THE GENOMIC AND METABOLOMIC PROFILING OF PANCREAS CANCER

A thesis submitted to the University of Manchester for the degree of
Doctor of Medicine in the Faculty of Medical and Human Sciences

2014

Sudip Sanyal
School of Medicine
Table of Contents

Chapter 1 - An Introduction to the Pathology and Genetics of Pancreas Cancer......... 18

1.1. Introduction ....................................................................................................................... 18
1.2. Risk Factors ....................................................................................................................... 19
1.3. Precursor Lesions ............................................................................................................... 20
1.4. Pathology ............................................................................................................................ 24
1.5. Genetic origins of pancreas cancer ................................................................................. 30
1.6. Molecular genetics of pancreas cancer ......................................................................... 31

1.6.1. Genome wide studies ................................................................................................. 32
1.6.2. Signalling pathways in the development of pancreatic cancer ......................... 34
1.6.3. Chromosomal Alterations ......................................................................................... 37
1.6.4. Microsatellite instability ............................................................................................ 38
1.6.5. Specific gene alterations ............................................................................................. 38
1.6.6. Telomere length abnormalities ................................................................................. 41
1.6.7. Epigenetic Silencing ................................................................................................. 41
1.6.8. Gene over-expression ................................................................................................. 42

Chapter 2 - A study of the natural history of pancreatic cystic neoplasms .......... 46

Introduction .......................................................................................................................... 46
Aims .......................................................................................................................................... 46
2.1. Pancreatic cystic tumours ............................................................................................. 46
2.2. Intraductal Papillary Mucinous neoplasms ............................................................. 48
2.3. Demographics ................................................................................................................. 48
2.4. Clinical features ............................................................................................................... 48
2.5. Classification .................................................................................................................... 49
2.6. Pathology .......................................................................................................................... 50

2.6.1. Morphology .................................................................................................................. 50
2.6.2. IPMN versus PanIN ..................................................................................................... 52
2.6.3. Architectural Detail ........................................................................................................ 52

2.7. Molecular pathology ....................................................................................................... 53
2.8. Practical Applications of molecular biology of IPMNs ........................................... 54
2.9. Management of cystic tumours .................................................................................... 56

2.9.1. Criteria for Surgery ..................................................................................................... 56
2.9.2. Diagnostic and Management Algorithm ............................................................... 57

2.10. Methods ......................................................................................................................... 57

2.10.1. Setting ......................................................................................................................... 57
2.10.2. Design ........................................................................................................................ 58
2.10.3. Methods ....................................................................................................................... 58
2.10.4. Study population ........................................................................................................ 58
2.10.5. Imaging pathways ....................................................................................................... 59
2.10.6. Cyst Characteristics .................................................................................................... 63

2.11. Management plan .......................................................................................................... 63

2.11.1. Inclusion and exclusion criteria .............................................................................. 64

2.12. Results ............................................................................................................................. 64

2.12.1. Clinical profile ............................................................................................................ 64
2.12.2. Results of investigations .......................................................................................... 67
2.12.3. Management and Outcomes .................................................................................... 68
2.12.4. Main duct IPMN ........................................................................................................ 69
2.12.5. Pancreatic cystic lesions ≤ 3 cm in size ................................................................. 71
2.12.6. Branch duct IPMN ..................................................................................................... 73
Chapter 3 - Poly A RT-PCR measurements of indicator genes in pancreas cancer and pancreatic cystic tumours

3.1. Introduction........................................................................................................................................ 81
3.2. Molecular Pathology of Pancreas Cancer .......................................................................................... 81
3.3. Development and use of Poly A PCR ................................................................................................. 82
  3.3.1. Aim of this study .......................................................................................................................... 83
3.4. Patients and Methods ........................................................................................................................ 83
  3.4.1. Study Design ................................................................................................................................. 83
  3.4.2. Patients ........................................................................................................................................ 83
  3.4.3. Data collection ............................................................................................................................... 87
  3.4.4. Sample Storage ............................................................................................................................. 88
  3.4.5. Ethical Approval ........................................................................................................................... 88
  3.4.6. Sample collection ......................................................................................................................... 89
3.5. Selection of Genes ............................................................................................................................... 90
3.6. Sequencing of primers and probes ..................................................................................................... 93
3.7. Materials and Methods (292) ......................................................................................................... 95
3.8. Results ............................................................................................................................................. 96
  3.8.1 RNA Extraction .............................................................................................................................. 96
  3.8.2. Pancreatic Juice ............................................................................................................................. 96
  3.8.3. Pancreatic Tissue .......................................................................................................................... 97
  3.8.4. Poly A PCR .................................................................................................................................. 97
  3.8.5. Real-time PCR ............................................................................................................................. 97
3.9. Discussion ....................................................................................................................................... 103
  3.9.1. Methodology ................................................................................................................................. 104
  3.9.2. Choice of methodology (Poly A PCR) ........................................................................................ 105
  3.9.3. Gene expression between pancreatic juice and tumour tissue in pancreas cancer ................. 106
  3.9.4. Gene expression between pancreatic juice and tumour tissue in pancreatic cystic tumours ... 107
3.10. Conclusion ...................................................................................................................................... 107

Chapter 4 - Metabolomic Studies in Pancreatic Cancer ............................................................................. 109

4.1. Introduction ...................................................................................................................................... 109
4.2. Approaches to investigating the metabolome ................................................................................... 110
  4.2.1 Background .................................................................................................................................. 110
  4.2.2. Overview of laboratory techniques ............................................................................................ 110
4.3. Review of metabolomic studies in pancreas cancer ......................................................................... 111
  4.3.1. Profiling of pancreatic tumor tissue .............................................................................................. 114
  4.3.2. Plasma profiling ........................................................................................................................... 115
  4.3.3. Serum profiling ............................................................................................................................ 116
  4.3.4. Profiling using saliva samples ...................................................................................................... 119
  4.3.5. Profiling using bile samples ......................................................................................................... 119
  4.3.6. Profiling using urine samples ...................................................................................................... 120
4.4. Animal studies ................................................................................................................................. 120
  4.4.1. Sprague-Dawley rat model .......................................................................................................... 120
  4.4.2. Transgenic pancreatic cancer rat model with expression of human K-rasG12V oncogene and induction is regulated by the Cre/lox system ........................................................................ 121
  4.4.3. Nude mouse xenograft model using injected pancreatic cancer cell SW1990 ......................... 122
  4.4.4. Perfused pancreatic cancer cell lines and induced pancreatic cancer ......................................... 122
4.5. Factors that influence metabolite levels ............................................................................................. 122
Conclusions and Future work ............................................................................................... 159

A registry of pancreatic cystic tumours .............................................................................. 159
Poly A RT-PCR measurement of indicator genes in pancreas cancer and pancreatic cystic
tumours .................................................................................................................................. 162
Metabolomic assessment of pancreatic cancer ................................................................. 165

References ............................................................................................................................... 168

Appendix .................................................................................................................................. 195

Word count: 52,155
List of Tables

Table 1: Risk factors in the development of pancreas cancer ....................................................... 20
Table 2: Classification of pancreatic neoplasms (49) ................................................................. 29
Table 3: Differences in molecular characteristics of the precursor lesions of pancreas cancer (Adapted from Cooper et al.) (190) ................................................................................................ 54
Table 4: Clinical profile of patients presenting with pancreatic cystic lesions ....................... 66
Table 5: Diagnosis after triple modality investigation and MDT review ........................................ 67
Table 6: Management of the whole population of suspected pancreatic cystic tumours .......... 68
Table 7: Management of pancreatic cystic lesions less than 3cm in size ..................................... 72
Table 8: Patient and sample characteristics .................................................................................. 85
Table 9: A list of the selected 'indicator' genes .............................................................................. 92
Table 10: List of primer sequences ............................................................................................... 93
Table 11: List of probe sequences ................................................................................................. 94
Table 12: Spearman's correlation coefficients in paired samples from patients with pancreas cancer ........................................................................................................................................... 101
Table 13: Comparison of gene expression levels between paired samples in pancreatic cystic tumours ........................................................................................................................................ 102
Table 14: Summary of metabolomic studies in pancreatic cancer .............................................. 111
Table 15: Patient demographics .................................................................................................. 127
Table 16: LCMS ESI + PCA results ............................................................................................. 133
Table 17: LCMS ESI - PCA scores .............................................................................................. 134
Table 18 - Significant positive ion variables ................................................................................. 147
List of Figures

Figure 1: Pancreatic intraepithelial Neoplasia: PanIN-1A (flat), PanIN-1B (papillary), PanIN-2 (papillary with nuclear change), and PanIN-3 (severe atypical cells with mitoses, budding and luminal necrosis). (31) .................................................................................................................... 21

Figure 2: Main duct IPMN demonstrating a pancreatic ductal diameter >1cm with mucin within the duct (34) ......................................................................................................................................... 22

Figure 3: Branch duct intraductal papillary mucinous neoplasm demonstrating mucinous cystic lesions in the branch duct without main duct dilatation (34) .................................................................................. 23

Figure 4: Mucinous cystic neoplasm with no ductal communication on cut section (34) ........ 24

Figure 5: Gross appearance of pancreas cancer........................................................................... 25

Figure 6: Cut section of pancreas cancer specimen..................................................................... 26

Figure 7: Well-differentiated adenocarcinoma with neoplastic tubular glands lined with cuboidal or cylindrical cells with large, irregular nuclei and clear mucinous cytoplasm (x200). (55) ............. 27

Figure 8: Poorly differentiated adenocarcinoma of the pancreas demonstrating small irregular glands, marked nuclear pleomorphism and no cytoplasmic mucin (x200). (56) ......................... 28

Figure 9: Sequence of genetic alterations involved in the progression of PanIN to pancreatic adenocarcinoma in pancreatic ductal cells (32, 72) ......................................................................................................................... 32

Figure 10: Genetic changes in pancreatic ductal adenocarcinoma depicting the core signalling pathways and the genes involved (adapted from Samuel, N. and Hudson, T.J., 2012) .......... 33

Figure 11: Core signalling pathways incorporating the addition of Axon guidance and chromatin regulation in the development of pancreatic ductal adenocarcinoma (adapted from Yachida et al., 2013) ............................................................. 34

Figure 12: Intracellular signalling pathways involved in pancreas cancer (84) ......................... 35

Figure 13: Mucinous cystic neoplasm with ovarian type stroma (x200)(162) ............................. 47

Figure 14: Types of pancreatic cystic tumours (183) ................................................................. 49

Figure 15: Main duct IPMN - showing a dilated duct with neoplastic proliferation of mucinous epithelium. The papillae are clearly identified (184). ................................................................. 50

Figure 16: IPMN demonstrating a transition zone between low-grade to high-grade dysplasia. (186) ................................................................................................................................. 51

Figure 17: The development of mucinous adenocarcinoma with associated desmoplasia. (186) 51

Figure 18: Pancreatic intraepithelial Neoplasia (PanIN-2) (187). ............................................. 52
Figure 19: The four histological subtypes of IPMN (189) ............................................................... 53

Figure 20: Strategy for the evaluation and management of pancreatic cystic neoplasms (226) ... 57

Figure 21: Main duct IPMN on CT scan (228) .................................................................................. 60

Figure 22: MRCP imaging of branch duct IPMN in the head and body of the pancreas. The MRI image clearly shows communication with the pancreatic ductal system without main pancreatic ductal dilatation (229) ..................................................................................................................... 60

Figure 23: Mucinous cystic neoplasm on CT. There is no demonstrable communication with the pancreatic ductal system (230) ................................................................................................................................. 61

Figure 24: Fish-eye ampulla on ERCP describes mucin pouting from the ampulla in IPMN (231) 61

Figure 25: Radial EUS examination showing main duct IPMN with a dilated main pancreatic duct with evidence of an intraductal tumour (arrows) (231) ......................................................................................... 62

Figure 26: Radial EUS examination showing branch duct IPMN with no main duct communication (231) .................................................................................................................................................. 62

Figure 27: Radial EUS examination showing mural nodules seen in a dilated main pancreatic duct (231) .................................................................................................................................................. 62

Figure 28: Patient recruitment ........................................................................................................ 84

Figure 29: Chart depicting the operative procedure performed ..................................................... 88

Figure 30: Histological diagnosis from the sample population ....................................................... 88

Figure 31: Histology of the research biopsies compared with that of the reported tumour from the final histological report. Blue represents concurrence between the research biopsies and the final pathological report while red indicates that the research biopsy and the final pathological report did not match (sampling error) .................................................................................................................. 90

Figure 32: This figure illustrates the median gene expression values (Ct values in log scale, after normalisation for GAPDH) in the entire cohort of pancreatic juice and tissue samples. ............... 97

Figure 33: This graph describes the median gene expression values (Ct values in log scale after normalisation for GAPDH) in paired pancreatic juice and tissue samples in pancreas cancer ..... 98

Figure 34: The figure illustrates the median gene expression values in paired pancreatic juice and tissue samples from patients with IPMN ........................................................................................ 99

Figure 35: Relative proportions of sample histology........................................................................ 129

Figure 36: PCA plot: Negative ion mode showing disparity between cases and controls .......... 135

Figure 37: PCA: positive ion mode showing disparity between cases and controls ................. 135
Figure 38: LCMS ESI positive PCA plots for individual cancer types versus the controls showing separation between the two groups. P – pancreas cancer, A – ampullary cancer, E – endocrine cancer

Figure 39: LCMS ESI negative PCA plots for individual cancer types versus the controls showing separation between the two groups.

