoekh T, Noorduyn L, Tytgat G, Gouma D, et al. The potential diagnostic use of K-ras codon 12 and p53 alterations in brush cytology from the pancreatic head region. The Journal of Pathology. 1998;186:247-53.
102. Blackford A, Serrano OK, Wolfgang CL, Pamigiani G, Jones S, Zhang X, et al. SMAD4 gene mutations are associated with poor prognosis in pancreatic cancer. Clinical Cancer Research. 2009;15(4674-4679).
103. Bootcov M, Bauskin A, Valenzuela S, et al MIC-1, a novel macrophage inhibitory cytokine, is a divergent member of the TGF-β superfamily. Proceedings of the National Academy of Sciences of the United States of America. 1997;94(11514-11519).
104. Klimp A, de Vries E, Scherphof G, Daemen T. A potential role of macrophage activation in the treatment of cancer. Critical Reviews in Oncology/Haematology. 2002; 44: 143–61;44:143-61.
105. Koopmann J, Rosenzweig CN, Zhang Z, et al. Serum markers in patients with resectable pancreatic adenocarcinoma: macrophage inhibitory cytokine 1 versus CA19-9. Clin Cancer Res. 2006;12:442-6.
106. Simeone D, Ji B, Banerjee M, Arumugam T, Li D, Anderson M, et al. CEACAM1, a novel serum biomarker for pancreatic cancer. Pancreas. 2007;34(4):436-43.
107. Ambros V. The functions of animal microRNAs. Nature. 2004;431(7006):350-5.
108. Ma M, Kong X, Weng M, Cheng K, Gong W, Quan ZW, et al. Candidate microRNA biomarkers of pancreatic ductal adenocarcinoma: meta-analysis, experimental validation and clinical significance. Journal of Experimental & Clinical Cancer Research. 2013;32(1):71.
109. Frampton AE, Castellano L, Colombo T, Giovannetti E, Krell J, Jacob J, et al. MicroRNAs cooperatively inhibit a network of tumor suppressor genes to promote pancreatic tumor growth and progression. Gastroenterology. 2014;146(1):268-77 e18.
110. Wang J, Chen J, Chang P, LeBlanc A, Donghui L, Abbruzzesse J, et al. MicroRNAs in Plasma of Pancreatic Ductal Adenocarcinoma Patients as Novel Blood-Based Biomarkers of Disease. Cancer Prevention Research. 2009;2(9):807-13.
111. Li A, Yu J, Kim H, K, Wolfgang C, Canto M, Hruban R, et al. MicroRNA array analysis finds elevated serum miR-1290 accurately