Figure 40: LCMS ESI positive PCA plots for samples sharing two-thirds of peaks between the individual cancers and controls.

Figure 41: LCMS ESI negative PCA plots for samples with two-thirds of peaks shared between individual cancers and controls.

Figure 42: The split in PCA components became more obvious when the graph was restricted to case samples only, with a clear demarcation seen in the negative mode. P – pancreas cancer, A – ampullary cancer, E – endocrine cancer

Figure 43: No corresponding split was observed in the PCA plots in the positive mode

Figure 44: PC-DFA plot showing a small degree of separation between the individual cancers on negative ionisation mode.

Figure 45: PC-DFA plot depicting a greater degree of separation between the three main tumour types.

Figure 46: PC-DFA loadings plot, negative mode with significant components highlighted with pink circles.

Figure 47: PC-DFA loadings plot, positive mode with significant components highlighted with pink circles.

Figure 48: PC-DFA varimax rotation plot indicating a separation between the three classes.
Abbreviations

1H-MRS = PROTON MAGNETIC RESONANCE SPECTROSCOPY
5-UMP = 5-URIDINE MONOPHOSPHATE
AIB1/NCOA3 = AMPLIFIED IN BREAST-1
AMP = ADENOSINE MONOPHOSPHATE
AMPK = ADENINE MONOPHOSPHATE-ACTIVATED PROTEIN KINASE
ARID1A = AT-RICH INTERACTIVE DOMAIN-CONTAINING PROTEIN 1A
ATM = ATAXIA TELANGIECTASIA MUTATED
C-MYC = CELLULAR HOMOLOG OF THE RETROVIRAL V- MYCONCOGENE
CACNA1G = CALCIUM CHANNEL, VOLTAGE-DEPENDENT, T TYPE, ALPHA 1G SUBUNIT
CAV1 = CAVEOLIN 1
CDKN2A = CYCLIN-DEPENDENT KINASE INHIBITOR 2A
CE-TOF-MS = CAPILLARY ELECTROPHORESIS TIME-OF-FLIGHT MASS SPECTROMETRY
CEA = CARCINOEMBRYONIC ANTIGEN
CHK2 = CHECKPOINT KINASE 2
CT = COMPUTED TOMOGRAPHY
CXCR4 = CHEMOKINE RECEPTOR TYPE 4
DNMT = DNA METHYLTRANSFERASE
DUSP6 = DUAL SPECIFICITY PHOSPHATASE 6
ECOG = EASTERN COOPERATIVE ONCOLOGY GROUP
EGFR = EPIDERMAL GROWTH FACTOR RECEPTOR
ERCP = ENDOSCOPIC RETROGRADE CHOLANGIOPANCREATOGRAPHY
ERK = EXTRACELLULAR SIGNAL-REGULATED KINASES
ESI-MS = ELECTROSPRAY IONIZATION MASS SPECTROMETRY
ESR1 = ESTROGEN RECEPTOR-1
EUS = ENDOSCOPIC ULTRASOUND
FDR = FALSE DISCOVERY RATE
FGF-BP = FIBROBLAST GROWTH FACTOR-BINDING PROTEIN
FGF-R = FIBROBLAST GROWTH FACTOR RECEPTOR
FI-FTICR-MS = FLOW-INJECTION FOURIER TRANSFORM ION CYCLOTRON RESONANCE MASS SPECTROMETRY
FI-MS/MS = FLOW-INJECTION TANDEM MASS SPECTROMETRY
GAS6 = GROWTH ARREST SPECIFIC 6
GATA6 = GATA BINDING FACTOR 6
GC/MS = GAS CHROMATOGRAPHY MASS SPECTROMETRY
GC/TOF-MS = GAS CHROMATOGRAPHY/TIME-OF-FLIGHT MASS SPECTROMETRY
GDP = GUANOSINE DIPHOSPHATE;
GPC = GLYPICAN 1
GTP = GUANOSINE TRIPHOSPHATE
HER2 = HUMAN EGFR 2
HER2/NEU = RECEPTOR TYROSINE-PROTEIN KINASE ERBB-2
HILIC = HYDROPHILIC INTERACTION CHROMATOGRAPHY
HMLH1 = MUTL HOMOLOG 1, COLON CANCER, NONPOLYPOSIS TYPE 2 (E. COLI)
HMQC OR HSQC = HETERONUCLEAR MULTIPLE/SINGLE QUANTUM COHERENCE SPECTROSCOPY
HMSH2 = MUTS PROTEIN HOMOLOG 2
HNPCC = HUMAN NON-POLYPOSIS COLORECTAL CANCER
HR-MAS = HIGH-RESOLUTION MAGIC ANGLE SPINNING PROTON MAGNETIC RESONANCE SPECTROSCOPY
HTERT = HUMAN TELOMERASE REVERSE TRANSCRIPTASE
HUSERMET = HUMAN SERUM METABOLOME IN HEALTH AND DISEASE
IGF-1 = INSULIN GROWTH FACTOR 1
IPMN = INTRADUCTAL PAPILLARY MUCINOUS NEOPLASM
K-RAS = V-KI-RAS2 KIRSTEN RAT SARCOMA VIRAL ONCOGENE HOMOLOG
LC/MS = LIQUID CHROMATOGRAPHY MASS SPECTROMETRY
LOH = LOSS OF HETEROZYGOSITY
MAGEA6 = MELANOMA ANTIGEN FAMILY A, 6
MAP2K4 = DUAL SPECIFICITY MITOGEN-ACTIVATED PROTEIN KINASE 4
MCN = MUCINOUS CYSTIC NEOPLASM
MEK = MITOGEN-ACTIVATED PROTEIN KINASE
MEKK = MITOGEN-ACTIVATED PROTEIN KINASE KINASE
MIC1 = MACROPHAGE INHIBITORY CYTOKINE 1
MIRNA = MICRORNA
MLL3 = MIXED-LINEAGE LEUKEMIA 3
MMP = MATRIX METALLOPROTEINASE
MRCP = MAGNETIC RESONANCE CHOLANGIOPANCREATOGRAPHY
MRNA = MESSENGER RNA
MSI = MICROSATELLITE INSTABILITY
MTOR = MAMMALIAN TARGET OF RAPAMYCIN
MUC1 = MUCIN 1, CELL SURFACE ASSOCIATED
MUC2 = MUCIN 2, CELL SURFACE ASSOCIATED
MUC4 = MUCIN 4, CELL SURFACE ASSOCIATED
MUC5AC = MUCIN 5AC
MYB = MYELOBLASTOSIS FAMILY
NAD = NICOTINAMIDE ADENINE DINUCLEOTIDE
NALCN = SODIUM LEAK CHANNEL, NON-SELECTIVE
NF-KAPPAB = NUCLEAR FACTOR KAPPA-LIGHT-CHAIN-ENHANCER OF ACTIVATED B CELLS
NMR = NUCLEAR MAGNETIC RESONANCE
P16 = CYCLIN-DEPENDENT KINASE INHIBITOR 2A
PANIN = PANCREATIC INTRAEPITHELIAL NEOPLASIA
PC-DFA = PRINCIPAL COMPONENT - DISCRIMINANT FACTOR ANALYSIS
PCA = PRINCIPAL COMPONENT ANALYSIS
PCR = POLYMERASE CHAIN REACTION
PENK = PROENKEPHALIN
PI3K = PHOSPHATIDYLINOSITOL 3- KINASE
PLAT = TISSUE PLASMINOGEN ACTIVATOR
PPENK = PREPROENKEPHALIN
RARB = RETINOIC ACID RECEPTOR, BETA,
RP-LC/ESI-MS = REVERSED-PHASE LIQUID CHROMATOGRAPHY/ELECTROSPRAY IONIZATION MASS SPECTROMETRY
RR = RELATIVE RISK
S100A4 = S100 CALCIUM BINDING PROTEIN A4
SCA = SEROUS CYSTADENOMA
SF3B1 = SPLICING FACTOR 3B SUBUNIT 1
SHH = SONIC HEDGEHOG
SLC16A4 = SOLUTE CARRIER FAMILY 16, MEMBER 4
SMAD/DPC4 = MOTHERS AGAINST DECAPENTAPLEGIC HOMOLOG 4/ DELETED IN PANCREATIC CARCINOMA, LOCUS 4
STAR = STEROIDOGENIC ACUTE REGULATORY PROTEIN
STK11 = SERINE/THREONINE KINASE 11
TFF = TREFOIL FAMILY
TGFB2 = TRANSFORMING GROWTH FACTOR – BETA RECEPTOR 2
TGFB = TRANSFORMING GROWTH FACTOR -BETA
THBS1 = THROMBOSPONDIN 1
TIMP3 = HUMAN TISSUE INHIBITOR OF METALLOPROTEINASES-3
TOCSY = TOTAL CORRELATION SPECTROSCOPY
UHPLC-MS = ULTRA HIGH PERFORMANCE LIQUID CHROMATOGRAPHY MASS SPECTROMETRY
UMP = URIDINE MONOPHOSPHATE
UPLC = ULTRA PERFORMANCE LIQUID CHROMATOGRAPHY
VEGF = VASCULAR ENDOTHELIAL GROWTH FACTOR RECEPTOR
WHO = WORLD HEALTH ORGANISATION
WNT = WINGLESS-RELATED INTEGRATION SITE
ZIM2 = ZINC FINGER, IMPRINTED 2
Abstract

Despite the considerable expansion of knowledge in the development of pancreatic cancer, there has been little progress made in facilitating an early diagnosis of this disease and predicting an accurate response to treatment. We aim to translate this knowledge to clinical practice by using a prospective database of precursor cystic lesions in pancreas cancer, assessing the use of over-expressed genes in pancreatic juice as a surrogate marker of these pancreas cancer and finally, downstream of these changes at the genetic level, use metabolomic techniques to look for biomarkers in pancreas cancer in serum.

In the first study, we investigate the natural history of pancreatic cystic neoplasms, specifically IPMNs, using a prospectively collected database to examine the profiles and outcomes of main duct IPMN, branch duct IPMN and cystic lesions measuring less than 3 cm in size. A total of 99 patients with suspected pancreatic cystic tumours were enrolled over 3 years. Median follow-up was 24 months (range 0 – 124). Cystic tumours comprised of 13 MD-IPMN, 40 BD-IPMN, 11 MCN and 8 adenocarcinomas among others. The complete cohort showed an overall risk of adenocarcinoma of 8%. Main duct IPMN showed a cumulative risk of 46% with evidence of progression of disease in a further 23%. The associated mortality in MD-IPMN was related to the underlying adenocarcinoma and was 38% in our group. The incidence of adenocarcinoma in branch duct IPMN was 11% with disease progression seen 13.8%. Evidence of extra-pancreatic malignancies was seen in 37.7% of patients with IPMN.