distingushes patients with low-stage pancreatic cancer form healthy disease controls. Clinical Cancer Research. 2013;19(13):3600-10.
112. Melo S, Luecke L, Kahlert C, Fernandez A, Gammon S, Kaye J, et al. Glypican-1 identifies cancer exosomes and detects early pancreatic cancer. Nature. 2015;523:177-82.
113. Kisiel J, Yab T, Taylor W, et al Stool DNA testing for the detection of pancreatic cancer: assessment of methylation marker candidates. Cancer. 2012;118:2623-31.
114. German J, Hammock B, Watkins S. Metabolomics: building on a century of biochemistry to guide human health. Metabolomics : Official journal of the Metabolomic Society. 2005;1(1):3-9.
115. Nicholson J, Lindon J, Holmes E. 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. Xenobiotica; the fate of foreign compounds in biological systems. 1999;29(11):1181-9.
116. Wishart D, Tzur D, Knox C, Eisner R, Guo A, Young N, et al. HMDB: the Human Metabolome Database. Nucleic acids research. 2007;35(Database issue):D521-6.
117. German J, Hammock B, Watkins S. Metabolomics: building on a century of biochemistry to guide human health. Metabolomics : Official journal of the Metabolomic Society. 2005;1:3-9.
118. Palmnas M, Vogel HJ. The Future of NMR Metabolomics in Cancer Therapy: Towards Personalizing Treatment and Developing Targeted Drugs? Metabolites. 2013;3:373-96.
119. Dettmer K, Aronov A, Hammock B. Mass Spectrometry-Based Metabolomics. Mass Spectrometry Reviews. 2007;26:51-8.
120. Saude E, Slupsky C, Sykes B. Optimization of NMR analysis of biological fluids for quantitative accuracy. Metabolomics : Official journal of the Metabolomic Society. 2006;2(3):113-23.
121. Bloch F, Hansen WW, Packard M. Nuclear Induction. Physical Review. 1946;69(3-4):127-.
122. Jacobsen N. NMR Spectroscopy Explained: Simplified Theory, Applications and Examples for Organic Chemistry and Structural Biology. New Jersey: John Wiley & Sons; 2007.
123. Claridge T. High-Resolution NMR Techniques in Organic Chemistry. Baldwin JE, Williams RM, editors. Oxford, UK: Elsevier Science Ltd; 1999.
124. Trygg J, Holmes E, Lundstedt T. Chemometrics in metabonomics. Journal of proteome research. 2007;6(2).
125. Craig A, Cloareo O, Holmes E, Nicholson JK, Lindon J. Scaling and Normalization Effects in NMR Spectroscopic Metabonomic Data Sets. Analytical Chemistry. 2006;78(7):2262-7.
126. Keun H, Ebbels T. Improved analysis of multivariate data by variable stability scaling: application to NMR-based metabolic profiling. Analytica Chimica Acta. 2003;490(1-2):265-76.
127. Van den Berg R, Hoefsloot H, Westerhuis JA, Smilde A, van der Werf M. Centering, scaling, and transformations: improving the biological information content of metabolomics data. BMC Genomics. 2006;7:142.

128. Westerhuis J, Hoefsloot H, Smit S, Vis D, Smilde A, Van Velzen E, et al. Assessment of PLSDA cross validation. Metabolomics : Official journal of the Metabolomic Society. 2008;4:81-9.
129. Bezabeh T, Ijare O, Albiin N, Arnelo U, Lindberg B, Smith I. Detection and quantification of D-glucuronic acid in human bile using 1H NMR spectroscopy: relevance to the diagnosis of pancreatic cancer. Magma. 2009;22(5):267-75.
130. Zhang L, Jin H, Guo X, Yang Z, Zhao L, Tang S, et al. Distinguishing pancreatic cancer from chronic pancreatitis and healthy individuals by (1)H nuclear magnetic resonance-based metabonomic profiles. Clinical biochemistry. 2012;45(13-14):1064-9.
131. Bathe OF, Shaykhutdinov R, Kopciuk K, Weljie A, McKay A, Sutherland F, et al. Feasibility of identifying pancreatic cancer based on serum metabolomics. Cancer Epidemiology, Biomarkers & Prevention. 2011;20(1):140-7.
132. OuYang D, Xu J, Huang H, Chen Z. Metabolomic Profiling of Serum from Human Pancreatic Cancer Patients Using 1H NMR Spectroscopy and Principal Component Analysis. Appl Biochem Biotechnol. 2011;165(1):148-54.
133. Tesiram YA, Lerner M, Stewart C, Njoku C, Brackett DJ. Utility of nuclear magnetic resonance spectroscopy for pancreatic cancer studies. Pancreas. 2012;41(3):474-80.
134. Davis V, Schiller D, Eurich D, Bathe O, Sawyer M. Pancreatic Ductal Adenocarcinoma is Associated with a Distinct Urinary Metabolomic Signature. Annals of Surgical Oncology. 2012.
135. Napoli C, Sperandio N, Lawlor R, Scarpa A, Molinari H, Assfalg M. Urine metabolic signature of pancreatic ductal adenocarcinoma by (1)h nuclear magnetic resonance: identification, mapping, and evolution. Journal of proteome research. 2012;11(2):1274-83.
136. Wang J, Ma C, Liao Z, Tian B, Lu J. Study on chronic pancreatitis and pancreatic cancer using MRS and pancreatic juice samples. World Journal of Gastroenterology. 2011;17(16):2126-30.
137. Bland M. An introduction to medical statisitcs. 3 ed. Oxford: Oxford University Press; 2005.
138. Broadhirst D, Kell D. Statistical strategies for avoiding false discoveries in metabolomics and related experiments. Metabolomics : Official journal of the Metabolomic Society. 2006;2(4):171-96.
139. Nicholson J, Foxall P. 750 MHz 1H and 1H-13C NMR spectroscopy of human blood plasma. Analytical Chemistry. 1995;67:793-811.
140. Lenz E, Bright J, Wilson I, Morgan S, Nash A. NMR-based metabonomic study of urine and plasma samples obtained from healthy human subjects. Journal of Pharmaceutical and Biomedical Analysis. 2003;33(5):1103-15.