In the second study, we explore the feasibility of gene expression profiling from RNA isolated from matched pancreatic juice and tumour tissue in patients with pancreatic cancer and pancreatic cystic tumours. RNA was isolated and Poly(A) PCR was used to globally amplify the RNA. RT-PCR was used to measure expression levels of 18 genes common to both pancreas cancer and pancreatic cystic tumours. Spearman’s rank correlation test was used to examine the relationship of gene expression between pancreatic juice and tissue. One gene out of eighteen, MSLN (p<0.008), showed significant correlation in the expression levels between paired pancreatic juice and tissue samples in pancreas cancer. In the cystic tumour group, only one gene MMP-7 (p<0.01), showed a significant correlation between paired juice and tissue samples. When the whole cohort was analysed for the false discovery rate, these genes did not exhibit statistically significant correlation between the samples. RNA analysis of pancreatic juice is feasible using the Poly(A) cDNA technique and correlation of gene expression is shown to exist, albeit with low sensitivity, indicating its potential use in clinical practice with small tissue and juice samples.

In the final study, we performed a literature review on the use of metabolomics in pancreas cancer. We performed metabolic profiling of serum samples from selected cancer patients and noncancerous controls using UPHLC-MS to generate and compare the metabolic profiles in serum samples from a cohort of patients with pancreas cancer, ampullary cancer and endocrine cancer. Thirty nine serum samples (including 19 pancreatic cancers, 9 ampullary cancers and 5 endocrine cancers) and 21 matched HUSERMET controls were analysed using Ultra high performance liquid chromatography mass spectrometry (UHPLC-MS) in both positive and negative ESI modes. The output was generated as a data matrix of mass spectral features with related accurate m/z and retention time pairs. The data was then signal corrected and individual peaks were normalised and the resultant spectra were compared against a metabolite reference library and analysed using univariate and multivariate statistical tests. We found a disparity in the metabolite peaks between the cases and controls on PCA that did not permit the accurate interpretation of the data in the case study set compared to the control set. No obvious reason other than metabolite degradation during storage could account for this difference. PC-DFA analysis of metabolite peaks between pancreas cancer, ampullary cancer and endocrine cancer showed significant difference between endocrine cancers and the other two groups. Significant ESI positive metabolites included those involved in lipid pathways and metabolites involved in glucose metabolism. There is encouraging scope for studies using prospective controls to identify and develop metabolic biomarkers in pancreas cancer.
Declaration

No portion of the work referred in the thesis has been submitted in support of an application for another degree or qualification of this or any other University or other institute of learning.

Copyright

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=487), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see http://www.manchester.ac.uk/library/aboutus/regulations) and in The University’s policy on Presentation of Theses.
Dedication

To Jenifer, without whose love and support along with generous cups of tea and biscuits, this work would not have been done.
Acknowledgements

This study would not have been possible if it were not for the goodwill and generosity of each and every patient who took part in this study. At what was a difficult time in their lives, their kindness and willingness to help others was inspiring.

To Dr Richard Byers and Professor Ajith Siriwardena, for their guidance, understanding and encouragement throughout this work and for funding this project. I am grateful for the patience you have shown in me.

Dr Byers for providing me with constant and persistent support and direction; for keeping me on track at all times and for introducing me to the world of genetics in cancer. This thesis would not have been done without your help.

Prof Siriwardena for offering me the opportunity to undertake a research project in pancreatic cancer and helping me develop a set of skills in academic surgery. Also, for your patience and understanding throughout and for supporting my aim to become a specialist registrar.

Roy Goodacre, Andrew Vaughan, Nicholas Rattray, Yun Xu and Will Allwood for the guidance and collaboration with the metabolomic throughput and analysis.

To my fellow researcher, Asela Bandara for the help in collecting samples, sorting out databases and willingness to help.

Steve Richardson, Brenda and the gang, for cups of tea, good chat and general help with the Poly A PCR processing.

Jean and David and Mum and Dad for their encouragement and general nagging without which, I am sure this thesis would undoubtedly have taken twice as long.

Finally, for her unfailing good cheer, motivation and support during this long and arduous project, my best and closest friend, my wife, deserves all the credit for keeping me going.
Presentations arising from this work

   
   **Sanyal S**, Siriwardena A, Byers RJ.

2. *Natural history of small < 3cm cystic lesions of the pancreas.*
   
   Presented at the American Pancreatic Association, Miami 2012.
   
   **Sanyal S**, Siriwardena A

3. *Assessment of the accuracy of EUS, EUS cytology and CT compared with a gold standard in the diagnosis of pancreatic lesions.*
   
   Presentation at the European Pancreas Club, Prague, June 2012.
   
   **Sanyal S**, Puleston J, Rana D, Holbrook M.

4. *A registry for patients with pancreatic cystic tumours: Experience and clinical outcomes.*
   
   Poster presented at the International Association of Pancreatology, Kochi, Feb 2011.
   
   **Sanyal S**, Siriwardena A, Sheen AJ.

5. *Selective non-resectional management of pancreatic cystic tumours: A cohort study based on a cystic tumour registry.*
   
Chapter 1 - An Introduction to the Pathology and Genetics of Pancreas Cancer

1.1. Introduction

Pancreas cancer is the 9th most common cause of cancer in the UK (1). Although only 8000 cases are diagnosed annually, the overall 5-year survival is about 3% and resection is only possible in 5-15% (2, 3). It has been estimated that the lifetime risk of developing pancreatic cancer is 1 in 77 for men and 1 in 79 for women in the UK (3, 4). The relative 1-year and 5-year survival rates for all stages are 16% and 3% respectively (5). The high mortality rate in pancreas cancer is largely due to the gradual onset of symptoms leading to difficulty in identifying the disease at an earlier stage and the aggressive nature of the cancer with early metastatic spread.

Surgery is associated with improved survival with a five-year survival of about 20% (6, 7). Patients with early-stage cancers survive longer; with cumulative 3-year survival rates for patients with stage I or stage II cancers of 50.4% and 45.5% respectively and survival rates for patients with stage III and stage IV cancers of 17.6% and 0-5%, respectively (6). In addition to surgery, chemotherapy has also been shown to improve survival in pancreas cancer. Adjuvant chemotherapy using 5-fluorouracil/ Folinic acid was shown to confer a significant survival benefit in resectable pancreas cancer (ESPAC-1 trial) (8). Since 1997, gemcitabine has been the standard first line treatment for patients with unresectable locally advanced or metastatic pancreas cancer. Gemcitabine has also shown an increase in disease free survival in patients undergoing curative pancreatic cancer resections (9, 10). FOLFIRINOX (a combination of fluorouracil, leucovorin, irinotecan and oxaliplatin) in comparison with gemcitabine alone, showed a survival benefit in patients with metastatic pancreas cancer who are under the age of 75 years and have a good ECOG performance status, despite a higher incidence of side effects (11). The addition of albumin-bound paclitaxel to gemcitabine has also been shown to significantly increase disease free survival as compared to gemcitabine alone but with increased myelosuppression and peripheral neuropathy (12).
Many patients are unsuitable for resection, either because the disease is locally advanced or has metastasized. The median survival of patients with locally advanced or metastatic disease is 6-11 months and 2-6 months respectively (13, 14). Thus, the overall outlook of pancreatic cancer may be improved if patients were diagnosed at an earlier stage or if those patients with lesions known to be pre-malignant could be offered surgery at an earlier stage of the disease. Pathological studies have identified three possible precursor lesions progressing to pancreatic cancer. These include pancreatic intraepithelial neoplasia (PanIN), intraductal papillary mucinous neoplasm (IPMN) and mucinous cystic neoplasm (MCN) (15). In order to detect lesions earlier, research has focused on understanding precursors of pancreas cancer and their molecular changes during progression towards invasive cancer. This has potentially opened up avenues for exploration to facilitate the early detection of molecular changes involved in the progression to invasive disease and the development of modalities blocking the progression to invasive disease. In addition, the analysis of metabolic pathways altered in comparison to normal controls, may help identify metabolic pathways involved in the progression to cancer. This makes metabolites suitable to be studied as possible markers during an altered metabolic state induced by diseases such as cancer.

The introduction will focus on the literature related to genetics of pancreas cancer and its precursor lesions and specifically look at the current molecular changes involved in the progression to cancer.

1.2. Risk Factors

The risk factors associated with pancreatic cancer are smoking (risk ratio of 1.74) (16), increasing age and certain syndromes such as Peutz-Jeghers syndrome (STK11 suppressor gene mutation) (17), autosomal dominant mutations in BRCA2 gene (18) among others. A pooled analysis of 14 prospective cohort studies showed a modest increase in the risk of pancreatic cancer with consumption of 30 or more grams of alcohol per day (19). A recent meta-analysis showed that consumption of three or more drinks of alcohol per day is associated with a 20–30% increased risk of pancreatic cancer (20). Pancreatitis is a well-known risk factor for pancreatic cancer. From a meta-analysis of 6 cohort studies, pooled relative risk estimate for chronic pancreatitis patients to get pancreatic cancer is found to be 13.3. The relative risk is particularly high when pancreatitis
develops at an early age, such as hereditary pancreatitis (RR 69) and tropical pancreatitis, (RR 100), with at least 50-fold greater rates of pancreatic cancer vs. the general population (21). About 5-10% of patients with pancreatic cancer may have a family history of the disease (22). The risk of pancreas cancer is 62-76% higher in people with a first degree relative with the disease (23). The risk is higher in those with a greater number of first-degree relatives affected or those who have the disease at a younger age (24). Radiotherapy for other cancers is also linked to an increased risk of pancreas cancer (25). Other risk factors are detailed below in Table 1 (26).

Table 1: Risk factors in the development of pancreas cancer

<table>
<thead>
<tr>
<th>Risk factors</th>
<th>Modifiable</th>
<th>Non-modifiable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cigarette smoking</td>
<td>Age</td>
<td></td>
</tr>
<tr>
<td>Obesity</td>
<td>Male sex</td>
<td></td>
</tr>
<tr>
<td>High intake of animal fat</td>
<td>First degree relatives with cancer</td>
<td></td>
</tr>
<tr>
<td>Occupational exposure (Nickel, Chloride hydrocarbons from petroleum products and wood pulp)</td>
<td>Chronic pancreatitis (Tropical/familial/chronic)</td>
<td></td>
</tr>
<tr>
<td>Partial gastrectomy</td>
<td>Inherited genetic predisposition</td>
<td></td>
</tr>
<tr>
<td>Diabetes mellitus of recent onset</td>
<td>Non-“O” blood group</td>
<td></td>
</tr>
<tr>
<td>Previous exposure to ionising radiation</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1.3. Precursor Lesions

There are three histologically recognized precursor lesions of pancreatic cancer, namely, pancreatic intraepithelial neoplasia, intraductal papillary mucinous neoplasm and mucinous cystic neoplasm (27-29).

Pancreatic intraepithelial neoplasia are microscopic, flat or papillary, noninvasive epithelial neoplasms arising in the pancreatic ducts, characterized by columnar to cuboidal cells with varying amounts of mucin and degrees of cytological and architectural atypia (30, 31). Higher grades of PanIN have been demonstrated to be associated with progression to invasive adenocarcinoma (15, 32) corroborated by similarities in molecular profiles in higher grade PanLNs (PanIN 3) and invasive adenocarcinoma (33).
A unified nomenclature was developed in 1999 to classify pancreatic intraepithelial neoplasia (PanINs). This is detailed below in Figure 1 (31).

PanIN-1A lesions are flat, tall columnar cells with basally located nuclei

PAN-1B lesions show papillary architecture.

PanIN-2 lesions exhibit nuclear abnormalities, such as a loss of polarity or nuclear crowding; PanIN-3 lesions have marked nuclear and cytological abnormalities (previously referred to as carcinoma in situ of the pancreas).

Invasion through the basement membrane marks the transition from PanIN-3 to invasive carcinoma. PanIN lesions demonstrate a progressive accumulation of genetic aberrations that mirror their histologic progression.

Figure 1: Pancreatic intraepithelial Neoplasia: PanIN-1A (flat), PanIN-1B (papillary), PanIN-2 (papillary with nuclear change), and PanIN-3 (severe atypical cells with mitoses, budding and luminal necrosis). (31)
Intraductal papillary mucinous neoplasms (IPMN) are grossly visible, predominantly papillary (rarely flat), noninvasive, mucin producing epithelial neoplasms arising from the main pancreatic duct or branch ducts, with varying degrees of ductal dilatation. These lesions can be categorized as main duct IPMN, branch duct IPMN or mixed IPMN (with features of both main and branch duct lesions). IPMNs usually produce a lesion greater than 1 cm in diameter, and include a variety of cell types with a spectrum of cytological and architectural atypia (30, 31). Operative specimen photographs of main duct IPMN are shown below Figure 2 and those of branch duct IPMN in Figure 3.