Patient Information Sheet 1 – 27/04/2012 – Version 2

Part 1:

Metabolomics & Novel Biomarkers for Pancreatic Cancer (Pancreatic resection group)

You are invited to take part in a RESEARCH study. Before you decide, it is important you understand why the research is being done and what it will involve. Please take time to read the following information carefully. Ask us if there is anything that is not clear, or if you would like more information. Take time to decide whether you wish to take part.

Part 1 tells you the purpose of this study and what will happen to you if you take part

Part 2 gives you more information about the conduct of the study

Thank you for reading this

What is the purpose of the study?

Cancer of the pancreas may be difficult to diagnose and often presents late. We aim to identify novel markers (biomarkers) to aid diagnosis and guide treatment. The markers of interest are small molecules which are released during metabolism. Metabolism is the set of chemical reactions that happen in the body to sustain life. We know that metabolism is altered in a disease state and that different products of metabolism are released. Our intent is to determine a signature that could be used to distinguish pancreatic cancer from non-cancerous conditions of the pancreas or biliary tree.

Why have I been invited?

You have been chosen for this study because you are about to undergo surgery for suspected pancreatic cancer. We plan to analyse a sample of your blood, bile, pancreatic fluid and urine for various small molecules or products of metabolism. We will then compare your metabolite levels with patients who have undergone an operation for a non-cancerous condition. We plan to include a total of 100 patients with cancer of the pancreas and 50 patients who have undergone an operation for a non-cancerous condition.

Do I have to take part?

It is entirely up to you if you wish to participate or not. We would like to reassure you that if you do not wish to participate this will not affect the standard of your care in any way. You are free to withdraw at any time, without giving a reason.

What will happen to me if I take part?

If you agree to participate we will ask you to sign a consent form. During your operation we will withdraw 2 teaspoons of blood from a line that will have already been inserted for the purpose of your operation by your anaesthetist. In addition your operating surgeon will collect a sample of bile and pancreatic fluid, which is released but normally discarded during your operation. A sample of urine will be collected from your urine catheter bag. You will be involved in the study only on the operative day.

What are the potential benefits of taking part?

Although there are no direct benefits to you from participating in this study, it will provide us with important information which may be of great value to improve the care of future patients.

What are the possible disadvantages and risks of taking part?

There are no particular risks or disadvantages of taking part in this study. All your care will proceed as planned. Now new treatment will be added or withheld.

Will my taking part be kept confidential?

Yes. We follow ethical and legal practice and all the information about you will be handled in confidence. The details are included in Part 2.

What if I have concerns?

Any complaint about the way you have been dealt with during the study will be addressed. The detailed information on this is given in Part 2.

This completes Part 1. If the information in Part 1 has interested you and you are considering participation, please read the additional information in Part 2 before making any decision.

Part 2:

What will happen if I don't want to carry on with the study?

You are free to withdraw from the study any time you want. It would be your decision to allow us or not to analyse any information which we may have already obtained from your samples. Any stored samples identifiable as yours could be destroyed if you wish.

What if there is a problem?

If you have any concerns about this study or the way it has been organised you should contact the principal investigator (Mr Andrew Smith) at the address or telephone number below. In the unlikely event that something goes wrong and you are harmed during the research the normal NHS complaints mechanism will still be available to you.