![Dilated main pancreatic duct with mucin within](image)

**Figure 2:** Main duct IPMN demonstrating a pancreatic ductal diameter >1cm with mucin within the duct (34)
Figure 3: Branch duct intraductal papillary mucinous neoplasm demonstrating mucinous cystic lesions in the branch duct without main duct dilatation (34)

The genetic analysis of IPMN demonstrates many of the molecular changes seen in pancreatic ductal adenocarcinoma, although these appear to be slightly different to those seen in PanIN (35-42).

Mucinous cystic neoplasms (MCN) are distinctive mucin producing epithelial neoplasms with a characteristic ovarian stroma, which usually do not communicate with the pancreatic duct (43, 44). The lack of ductal communication and the minimal risk of recurrence after resection combined with the potential for neoplastic transformation makes surgery the treatment of choice for these lesions. The risk of neoplastic transformation is in the range of 7-30% (45-47). A postoperative cut specimen is depicted in to demonstrate the lack of ductal communication in Figure 4.
It has been proposed that MCNs may originate from ectopic ovarian tissue enveloped by pancreatic cells during embryogenesis and potentially facilitate the development of MCNs (43). Global gene expression analysis of MCNs has shown over-expression of several genes involved in estrogen metabolism, including steroidogenic acute regulatory protein (STAR) and estrogen receptor-1 (ESR1) by the ovarian-type stroma (48). Therefore, knowledge of the mechanisms of progression of these precursor lesions to invasive cancer could potentially assist in developing molecular diagnostic tools facilitating an earlier diagnosis and therefore treatment of pancreatic cancer.

1.4. Pathology

The tumours occur more commonly in the head of the gland (60%), followed by the body and tail (15%) and can occur diffusely in 20% (49). Cancers of the pancreatic head often demonstrate invasion and stenosis of the bile duct and pancreatic duct causing a simultaneous ductal dilatation referred to as the ‘double duct’ sign. This double duct dilatation, when seen on ERCP has been reported to predict pancreas or ampullary cancer with a high degree of certainty (50).
Tumours of the body and tail of the pancreas are often larger and demonstrate a greater degree of local invasion. Pancreas cancer is a disease that shows evidence of systemic spread in 80% of cases at the time of diagnosis. The most common sites for distant metastases are the liver (80%), peritoneum (60%), lung and pleura (50–70%), and adrenal glands (25%) (51). Grossly, they are firm, yellow-white masses often demonstrating vascular, neural and lymphatic invasion as seen in Figure 5. Cut sections show hard, grey mass with occasional areas of microcystic change, but areas of necrosis or haemorrhage are uncommon (Figure 6). Histologically, pancreatic cancers are tumours arising from pancreatic ductal epithelia, with production of mucin and expression of a characteristic cytokeratin (CK) pattern. The cytokeratin phenotype of an adenocarcinoma can assist in the identification of its primary site. Pancreas cancer shows strong reactivity for cytokeratin-7 (CK7) while less so for CK17 and CK20 (50% for CK17 and 20-62% for CK20) (52-54).

Figure 5: Gross appearance of pancreas cancer
Microscopically, ductal adenocarcinomas are characterized by low columnar, mucin containing cells in glands infiltrating into the surrounding tissues, associated with a surrounding desmoplastic stromal component (37). Most pancreatic cancers are histologically moderate to well differentiated tumours; however variations in the degree of differentiation can be seen at different foci within the same sample. Well-differentiated carcinomas consist of large duct-like structures, combined with medium-sized neoplastic glands in a tubular or cribriform pattern. Occasional irregular papillary projections may be present. The columnar cells contain pleomorphic nuclei with loss of polarity and prominent nucleoli as seen in Figure 7.
Figure 7: Well-differentiated adenocarcinoma with neoplastic tubular glands lined with cuboidal or cylindrical cells with large, irregular nuclei and clear mucinous cytoplasm (x200). (55)

Moderately differentiated carcinomas predominantly show a mixture of medium sized duct-like and tubular structures of variable shape in a desmoplastic stroma. In comparison to well-differentiated carcinoma, the cells exhibit a greater variation in nuclear size, chromatin structure and prominence of nucleoli with frequent mitosis. Poorly differentiated carcinomas are composed of a mixture of small irregular glands as well as solid tumour cell sheets and nests, replacing acinar tissue. Microscopically, the cells show marked pleomorphism with little or no mucin production and brisk mitotic activity (Figure 8).
Pancreatic cancer cells are associated with the aberrant production of multiple mucins; using the unique property of mucins to interact with the cellular microenvironment, survive and proliferate and to metastasise to different locations (57). MUC1 is a marker of intralobular ductal cells is consistently expressed in pancreatic ductal adenocarcinoma (42). MUC2 on the other hand, is not found in the normal pancreas and is also not present in ductal adenocarcinoma while MUC4 and MUC5AC, which are not found in the normal pancreas, are expressed (58). The WHO classification of pancreatic cancers subtypes is detailed below.
Table 2: Classification of pancreatic neoplasms (49)

<table>
<thead>
<tr>
<th>Exocrine Neoplasms</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Serous Cystic neoplasms</strong></td>
</tr>
<tr>
<td>Microcystic Serous cystadenoma</td>
</tr>
<tr>
<td>Macrocystic serous cystadenoma</td>
</tr>
<tr>
<td>Solid serous cystadenoma</td>
</tr>
<tr>
<td>Von Hippel Lindau-associated serous cystic neoplasms</td>
</tr>
<tr>
<td>Serous cystadenocarcinoma</td>
</tr>
<tr>
<td><strong>Mucinous cystic neoplasms</strong></td>
</tr>
<tr>
<td>With low, moderate or high grade dysplasias</td>
</tr>
<tr>
<td>With associated invasive adenocarcinomas</td>
</tr>
<tr>
<td><strong>Intraductal neoplasms</strong></td>
</tr>
<tr>
<td>Intraductal papillary mucinous neoplasm</td>
</tr>
<tr>
<td>- With moderate dysplasia</td>
</tr>
<tr>
<td>- With associated invasive adenocarcinoma</td>
</tr>
<tr>
<td>Intraductal oncocytic papillary neoplasm</td>
</tr>
<tr>
<td>Intraductal tubular neoplasm</td>
</tr>
<tr>
<td>- With low grade or high grade dysplasia</td>
</tr>
<tr>
<td>- With associated invasive adenocarcinoma</td>
</tr>
<tr>
<td><strong>Pancreatic Intraepithelial Neoplasia (PanIN)</strong></td>
</tr>
<tr>
<td>PanIN-1A, PanIN-1B, PanIN-2 and PanIN-3</td>
</tr>
<tr>
<td><strong>Invasive Ductal adenocarcinoma</strong></td>
</tr>
<tr>
<td>Tubular adenocarcinoma</td>
</tr>
<tr>
<td>Adenosquamous carcinoma</td>
</tr>
<tr>
<td>Signet ring cell carcinoma</td>
</tr>
<tr>
<td>Colloid adenocarcinoma</td>
</tr>
<tr>
<td>Hepatoid carcinoma</td>
</tr>
<tr>
<td>Medullary carcinoma</td>
</tr>
<tr>
<td>Undifferentiated (anaplastic) carcinoma</td>
</tr>
<tr>
<td>Undifferentiated carcinoma with osteoclast like giant cells</td>
</tr>
<tr>
<td><strong>Acinar Cell Neoplasms</strong></td>
</tr>
<tr>
<td>Acinar cell cystadenoma</td>
</tr>
<tr>
<td>Cell carcinoma</td>
</tr>
<tr>
<td>Acinar cell cystadenocarcinoma</td>
</tr>
</tbody>
</table>
Endocrine Neoplasms

<table>
<thead>
<tr>
<th>Microadenomas</th>
</tr>
</thead>
<tbody>
<tr>
<td>Well differentiated pancreatic neoplasm</td>
</tr>
<tr>
<td>Poorly differentiated endocrine carcinoma (small and large cell carcinomas)</td>
</tr>
</tbody>
</table>

Epithelial neoplasms with multiple directions of differentiation
- Mixed acinar-endocrine carcinoma
- Mixed acinar-ductal carcinoma
- Mixed ductal-endocrine carcinoma
- Mixed acinar-ductal-endocrine carcinoma
- Pancreatoblastoma

Epithelial neoplasms of uncertain direction of differentiation
- Solid pseudo-papillary neoplasm

Miscellaneous epithelial neoplasms
- Teratoma
- Lymphoepithelial cyst
- Epidermoid cyst in intra-pancreatic heterotopic spleen

1.5. Genetic origins of pancreas cancer

A series of consistent genetic changes occurs in the evolution of pancreatic cancer (59, 60). In addition, support is provided by progressive changes seen in precursor lesions that are consistent with those seen in progression to adenocarcinoma (61), while in familial pancreatic adenocarcinoma, the genetic changes have been well characterized (62). Lastly, genetically engineered mouse models with inactivated oncogenes replicate the range of changes occurring in pancreatic adenocarcinoma (63, 64).

The cell of origin of pancreas cancer is still unknown. Emerging evidence using a xenograft model implanted in immunocompromised mice with pancreatic adenocarcinoma suggests that pancreatic cancer may arise from a cancer stem cell characterized by a CD44, CD 24+ ESA+ phenotype (65). It is to be emphasized that this study, though encouraging in identifying highly neoplastic cells with the ability to renew and propagate, was primarily experimental and had several limitations when applied to naturally evolving human pancreas adenocarcinoma. Limitations of the study include the specific phenotype of the candidate stem cells chosen, the type and location of implantation into the model and whether the stem cells in vivo would more
accurately be represented by a mutated stem cell or a de-differentiated cell that has regained stem cell like properties. No clear candidate pancreatic stem cell has emerged despite extensive attempts at isolating one. The two potential candidate cells include either the centroacinar cells, the only cells that retain Notch activation or differentiated acinar cells that harbor facultative progenitor properties (66-69). Activation of the notch-signaling pathway implies the persistence of a precursor like transcriptional program and represents de-differentiation in the acinar cells. The subsequent acquisition of oncogenic mutations permits independent expansion of these deregulated stem/progenitor cells, culminating in neoplasia (70). Though the centroacinar cell is an attractive candidate, it is unclear whether this cell is a precursor for either ductal or acinar cell adenocarcinoma. Also, uncertainty exists regarding the role of these cells in the precursor lesions seen in the progression to cancer.

1.6. Molecular genetics of pancreas cancer

Pancreas cancer is a disease of inherited and somatic mutations. The progression from normal ductal epithelium through to PanIN to adenocarcinoma is mediated through a series of progressive, non-random, cascading genetic signals producing alterations in cellular morphology. The genetic abnormalities include alterations in chromosome or gene copy number, microsatellite instability, epigenetic silencing, intragenic point mutations, and gene overexpression secondary to increased transcription (71).

The progressive sequence of genetic alterations in pancreatic ductal cells associated with progression from PanIN-1A through PanIN-3 advancing through to adenocarcinoma is detailed in Figure 9 below (32).
1.6.1. Genome wide studies

Next generation, high throughput technologies have revolutionized genome-wide studies of pancreas cancer (73-76). These large volume sequencing studies have underlined genetic changes ranging from base pair alterations (substitutions, insertions or deletions) involving single or multiple base pairs; single chromosome alterations (deletions, inversions, amplifications or fold-back inversions) to large chromosomal re-arrangements (translocation or insertion between two non-homologous chromosomes) (77). A comprehensive genetic analysis of 24 pancreatic cancers found that pancreatic cancers contain an average of 63 genetic alterations, the majority of which were point mutations. These alterations defined a core set of 12 cellular signaling pathways and processes that were each genetically altered in 67 to 100% of the tumors (74). The core signaling pathways including the principal genes involved in the progression to invasive pancreatic ductal adenocarcinomas are depicted below in Figure 10 (71).
A literature search aiming to catalogue molecules with evidence of overexpression at mRNA level, protein level, or both identified a total of 2516 genes involved in pancreatic cancer. For selection into the compendium, there had to be a minimum of a 2-fold over-expression of mRNA in pancreas cancer compared with normal samples. Molecules with a fold change of less than two were included only if they were detected by multiple methods. Of these, 1868 genes were reported as overexpressed only in the mRNA analyses, while 441 were reported as overexpressed both at mRNA and protein levels, and 207 molecules were reported as overexpressed only at the protein level. The authors reported that more than 200 genes were overexpressed in pancreatic cancers in four or more studies and identified these molecules as candidates for focused validation (78). The molecules selected for focused validation were detected by multiple methods at both mRNA and protein levels. All the selected molecules were membrane associated or secreted molecules and had immunohistochemical evidence, thereby increasing their likelihood as potential biomarkers.
More recently, Biankin et al performed exome sequencing and copy number analysis to define genomic aberrations in ‘early stage’ pancreatic adenocarcinoma. Early stage pancreatic adenocarcinoma was defined as pre-operative clinical stage I and stage II cancers. Analysis identified 2,016 non-silent mutations and 1,628 copy-number variations. The study confirmed known mutations (KRas, p53, CDKN2A, Smad4, MLL3, TGFBR2, ARID1A and SF3B1), and also identified additional mutated genes including those involved in chromatin modification (EPC1 and ARID2), DNA damage repair (ATM) and other mechanisms (ZIM2, MAP2K4, NALCN, SLC16A4 and MAGEA6), also identified in a previous study (79). They also identified frequent somatic aberrations in genes described traditionally as embryonic regulators of axon guidance, particularly SLIT/ROBO signaling (75). This can be added to the principal group of signaling pathways responsible for the development of pancreatic cancer as is shown in Figure 11. (71)

![Figure 11: Core signalling pathways incorporating the addition of Axon guidance and chromatin regulation in the development of pancreatic ductal adenocarcinoma (adapted from Yachida et al, 2013)](image)

1.6.2. Signalling pathways in the development of pancreatic cancer

Alterations in a few of the core signal transduction pathways involved in pancreatic carcinogenesis are detailed below (80, 81). These pathways represent the most commonly altered pathways in the evolution of pancreas cancer and have been shown to characterize the molecular changes in both pancreas cancer and its precursor lesions.
1.6.2.1. The Ras, MAP2K, and MEK pathway

i) Kras

K-Ras encodes a set of membrane-bound GTP-binding proteins that is activated by signaling molecules, such as epidermal growth factor receptor (EGFR). Mutated K-Ras acts to continuously activate the Ras protein, producing a permanently activated series of downstream signalling cascades (including the Ras pathway, the mitogen-activated protein kinase and PIK3A-Akt signals). These then proceed to trigger processes including transcription, translation, cell cycle progression, cell survival, and motility (82, 83) (Figure 12).

Figure 12: Intracellular signalling pathways involved in pancreas cancer (84)

(Abbreviations: EGFR = epidermal growth factor receptor; ERK = extracellular signal-regulated kinases; FGF-R = fibroblast growth factor receptor; GDP = guanosine diphosphate; GTP = guanosine triphosphate; HER2 = human EGFR 2; IGF = insulin growth factor; MEK = mitogen-activated protein kinase; MEKK = mitogen-activated protein kinase kinase; MMP = matrix metalloproteinase; mTOR = mammalian target of rapamycin; NF-kappaB = nuclear factor kappa-light-chain-enhancer of activated B cells; VEGF = vascular endothelial growth factor receptor)

ii) EGFR, VEGFR, and IGFR-1 pathway (85, 86)

EGFR and vascular endothelial growth factor receptor (VEGFR) are tyrosine kinase receptors that are stimulated on binding EGF and VEGF. They promote cell proliferation, survival, and angiogenesis. EGFR can be triggered by a variety of genetic alterations including mutations,
overexpression of receptor ligands, or loss of inhibitory regulation producing downstream signaling events through Ras/Raf/MEK/MAPK, PI3K–AKT, and the STAT family of proteins (86). IGFR promotes survival and antiapoptotic effects in tumor cells through both K-Ras-dependent and independent (PI3K–Akt, MAPK, and STAT) downstream signaling cascades (85, 87). VEGFR promotes neo-angiogenesis within the pancreatic cancer and is directly associated with proliferation (88, 89).

iii) PI3K/Akt signaling pathway

The phosphatidylinositol 3-kinase/Akt signalling pathway is involved in the proliferation, angiogenesis and resistance to apoptosis of pancreatic cancer cells. EGFR/IGF1R activates PI3K, which in turn activates Akt, which induces downstream targets including mTOR and the transcription factor NFκB (90). Loss of the regulatory protein that normally inactivates PI3K, phosphatase and tensin homolog (PTEN) is also seen (91).

1.6.2.2. Pathways involved in the Peri-tumoural Stromal reaction

The pathways mainly involved in stromal reactions include the TGFβ and Hedgehog signaling pathways.

i) Transforming growth factor β (TGFβ)

Pancreatic cancer is associated with a significant peri-tumoural stromal reaction with fibroblast proliferation (49). Evidence has shown that pancreatic cancer cells are actively involved in directing the desmoplastic response and this is driven by TGFβ (92). TGFβ is a cytokine secreted by epithelial, endothelial and mesenchymal cells. By binding to its receptor it prompts the phosphorylation of Smad2 and Smad3, thereby activating gene transcription (93). TGFβ is known to promote a microenvironment that plays a central role in stroma production, angiogenesis, and tumor-induced immunosuppression (94).

ii) Sonic hedgehog pathway (SHH)

Another pathway that contributes to the host-stromal response is the sonic hedgehog pathway. SHH is expressed in about 70% of human pancreatic cancers and influences differentiation and
motility of human pancreatic stellate cells and fibroblasts by mediating cell cycle changes and inducing neo-vascularization (95, 96).

1.6.2.3 Embryonic Signalling pathways

Interactions between embryonic signaling pathways are significant in the differentiation and development from embryo to adult. There is evidence to intimate that these pathways may remain operational within a subset of adult organs and that deregulation of their activity contributes to the development and progression of certain tumours (97, 98). In addition to the Hedgehog and TGFβ discussed above, Notch and Wntβ-catenin are additional pathways involved in the pathogenesis of pancreatic cancer.

a) The Notch genes encode for a set of transmembrane receptors important in the development of tissue development, proliferation, differentiation and apoptosis (99). Notch signalling occurs downstream of Ras, EGFR and TGFβ signalling in pancreatic carcinogenesis and promotes tumour vascularization (100).

b) Wnt signaling is involved in proliferation and prevention of apoptosis and as such is part of normal embryonic development and homeostatic self-renewal (101). The hedgehog and Smad4 pathways have been demonstrated to regulate this pathway (102).

1.6.3. Chromosomal Alterations

Chromosomal alterations have been identified in most types of pancreatic cancer. Techniques such as karyotyping, comparative genomic hybridization and allelotyping are used to map out chromosomal abnormalities. Studies have shown multiple, non-random chromosomal alterations in most cases. Chromosome losses are more common than chromosomal gains in pancreatic cancer. The most common regions of genomic loss in primary pancreatic cancers and pancreatic cancer xenografts include chromosome arms 9p, 17p, 18q, 3p, 8p, and 6q (103-105). Frequent gains of DNA, secondary to unbalanced chromosome rearrangements have been observed in primary pancreatic cancers on chromosome arms 3q, 5p, 7p, 8q, and 20q (106). Various studies have identified gains in genetic copy number at the 18q11.2 locus, which contains GATA6, which contributes progression through PanIN to pancreatic cancer (107, 108). Many regions of
chromosomal depletion also contain tumor suppressor genes, for example, locus 18q21 contains the Smad4/dpc4 gene and p53 (locus 17p13) (109).

### 1.6.4. Microsatellite instability

Although chromosomal abnormalities are common, other genetic abnormalities include DNA mismatch and resultant microsatellite instability (MSI) (110). MSI is characterized by hypermutability at minisatellite repeat sequences and is associated with poor differentiation, lack of \( KRAS2 \) and \( p53 \) mutations, and inherited tumors associated with the human non-polyposis colorectal cancer (HNPCC) syndrome (111, 112). MSI cancers with these associations have a specific histologic pattern and have been termed medullary cancers (113).

The analysis of chromosomal alterations has led to understanding of the impact of tumor suppressor genes and oncogene activation.

### 1.6.5. Specific gene alterations

**Oncogenes**

The accumulation of somatic mutations, in the absence of a familial or hereditary tendency, drives the proliferation and growth of pancreatic cancer. These mutations can affect genes that ‘drive’ the progression of cancer as opposed to ‘passenger’ mutations, which occur at fragile sites in the genome, but do not drive growth or metastasis. Based on the frequency of genetic changes in pancreas cancer, a genetic ‘topographic’ map can be generated measuring the four commonest gene changes as ‘mountains’ (incorporating the gene changes in K-ras, CDKN2A/p16, Smad/DPC4 and p53) and the multitude of various other changes as ‘hills’ or otherwise lower frequency driver genes (32, 114).

A comprehensive genetic analysis of 24 pancreatic cancers identified an average of 63 genetic alterations, the majority of which were point mutations. The authors investigated the genes to identify specific pathways involved and identified 12 core processes that were altered in 63-100% of the cancers studied (74). The pathways included genes in which genes were altered in numbers varying from single gene alterations (K-ras), or several genes (TGF-beta signalling...
pathway) to multiple genes (the GTPase signalling pathway, integrin signalling) were involved (84).

i) K-Ras

Activating mutations of K-ras are the single most common mutation seen in pancreas cancer (115). K-ras encodes a member of the RAS family of guanosine triphosphate (GTP)-binding proteins that mediate a wide range of cellular functions including proliferation, cell survival and cytoskeletal remodeling. Activating mutations impair the intrinsic GTPase activity of the KRAS2 gene product; this results in a protein that is constitutively active in intracellular signal transduction (63, 82). The high frequency of K-ras mutations supports its inclusion as an initiator of pancreatic cancer. In addition, it appears to have a role in cancer maintenance as well (83, 116).

ii) Additional candidate oncogenes

Additional oncogene candidates have been identified. These include among others, the gene encoding the oncogenic transcription factor C-Myc, MYB (chromosome 6q), AIB1/NCOA3 (chromosome 20q), and EGFR (chromosome 7p) (114)

Tumor Suppressor genes

i) CDKN2A/p16

p16/CDKN2A, also known as INK4A, is the most commonly inactivated tumor suppressor gene in pancreatic cancers (117). This loss may occur by a variety of mechanisms such as homozygous deletion (40%), intragenic mutation with loss of the second allele (40%) or epigenetic silencing of gene expression by promoter methylation (15%) (118). Its functions by inhibiting cell cycle progression through the G1eS checkpoint that is mediated by CDKs such as CDK4 and CDK6 (118).
ii) p53

Inactivation of the p53 gene on chromosome 17p is present in approximately 50-75% of pancreatic cancers (119). Its cellular functions include the regulation of the G1-S cell cycle checkpoint; maintenance of G2-M arrest and the induction of apoptosis. Loss of p53 leads to ongoing cellular function in the presence of DNA damage and results in the accrual of persistent genetic defects.

iii) Smad/DPC4

Smad/DPC4 is inactivated in approximately 55% of pancreatic cancers, either by homozygous deletion (30%) or by intragenic mutations and loss of the second allele (25%) (94). The Smad4 protein plays a critical role in signaling through the transforming growth factor- beta pathway. The TGF- beta pathway has profound growth-inhibitory effects by regulating the expression of specific target genes; therefore, loss of Smad4 in pancreatic cancer cells, provides a selective growth advantage (120, 121).

Additional tumour suppressor genes

Several tumor suppressor genes are inactivated in smaller proportions (5%-10%) of pancreatic cancers. These include:

a) LKB1/STK11 (chromosome 19p) (17)

b) TGF-beta/activin signaling pathway receptors such as TGFBR1 (122)

c) The gene MKK4 on chromosome 17p (encodes for a stress-activated protein kinase) is inactivated in some pancreatic cancer metastases (123).