Principal Investigator

Mr Andrew Smith
Consultant Pancreatic & General Surgeon
St James Institute of Oncology
St James Hospital
Beckett Street
Leeds
LS9 7TF
(0113) 2064719

Alternatively you may contact the Patient Advice and Liaison Service (PALS) at the Leeds Teaching Hospitals on (0113) 2067168

Will my taking part in this study be kept confidential?

We will keep all of your data strictly confidential. All of your samples for the purpose of research will be coded immediately after collection. No identifying information other than the code will be attached to your samples. The deciphering of the code to identify your personal details can only be done on a dedicated secure computer within St James Hospital by members of the research team. Any publication of the results will be completely anonymous.

Involvement of the General Practitioner (GP)

It is not necessary to inform your GP about your participation in this study

What will happen to any samples I give?

Your samples will be stored securely within The Leeds Teaching Hospitals NHS Trust. Analysis will take place within the School of Chemistry at the University of Leeds. The samples will not be used for any other future studies. At the end of the study, they will be disposed of securely. No samples will be transferred outside the UK.

Will any genetic tests be done?

No. There will not be any sort of genetic tests conducted using your blood samples.

What will happen to the results of the research study?

The results from this study may be presented at medical meetings or published in medical journals. We can assure you that you will not be identifiable in any of the results. We regret that we cannot inform you of your individual sample analysis results. However, upon request we are happy to provide an overview of the findings of the study once it is completed.

Who is organising and funding the study?

The study is organised and funded by the Department of Surgery within the Leeds Teaching Hospitals NHS Trust. The doctors involved in the study are not being paid for conducting the study

Who has reviewed the study?

This study has been reviewed by the Northampton Research Ethics Committee.

Patient Information Sheet 2 – 27/04/2012– Version 2

Part 1:

Metabolomics & Novel Biomarkers for Pancreatic Cancer
(Non-Cancer control group)

You are invited to take part in a RESEARCH study. Before you decide, it is important you understand why the research is being done and what it will involve. Please take time to read the following information carefully. Ask us if there is anything that is not clear, or if you would like more information. Take time to decide whether you wish to take part.

Part 1 tells you the purpose of this study and what will happen to you if you take part

Part 2 gives you more information about the conduct of the study

Thank you for reading this

What is the purpose of the study?

Cancer of the pancreas may be difficult to diagnose and often presents late. We aim to identify novel markers (biomarkers) to aid diagnosis and guide treatment. The markers of interest are small molecules which are released during metabolism. Metabolism is the set of chemical reactions that happen in the body to sustain life. We know that metabolism is altered in a disease state and that different products of metabolism are released. Our intent is to determine a signature that could be used to distinguish pancreatic cancer from non-cancerous conditions of the pancreas or biliary tree.

Why have I been invited?

You have been chosen for the study because you are about to undergo an operation for a non-cancerous condition of the biliary tree or pancreas. We plan to analyse a sample of your blood, bile and urine for various small molecules or products of metabolism. We will then compare your metabolite levels with patients who have undergone an operation for pancreatic cancer. We plan to include a total of 100 patients with cancer of the pancreas and 50 patients who have undergone an operation for a non-cancerous condition.

Do I have to take part?

It is entirely up to you if you wish to participate or not. We would like to reassure you that if you do not wish to participate this will not affect the standard of your care in any way. You are free to withdraw at any time, without giving a reason.

What will happen to me if I take part?

If you agree to participate we will ask you to sign a consent form. In the anaesthetic room we will withdraw 2 teaspoons of blood from a cannula or line that will be inserted for the purpose of your operation by your anaesthetist. A sample of bile will be aspirated from your gallbladder once it has been removed from your body. A sample of urine will also be collected from you prior to your operation. Your care will not be affected in any way by the collection of these samples. You will be involved in the study only on the operative day.

What are the potential benefits of taking part?

Although there are no direct benefits to you from participating in this study, it will provide us with important information which may be of great value to improve the care of future patients.

What are the possible disadvantages and risks of taking part?

There are no particular risks or disadvantages of taking part in this study. All your care will proceed as planned. Now new treatment will be added or withheld.