Genome maintenance and caretaker genes

These include genes that are involved in maintaining fidelity of DNA structure. They are seen to be modified in familial pancreatic cancer but are also seen as somatic mutations in non-hereditary
pancreas cancer. These genes may include hMLH1 and hMSH2 (DNA repair genes) and the BRCA2 genes/ FANC genes (Fanconi’s anaemia related genes) (110, 124).

1.6.6. Telomere length abnormalities

Telomeres are hexameric TTAGGG repeats at the ends of chromosome arms that confer stability to chromosomes during cell division and prevent the ends from becoming sticky. They are progressively lost during each round of cell division (125). Telomere lengthening is mediated by telomerase and critically short telomeres mediate a DNA damage response. If the telomeres are abnormally shortened, the initiation of continuous cycles of breakage-fusion-bridge that forms between chromosomes, results in regions of amplifications and deletions within the daughter cell genome (126). In addition, it has been shown that telomeres are abnormally shortened in all stages of PanIN compared to normal ductal epithelial cells, suggestive of the fact that this appears to be one of the primary molecular changes observed in the progression to cancer (127).

1.6.7. Epigenetic Silencing

Inactivation of tumor-suppressor genes occurs not only through homozygous deletion or single allelic deletion with a concomitant mutation of the second allele, but may also occur via heritable DNA modifications that do not involve alterations in DNA sequence referred to as epigenetic silencing (128). The main epigenetic mechanisms that may affect gene expression include DNA methylation, histone modification and microRNA expression.

i) DNA methylation

**DNA methylation** is the covalent binding of a methyl group to the 5-carbon of cytosine residues mediated by DNA methyltransferases (DNMTs). Approximately 80% of pancreatic cancers overexpress Dnmt1 protein (129). CpG islands are stretches of DNA with a high CG nucleotide content and are frequently located near the transcriptional start sites of genes (130). Aberrant hypermethylation of promoter CpG islands is tightly associated with gene silencing and may be associated with loss of tumor suppressor function in cancer (131). Genes that are silenced in pancreatic cancer by epigenetic changes include PENK, RARB, CDKN2A/P16, CACNA1G, TIMP3, THBS1 and hMLH1 (132, 133).
ii) DNA hypomethylation

*DNA hypomethylation* occurs at the 5’ regions of certain genes in pancreatic cancer and is associated with overexpression of the encoded protein. Thus, hypomethylation can result in loss of regulation and the promotion of gene and protein expression, in exact converse to those changes induced by DNA hypermethylation (134).

iii) MicroRNAs

**MicroRNAs (miRNA)** are small, non-protein coding RNA molecules (18 to 24 nucleotides) that regulate transcription of target messenger RNAs and are involved in the regulation of cellular differentiation, proliferation, and apoptosis (135, 136). The expression of miRNAs may be altered through epigenetic, genetic or transcriptional regulation (137-139). Pancreatic ductal adenocarcinomas have been shown to aberrantly express numerous miRNAs (140, 141).

### 1.6.8. Gene over-expression

**i) Mechanism of gene over-expression**

In addition to other genetic changes, specific genes may be over-expressed or silenced in pancreatic cancer. Altered gene expression in tumors can result from chromosomal rearrangements, epigenetic silencing, or mutations in genes that are upstream of a target gene. The use of global gene expression technology has identified a multitude of overexpressed genes in pancreatic ductal adenocarcinoma (59, 142).

**ii) Genes over-expressed in pancreatic cancer**

These include growth hormones and their receptors such as increased levels of fibroblast growth factor, insulin-like growth factor-1 (IGF-1), IGF-1 receptor, nerve growth factor, and vascular endothelial growth factor associated with increased tumor invasiveness (143-145). Tyrosine-kinase growth factor receptors and ligands are also up regulated in pancreatic cancer cells; for example, the gene *HER2/neu* encodes for a glycoprotein with tyrosine kinase activity (146).
iii) Gene expression in the early detection of pancreas cancer:

These overexpressed genes may serve as useful markers for screening if their gene products are released into the blood or pancreatic juice, or they may serve as targets for the development of new pancreatic cancer therapies. In addition, by studying the variations in gene expression in the surrounding stroma, it has been shown that there are regional differences in gene expression within the host desmoplastic response (147). This study characterized the expression of 12 genes identified in invasive pancreatic cancer tissues but not in pancreatic cancer cell lines. In situ hybridization demonstrated that eight genes were expressed within the stromal and/or angioendothelial cells involved in the desmoplastic response of the invasive tumour while the stromal cells specifically expressed four genes immediately adjacent to the invasive neoplastic cells. Thus, this regional variation in areas of the host desmoplastic response seems to depend on proximity of the stromal cells to the primary neoplastic epithelium.

Currently, the most common marker for pancreatic cancer is Ca19-9. Though its sensitivity in symptomatic patients is about 80%, this value is considerably diminished (~55%) in small, resectable cancers (<3 cm) (148). Global gene analysis has revealed a wide variety of possible target over-expressed genes that have potential as serum markers. These include among others, macrophage inhibitory cytokine 1 (MIC1) which is differentially overexpressed in invasive pancreatic cancers (149) and performs as well as Ca19–9 and other tumor markers in distinguishing patients with resectable peri-ampullary cancers from benign pancreatic lesions. The combination of MIC-1 and Ca19-9 yielded a greater accuracy than Ca19-9 alone in differentiating pancreas cancer form healthy controls and chronic pancreatitis (150-152). A recent meta-analysis of studies using MIC-1 as a biomarker suggested that it might have a high sensitivity and specificity in confirming the diagnosis of pancreas cancer (153). Despite this, MIC-1 is not routinely used as a clinical biomarker for screening as no study has shown that it facilitates the early detection of pancreas cancer. Also, MIC-1 is also raised in several other cancer types and there is no agreement on the cut-off value to diagnose pancreas cancer between groups. At this time, MIC-1 has not been used as a validated marker for the diagnosis of early pancreas cancer. A cohort of genes over-expressed in pancreatic cancer including neutrophil gelatinase associated lipocalin (NGAL) (152), osteopontin (154), tissue inhibitor of metalloproteinase-1 (155) and mesothelin genes (156, 157) have all been explored as potential
biomarkers for pancreas cancer; however, none of these have been shown to facilitate an early diagnosis of the disease and await further validation studies prior to being judged to be suitable for clinical use.

**Conclusion**

The progress in the understanding of the genetic biology of pancreatic cancer can help us to recognize patterns of metastasis and in the development of markers of early pancreas cancer. In autopsy studies from 7 patients with metastatic pancreas cancer, data generated from the sequencing of the genomes in cell lines from cancer metastases were used to investigate the clonal relationships between primary and metastatic cancers. This interesting study showed that clonal populations that give rise to distant metastases are embodied within the primary carcinoma, but originate from a non-metastatic clone. Using mathematical modeling, based on the number of cell divisions in the single lineage between tumour initiation and birth of the founder cell of the parental clone during which passenger mutations accumulated, the timing of the genetic evolution of pancreatic cancer was calculated as approximately 10 years from the instigating mutation and the development of the initiator cancer cell with an additional five years required for the ability to metastasize leading to disseminated spread and death approximately 2 years later (158). Currently, most patients are diagnoses in the period within 2 years of irredeemable spread and death (159). These insights demonstrate that advances in the molecular biology of pancreas cancer could potentially alter the clinical course of the disease by facilitating an earlier diagnosis if techniques adapted from this understanding of pancreatic cancer biology can be transmitted from the laboratory to the bedside.

**Overall Objectives**

There has been a considerable expansion of knowledge in the development of pancreatic cancer. Despite this, there has been little progress made in facilitating an early diagnosis of this disease and predicting an accurate response to treatment. We aim to translate some of the understanding we have gleaned in the molecular biology of pancreas cancer with a view to facilitate the early diagnosis of pancreas cancer in clinical practice. By studying the natural history and progression of precursor lesions we aim to identify patients who are at increased risk of developing pancreatic
cancer and clarify the incidence of adenocarcinoma in this group. By analyzing over-expressed genes in pancreatic cancer and its precursor lesions we hope to translate the advances in molecular biology to achieve an early diagnosis. Finally, we will also study the metabolic derangements that occur downstream of these changes at the genetic level.

This thesis has three discrete sections:

1. The first section investigates the natural history of pancreatic cystic neoplasms, specifically IPMNs, using a prospectively collected database. This study evaluates the diagnostic pathways, clinical profiles and outcomes of these precursor lesions. In particular, we investigate main duct IPMN and branch duct IPMN's and also focus on cystic lesions measuring less than 3 cm in size, as these lesions pose a dilemma in terms of their surgical management.

2. The second section aims to assess the clinical utility of over-expressed ‘indicator’ genes from matched pancreatic juice and pancreatic tumour tissue samples using Poly A PCR, obtained from patients undergoing surgery for pancreatic cystic neoplasms and pancreatic cancer. This is based on the null hypothesis that there is no relationship between genes over-expressed in pancreatic juice and the same genes in pancreatic tumour tissue.

3. The third study will review the current literature on the use of metabolomics as a tool in identifying biomarkers in pancreas cancer. We will evaluate the potential use of gas chromatography mass spectrometry in a cohort of patients with pancreas cancer, ampullary cancer and endocrine cancer. This study will seek to identify metabolites differentiating these conditions based on the null hypothesis that there is no difference in the metabolite profiles between the three conditions.
Chapter 2 - A study of the natural history of pancreatic cystic neoplasms

Introduction

Pancreatic cystic tumours comprise of a subset of pancreatic lesions incorporating intraductal papillary mucinous neoplasms, mucinous cystic neoplasms and serous cystadenomas. They are increasingly reported and studied due to their neoplastic potential. As precursors of pancreatic cancer, IPMN and MCN’s facilitate a valuable understanding towards the development of pancreatic cancer while SCA’s are most often rare and benign lesions that are most often managed conservatively.

Aims

We aim to assess the natural history of these lesions by means of a prospectively collected database. A greater insight into the progression of these lesions shall enable the early management of those patients with an increased likelihood of developing cancer. In addition, it will spare patients from undergoing major surgery with considerable morbidity and mortality, when the risk is substantially lower. This chapter will review the basic pathology and molecular development of pancreatic cystic tumours with special emphasis on IPMN’s. We will also prospectively study this cohort of patients evaluating the clinical profile, investigation and management strategies involved. We will focus attention on a group of patients with cystic tumours less than 3 cm in size who pose a dilemma in terms of their surgical management.

2.1. Pancreatic cystic tumours

Intraductal papillary mucinous neoplasms are grossly visible, predominantly papillary (rarely flat), noninvasive, mucin producing epithelial neoplasms arising from the main pancreatic duct or branch ducts, with varying degrees of ductal dilatation. IPMNs usually produce a lesion greater than 1 cm in diameter, and include a variety of cell types with a spectrum of cytological and architectural atypia (27, 30). Genetic analysis of IPMN demonstrates many of the molecular changes seen in pancreatic ductal adenocarcinoma, although these appear to be slightly different
to those seen in PanIN (36-42, 49). PanIN’s and IPMN share many features (42, 160). Both exhibit columnar, mucin producing epithelial cells that may arrange in an intraductal papillary pattern with varying degrees of atypia that progress to cancer in a series of progressive genetic changes. The papillary pattern in IPMN tends to be florid and produces large quantities of mucin while PanIN’s tend to be flat with small or micropapillary with small amounts of mucin if any (161). IPMN generally tends to involve the larger ducts while PanIN generally involves the smaller ducts less than 5mm in diameter. IPMN’s are macroscopically visible on radiological scans and grow in size while remaining restricted to the ducts while PanIN's are microscopic lesions that are detected on histology. However, genetically, smad4/ Dpc4 is less frequently inactivated in IPMN than in PanIN (39). Some IPMN also express MUC2 and this is not seen in PanIN, the cancers that develop in these MUC2 positive IPMN also tend to be the more favourable colloid carcinomas as opposed to ductal adenocarcinomas (42).

Mucinous cystic neoplasms are distinctive mucin producing epithelial neoplasms with a characteristic ovarian stroma, which usually do not communicate with the pancreatic duct (43, 49). It has been proposed that MCNs may originate from ectopic ovarian tissue enveloped by pancreatic cells during embryogenesis and potentially facilitate the development of MCNs (43) (Figure 13).

![Ovarian type stroma](image)

**Figure 13: Mucinous cystic neoplasm with ovarian type stroma (x200) (162)**
2.2. Intraductal Papillary Mucinous Neoplasms

The first series of patients with intraductal pancreatic papillary lesions with prominent duct dilatation and grossly visible mucin was reported in 1982 (163). In 1994, Sessa introduced the term intraductal papillary mucinous neoplasms for these lesions and this has remained the standard terminology (164).

2.3. Demographics

Intraductal papillary mucinous neoplasms comprise of less than 5% of exocrine pancreatic tumors though they are being recognized with increasing frequency. In one population based study the age- and sex-adjusted cumulative incidence for IPMN was 2.04 cases per 100,000 persons over a period from 1984 to 2005. On restricting this to residents over the age of 60 years, the point prevalence was 99.10 cases per 100,000 persons or one per 1009 persons (165). In one study, retrospective analysis of abdominal magnetic resonance scans over a year demonstrated incidental pancreatic cysts in 13.5% of 616 patients (166).

2.4. Clinical features

Intraductal papillary mucinous neoplasms occur more commonly in men than in women (3:2) and mean age at diagnosis is 60-65 years (167, 168). Presenting symptoms include abdominal pain, pancreatic insufficiency and new onset or worsening diabetes mellitus. These symptoms may be associated with obstructive pancreatitis caused by the clogging of the ducts with thick mucus. Jaundice is uncommon and is often associated with invasive IPMN. Up to a third of patients are asymptomatic at presentation and are found to have pancreatic cystic lesions detected on imaging studies (169). Seventy five percent of lesions arise in the head of the gland, followed by the body and tail respectively (169, 170). Thirty percent of lesions may also be multifocal within the pancreas and missed on incomplete histological sampling (160). Ten percent may also diffusely involve the whole gland (171). In addition, thirty percent of patients with IPMN may present with neoplasia elsewhere in the body (172). This study evaluated the presence of extrapancreatic neoplasms pre and post operatively in a group of 42 patients with benign and malignant IPMN. The extra-pancreatic neoplasia included colorectal adenomas in 21%, 12%
were adenocarcinomas and 10% were gastric carcinomas. However, this finding has not been corroborated by a multicenter observational study conducted to assess the prevalence of extra-pancreatic neoplasms following the diagnosis of IPMN and excluding neoplasms diagnosed prior to or synchronous with the IPMN. The authors did not demonstrate a difference in the incidence of extra-pancreatic neoplasms in patients with IPMN as compared to the general population following the diagnosis (173).

2.5. Classification

Anatomically, IPMNs are sub-classified depending on the calibre of the involved ducts into main duct, branch duct and a mixed duct type (Figure 14). In one study identifying predictors of progression to pancreas cancer, univariate analysis showed that presence of symptoms, a main pancreatic duct >15 mm and mural nodules were all significant predictors of malignancy in main or mixed duct IPMN. On multivariate analysis in branch duct IPMNs, the authors found that size >30mm and the presence of mural nodules were the strongest predictors of malignancy (174). Invasive adenocarcinoma is found in 20-50% of IPMNs (175). Main duct and branch duct IPMN have significant differences in prevalence of cancer ranging from 57-92% and 6-42% respectively (176-182).

---

Figure 14: Types of pancreatic cystic tumours (183)
2.6. Pathology

2.6.1. Morphology

Histologically, IPMNs are made up of tall columnar, mucin-producing, papillary-growing neoplastic cells on the inner surface of the ducts. Features supportive of the diagnosis of IPMN include macroscopically visible papillae, visible luminal mucin and tall papillae with stromal cores (Figure 15) (184). The walls of the involved ducts are usually fibrotic, pauci-cellular and devoid of acinar parenchyma (185). This is seen in Figure 16 and Figure 17. (186)

Figure 15: Main duct IPMN - showing a dilated duct with neoplastic proliferation of mucinous epithelium. The papillae are clearly identified (184).
Figure 16: IPMN demonstrating a transition zone between low-grade to high-grade dysplasia. (186)

Figure 17: The development of mucinous adenocarcinoma with associated desmoplasia. (186)
2.6.2. IPMN versus PanIN

As a type of intraductal neoplasia, IPMN is also somewhat similar to pancreatic intraepithelial neoplasia (PanIN). PanIN are microscopic (i.e. undetectable by routine radiological procedures) and non invasive pancreatic ductal epithelial neoplasms that usually involve ducts < 5mm in diameter and may be papillary or flat with varying amounts of mucin production. However, the papillae in PanINs are not as tall or complex as in IPMNs; they develop in branched, complex ducts and often low grade PanIN often surrounds higher grade lesions (i.e. they are histologically complex) (187) (Figure 18).

![Loss of polarity and nuclear crowding in PanIN2](image)

Figure 18: Pancreatic intraepithelial Neoplasia (PanIN-2) (187).

2.6.3. Architectural Detail

On the basis of predominant cellular architecture and the pattern cell differentiation, IPMN's can be subdivided into four distinctive microscopic subtypes (185). These include, the gastric foveolar type, the intestinal type (IN), pancreatobiliary (PB) and the oncocytic type (185, 188, 189). Approximately 50% of IPMNs are of IN morphology, 26% of gastric type, 13% oncocytic and 7% PB subtype. Data suggests that the PB type papilla is more prone to be associated with tubular type invasion while the IN type with colloid type invasion (44, 176, 188).
2.7. Molecular pathology

The molecular events involved in the pathogenesis of IPMN are being increasingly studied, in order to understand the genetic changes occurring during the progression to pancreas cancer. IPMNs share many of the molecular characteristics seen in PanIN’s and pancreas cancer to a lesser extent. The abnormalities in molecular pathways include changes in activation of oncogenic pathways, inactivation of tumour suppressor genes, deregulation of STK11/LKB1 and DUSP6, loss of heterozygosity of chromosomes and alterations in DNA methylation. The various differences in the genetic changes between pancreatic intraepithelial neoplasms, intraductal papillary mucinous neoplasms and pancreas cancer are summarized in the table below. (Table 3).

Figure 19: The four histological subtypes of IPMN (189)
### Table 3: Differences in molecular characteristics of the precursor lesions of pancreas cancer (Adapted from Cooper et al.) (190)

<table>
<thead>
<tr>
<th>Lesion</th>
<th>PanIN</th>
<th>IPMN</th>
<th>Pancreas cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td>KRas</td>
<td>Mutations in KRas in 44% of PanIN-2 and 87% of PanIN3 (191)</td>
<td>KRas codon 12 mutations seen in 30-80% of IPMN (164)</td>
<td>95% mutated in pancreas cancer (192, 193)</td>
</tr>
<tr>
<td>Other mutations</td>
<td>Telomere shortening is observed in PanIN1 (127), CDKN2 and HER2 mutations are seen in PanIN-2 while p53, Smad4, BRCA2 are seen in PanIN-3 (61, 194, 195)</td>
<td>P53, p16 mutations are observed in high grade IPMN (196, 197). Smad4 mutations appear in invasive IPMN (38). BRAF mutations are seen in 1-3% of IPMNs (198, 199). GNAS mutations appear to be specific to IPMN (200) IPMNs show RNF43 inactivating mutations (201).</td>
<td>CDKN2, p53, Smad4, BRCA2, p16 mutations are well established in pancreas cancer (94, 117, 202). GNAS mutations are not commonly seen in PDAC unless arising from IPMN</td>
</tr>
<tr>
<td>DUSP6</td>
<td>Not usually suppressed in PanINs</td>
<td>17% of IPMNs show abrogation of DUSP6 (203)</td>
<td>DUSP6 is frequently suppressed in pancreatic cancer (66%) (203, 204)</td>
</tr>
<tr>
<td>Chromosomal changes</td>
<td>Somatic copy number changes are seen in up to 20% of PanIN including monosomies of chromosomes 6, 17 and 18 (205)</td>
<td>IPMN with moderate or high-grade dysplasia or invasive carcinoma commonly show 5q, 6q and 11q loss as compared to PDAC (206).</td>
<td>Loss in chromosomes 18, 13, 12, 17 and 6 with gains in chromosomes 20 and 7 (103, 207)</td>
</tr>
<tr>
<td>miRNA changes</td>
<td>miR-21, miR-221, miR-196, miR222, let-7A expression increase with increasing grade of PanIN (208, 209).</td>
<td>miR-155 and miR-21 are over-expressed in high grade IPMN (210, 211)</td>
<td>miR-132, miR-212, miR-29C, miR-96, miR-143, miR-1290 are over-expressed compared to benign tissues (212-214).</td>
</tr>
</tbody>
</table>

### 2.8. Practical Applications of molecular biology of IPMNs

The aim of understanding molecular changes in cancer is to develop sensitive and specific diagnostic tests to facilitate an accurate early diagnosis and develop therapeutic measures to aid clinical management. The following genes are targets for therapeutic strategies.
i. **K-Ras:** K-Ras has long been tried as a marker for the early diagnosis of IPMNs as well as for pancreatic ductal adenocarcinoma, since mutations of KRAS prevail in almost all of these neoplasms. However, despite a number of studies assessing its utility in diagnosis, the clinical utility of K-Ras remains unknown (215).

ii. **PIK3CA:** PIK3A mutations appear to be important in a subset of IPMN carcinogenesis. Drugs targeting the PI3K pathway have been developed and tested, and among such drugs, temsirolimus and everolimus, derivatives of rapamycin that inhibit the mammalian target of rapamycin, have been approved for clinical use in the United States (216).

iii. **Alterations of the DNA copy number** seem to be an early event for the development of IPMNs. Chromosomal alterations in EUS fluid could help differentiate low-grade lesions from high-grade lesions in combination with cytology. Fritz et al used microarray-based comparative genomic hybridization to evaluate IPMNs ranging from mild to moderate grade dysplasia to IPMN associated carcinoma. They found that none of the low grade IPMNs showed any detectable chromosomal aberrations. IPMNs with moderate and high-grade dysplasia showed frequent copy number alterations. They found a higher incidence of loss of chromosomes in 5q, 6q and 11q with high-grade dysplasia or invasion compared to pancreas cancer. Chromosomal 6q loss was found in 10/13 samples of patients with moderate dysplasia or malignancy. Cell free DNA in cyst fluid could be analyzed to demonstrate 6q 5q and 11q LOH by fluorescent in situ hybridization could be used as a panel in differentiating IPMN from ductal adenocarcinoma (206, 217).

iv. **SHH:** Recent studies suggest that SHH measurement of pancreatic juice may provide some advantages in the treatment or follow-up of a subset of patients with IPMN. Sonic hedgehog blockers such as cyclopamine are available (218). A water soluble derivative of parthenolide has been targeted against activated nuclear factor–kappa B have also been tried in small scale trials in pancreas cancer (219).

v. Detection of aberrantly methylated genes in pancreatic juice obtained at EUS may be useful preoperatively to predict the presence of an invasive focus in IPMNs. The functional abrogation of genes like p16, p53 and DPC4 are routinely observed in pancreas cancer.
However, IPMNs do not show mutations in these genes commonly, thereby lending
consideration of the possibility of transcriptional silencing of these genes in the progression of
IPMN to cancer. Sato et al analysed the status of 7 CpG islands (including ppENK, p16, p53)
using methylation specific PCR in 51 IPMNs of different grades. They found that more than
80% of iPMN exhibited hypermethylation of at least one of the CpG islands with increased
hypermethylation of ppENK and p16 observed in high-grade IPMN as compared to low grade
IPMN. They also observed that the average number of hypermethylated loci was greater in
high-grade lesions than low-grade IPMN (220).

vi. Human telomerase reverse transcriptase (hTERT) expression has been shown to indicate
malignant transformation in IPMN (221). The immunohistochemical detection of telomerase
activity in pancreatic juice specimens can differentiate between borderline IPMN and
carcinoma (222). MUC4 and MUC5AC have also been put forward as potential markers in
distinguish between benign and malignant IPMNs (223).

2.9. Management of cystic tumours

A registry of patients with suspected pancreatic cystic tumors was maintained with patients being
registered on visits to the outpatient department and multidisciplinary meetings. The collected
dataset included details of referral, radiological and endoscopic findings, pathological data and
outcomes.