Will my taking part be kept confidential?

Yes. We follow ethical and legal practice and all the information about you will be handled in confidence. The details are included in Part 2.

What if I have concerns?

Any complaint about the way you have been dealt with during the study will be addressed. The detailed information on this is given in Part 2.

This completes Part 1. If the information in Part 1 has interested you and you are considering participation, please read the additional information in Part 2 before making any decision.

Part 2:

What will happen if I don't want to carry on with the study?

You are free to withdraw from the study any time you want. It would be your decision to allow us or not to analyse any information which we may have already obtained from your samples. Any stored samples identifiable as yours could be destroyed if you wish.

What if there is a problem?

If you have any concerns about this study or the way it has been organised you should contact the principal investigator (Mr Andrew Smith) at the address or telephone number below. In the unlikely event that something goes wrong and you are harmed during the research the normal NHS complaints mechanism will still be available to you.

Principal Investigator

Mr Andrew Smith
Consultant Pancreatic & General Surgeon
St James Institute of Oncology
St James Hospital
Beckett Street
Leeds
LS9 7TF
(0113) 2064719

Alternatively you may contact the Patient Advice and Liaison Service (PALS) at the Leeds Teaching Hospitals on (0113) 2067168

Will my taking part in this study be kept confidential?

We will keep all of your data strictly confidential. All of your samples for the purpose of research will be coded immediately after collection. No identifying information other than the code will be attached to your samples. The deciphering of the code to identify your personal details can only be done on a dedicated secure computer within St James Hospital by members of the research team. Any publication of the results will be completely anonymous.

Involvement of the General Practitioner (GP)

It is not necessary to inform your GP about your participation in this study

What will happen to any samples I give?

Your samples will be stored securely within The Leeds Teaching Hospitals NHS Trust. Analysis will take place within the School of Chemistry at the University of Leeds. The samples will not be used for any other future studies. At the end of the study, they will be disposed of securely. No samples will be transferred outside the UK.

Will any genetic tests be done?

No. There will not be any sort of genetic tests conducted using your blood samples.

What will happen to the results of the research study?

The results from this study may be presented at medical meetings or published in medical journals. We can assure you that you will not be identifiable in any of the results. We regret that we cannot inform you of your individual sample analysis results. However, upon request we are happy to provide an overview of the findings of the study once it is completed.

Who is organising and funding the study?

The study is organised and funded by the Department of Surgery within the Leeds Teaching Hospitals NHS Trust. The doctors involved in the study are not being paid for conducting the study

Who has reviewed the study?

This study has been reviewed by the Northampton Research Ethics Committee.

The Leeds Teaching Hospitals NHS
NHS Trust

Consent Form 1 – 27/04/2012 – Version 2.0

Title of project: Metabolomics & Novel Biomarkers for Pancreatic Cancer (Pancreatic resection group)

Name of researcher: Mr Andrew Smith, Mr Rowland Storey

	Please initial
1. I confirm that I have read and understand the information sheet 1 dated 27/04/2012 – version 2.0 for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.	☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, without my medical care or legal rights being affected.	☐
3. I understand that the relevant sections of my medical notes and data collected during the study may be looked at by the individuals involved in this study, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records	☐
4. I agree to take part in the above study.	☐

Name of Patient	Date	Signature

Name of Person taking consent	Date	Signature

The Leeds Teaching Hospitals NHS
NHS Trust

Consent Form 2 – 27/04/2012 – Version 2.0

Title of project: Metabolomics & Novel Biomarkers for Pancreatic Cancer (Non-cancer control group)

Name of researcher: Mr Andrew Smith, Mr Rowland Storey

	Please initial
1. I confirm that I have read and understand the information sheet 2 dated 27/04/2012 – version 2.0 for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.	☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, without my medical care or legal rights being affected.	☐
3. I understand that the relevant sections of my medical notes and data collected during the study may be looked at by the individuals involved in this study, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records	☐
4. I agree to take part in the above study.	☐

_____	_____	_____
Name of Patient	Date	Signature
_____	_____	_____
Name of Person taking consent	Date	Signature