2.9.1. Criteria for Surgery

Criteria for resection included all main and mixed duct tumors in prospective surgical candidates.
Candidates were assessed prior to surgery by cardiopulmonary exercise testing for fitness for
surgery. Branch duct lesions were resected if the patient was symptomatic, or had lesions >3cm
and/or mural nodules. Asymptomatic cystic lesions without main pancreatic duct dilatation
(>6mm), mural nodules and <30mm in size have a low risk of progression to invasive cancer and
were followed up at 12-36 month intervals (224). It has been recommended that follow up
intervals be shortened depending on the increasing diameter of the cyst and the interval
lengthened after 2 years of non progression (224, 225).
2.9.2. Diagnostic and Management Algorithm

Figure 20: Strategy for the evaluation and management of pancreatic cystic neoplasms (226)

2.10. Methods

2.10.1. Setting

Hepato-Pancreato-Biliary (HPB) Surgery and Gastroenterology services of an academic medical centre serving a predominantly urban population of 3.2 million people. During the period of recruitment for the present study, this centre was one of two HPB units serving the conurbation.
2.10.2. Design

Observational study of a clinical cohort recruited from a single centre.

2.10.3. Methods

A computerized database for recording of the management of patients with cystic tumours was established in October 2009. Patients were identified from the outpatient clinics of both the surgical and gastroenterology services. Data were collected on patients referred to these clinics over period from October 2002 to September 2012. The dataset was closed for analysis on January 2013. All data were recorded, including age, sex, the lag time between first referral to a specialist clinic and confirmation of the diagnosis, clinical presentation, morphological features on radiological investigations, predictive criteria of malignancy, treatment (medical, endoscopic or surgical), cytological (and/or pathological) findings, follow-up examinations, events during follow up and blood and cyst fluid markers of malignancy, integrated clinical diagnosis and subsequent management. Co-morbidities were recorded at time of baseline presentation (these were not prospectively protocol-defined but recorded according to the descriptions provided in the case notes). Two types of follow-up were established for patients. Thirty-nine patients with a diagnosis of a pancreatic cystic neoplasm, who presented between 2002-2009, were re-assessed based on the Sendai criteria and placed in the follow-up program and sixty patients were prospectively assessed and followed up for the presence of a pancreatic cystic neoplasm and enrolled into the database. The increased number of pancreatic cystic tumours diagnosed from 2009 onwards can be attributed to greater recognition of these tumours as well as the development of a cystic tumour database to follow these patients up. This increased number of patients picked up over the ten years has been seen in other centres as well (227).

2.10.4. Study population

Ninety-nine patients were evaluated for suspected pancreatic cystic neoplasms. Patients with a clinical course suggestive of pancreatic pseudocyst were excluded. The diagnosis was established on a combination of radiological or endoscopic data.
2.10.5. Imaging pathways

All patients underwent computed tomography. The majority of scans were undertaken to a pre-defined standard pancreas protocol. Endoscopic retrograde cholangiopancreatography (ERCP) with placement of a plastic endobiliary stent was undertaken in patients presenting with jaundice. Endoscopic ultrasonography (EUS) was planned in all patients although co-morbidities prevented some patients from undergoing this procedure. A subset of patients underwent magnetic resonance cholangiopancreatography (MRCP) to assist in clarifying the anatomy of the cystic lesions. Lesions were classified according to their relation to the main pancreatic duct and structural characteristics into main duct intraductal papillary mucinous neoplasms (Figure 21), branch duct IPMN (BD IPMN) (Figure 22) and mucinous cystic neoplasms (Figure 23) and serous cystic neoplasms. Radiological criteria were defined as main duct when diffuse or segmental duct dilatation was present with a main pancreatic duct diameter of 10mm or more; branch duct when cystic lesions (either single or multiple) were communicating with pancreatic ductal branches with main duct dilatation < 10mm. Mucinous cystic neoplasms were defined as single cysts with no demonstrable communication with the main pancreatic duct and present in women.

Imaging features were identified on a combination of CT, EUS, ERCP or MRCP examinations. An attempt was made to clarify the diagnosis by means of at least 3 investigations – CT and/or MRCP, EUS and FNA cytology. Investigations were necessarily tailored depending on each individual patient’s comorbidities, suitability for a particular investigation and choice. Certain patients were not suitable for some investigations due to anatomical features precluding EUS and FNAC or MRCP due to contraindications were not subjected to this ‘triple assessment’. Patients with equivocal lesions on initial examination were followed up and excluded if subsequent findings on morphological examinations revealed a benign pathology such as a simple cyst or chronic pancreatitis with a pseudocyst.
Figure 21: Main duct IPMN on CT scan (228)

Figure 22: MRCP imaging of branch duct IPMN in the head and body of the pancreas. The MRI image clearly shows communication with the pancreatic ductal system without main pancreatic ductal dilatation (229)
All patients with suitable anatomy underwent EUS examinations to obtain a cytological and biochemical sample. Occasional patients underwent an endoscopic examination as part of a workup for an unrelated condition, wherein, the presence of mucous pouting from the ampulla could be observed (Figure 24). On EUS, features sought for included number of lesions, location, size, presence of a mass (Figure 25), ductal communication (Figure 26), septations and mural nodules (Figure 27) (231). The presence of mucin pouting from the ampulla on ERCP is also suggestive of IPMN.

Figure 24: Fish-eye ampulla on ERCP describes mucin pouting from the ampulla in IPMN (231)
Figure 25: Radial EUS examination showing main duct IPMN with a dilated main pancreatic duct with evidence of an intraductal tumour (arrows) (231)

Figure 26: Radial EUS examination showing branch duct IPMN with no main duct communication (231)

Figure 27: Radial EUS examination showing mural nodules seen in a dilated main pancreatic duct (231)
The character of the aspirate (mucinous or serous) was also noted. Cytological samples were stained with alcian blue and Giemsa stains. On cytological examination, the samples were classified as malignant, mucinous with severe, moderate or mild dysplasia, benign serous or indeterminate in nature. The aspirated samples, if sufficient in volume, were also assessed for CEA and amylase levels.

### 2.10.6. Cyst Characteristics

The cyst characteristics that were assessed on CT included cyst size, location, main pancreatic duct dilatation, pancreatic duct communication, whether single or multifocal, simple or complex, or if there was a solid component or presence of intramural nodules. In addition, the presence of cyst calcification and lymph nodal enlargement was recorded. These features were corroborated on EUS, with emphasis on the establishment of pancreatic ductal communication and the presence of a mucinous or serous aspirate. Finally, the aspirate cytology was used to further categorize the cyst into mucinous, serous or malignant pathology.

### 2.11. Management plan

Patients underwent at least two initial investigations to confirm the diagnosis (imaging, endoscopic and cytological) at the first presentation. They were discussed in a multidisciplinary team meeting and a plan of management was agreed on with patients being allocated either to surgery or for conservative management based on the cyst pathology, risk of progression to malignancy, co-morbid status and assessment of fitness for surgery. In the group of patients managed conservatively, yearly clinical and radiological follow up was instituted. The small proportion of patients who were not fit for surgery or declined an operation were given an open yearly follow up to ensure symptom control and assessment of disease progression if they wished. This group of patients was aware that further imaging would be useful only in terms of planning for symptom relief and would not change their management and were consented for further imaging when performed.

Further follow up took the form of CT imaging of the pancreas at intervals of 12 months. MR scans were used in a proportion of patients wherein the preliminary scan used was an MR of the
pancreas. In the event of a change in features on the scans, EUS was used to confirm changes and assess cytology. All scans were reviewed an independent gastrointestinal radiologist blinded to the initial report, to ensure that scans were reported in the appropriate format.

2.11.1. Inclusion and exclusion criteria

Patients were included if they satisfied findings on CT revealing cysts with or without pancreatic ductal communication, confirmed on EUS with cytology and biochemical analysis. MRCP and ERCP were used when the nature of lesions could not be determined on the two modalities stated above or if the procedures were required for additional pathology. Patients were excluded if they were subsequently judged in a multidisciplinary (MDT) meeting to have chronic pancreatitis with a pseudocyst or if the diagnosis was not confirmed on imaging modalities and cytology.

Descriptive statistics were used for all study variables including mean and standard deviations for continuous data and frequency distributions for categorical data. Analysis was performed to identify associations between cyst and patient characteristics and outcome data was measured using the Mann-Whitney U test. All analyses were performed using SPSS (version 20.0, SPSS Inc. Chicago US).

2.12. Results

2.12.1. Clinical profile

*Presenting population of suspected pancreatic cystic tumours.*

A total of 99 patients with suspected pancreatic cystic tumours were enrolled into the database over the 3-year period. The median age of the patients was 65 (range 35 – 89) years. The group included 38 males (38%) and 61 females (62%). Data on delay from the date of referral to the first outpatient clinic was available for 75 patients (76%). In the remaining 24 patients, referrals made through other hospitals precluded an accurate calculation of the date of referral to the first outpatient clinic. The median delay from specialist referral to attendance at the HPB/Gastroenterology clinic was 1 month (range 0 – 75). Eighty-six patients (87%) had their case reviewed at a multidisciplinary team meeting prior to initiation of treatment plan. No
multidisciplinary meeting data could be found in thirteen patients. This subset included only patients followed up from between 2002 to 2009, wherein the meeting data was not routinely recorded. The median period of follow-up was 24 months (range 0 – 124). For the purpose of this study, the recorded follow up was evaluated at 24 months to ensure assessment of the dataset and outcomes. Patients will continue to be assessed prospectively and followed up at the end of this study. Forty-two patients (42%) were asymptomatic at time of presentation (in these patients, the most frequent mode of presentation was an incidentally discovered lesion). Twenty-two patients (22%) presented with abdominal pain and 12 (12%) with jaundice. Other presentations included new-onset diabetes mellitus, weight loss and rigors.

Table 4 summarizes the clinical profile of patients with suspected pancreatic cystic tumours.
Pancreatic surgery carries an independent risk of mortality and morbidity and is influenced by an individual’s comorbid status. Individuals were assessed for surgery using cardiopulmonary exercise testing (CPET). CPET testing is a non-invasive simultaneous testing of the cardiovascular and respiratory capacity of a patient during exercise. The test results also take into account any cardiac or respiratory medications and any previous cardiac or thoracic surgery (232). CPET testing allows an accurate representation of the ability of the patient to tolerate the

![Table 4: Clinical profile of patients presenting with pancreatic cystic lesions](image)
cardiovascular strain that surgery imposes and as such can categorise the risk of major surgery in an individual patient. Other organ dysfunction scores can also be used such as the multiple organ dysfunction score (MODS) (233) and the sequential organ failure assessment score (SOFA) (234). Both these scores have been validated in the critically ill patient in the ICU setting and in general, represent predictors for post-operative morbidity and mortality. The physiological and operative severity score for enumeration of mortality and morbidity (P-POSSUM) score is another score to classify an individual's preoperative and post operative risk; but has been shown to overestimate the morbidity and mortality for patients undergoing pancreatic surgery (235). We found that a total of 56% of patients had moderate to severe cardiovascular, respiratory or renal disease adversely limiting their fitness for surgery. In this high-risk group, those with lesions suggestive of benign or ‘not suspicious’ pathology were managed conservatively.

2.12.2. Results of investigations

All 99 (100%) patients underwent an index CT scan. EUS with cytology was performed in 70 (71%) with comorbidity, altered anatomy or surgery being the primary causes for the procedure not being carried out in 29 patients. In patients not undergoing EUS for diagnosis, the diagnosis was founded on either surgical histology or sequential imaging with or without a second imaging modality (MRI). ERCP with stent placement was undertaken in 11 patients. The diagnoses in the total population of patients presenting with suspected pancreatic cystic neoplasms are detailed in Table 5.

Table 5: Diagnosis after triple modality investigation and MDT review

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Total number of patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main duct IPMN</td>
<td>13</td>
</tr>
<tr>
<td>Branch duct IPMN</td>
<td>40</td>
</tr>
<tr>
<td>Mucinous cystic</td>
<td>11</td>
</tr>
<tr>
<td>Serous cystadenoma</td>
<td>4</td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>8</td>
</tr>
<tr>
<td>Renal cell carcinoma</td>
<td>1</td>
</tr>
<tr>
<td>Endocrine cancer</td>
<td>1</td>
</tr>
<tr>
<td>Metastatic cancer</td>
<td>1</td>
</tr>
<tr>
<td>Benign cyst</td>
<td>6</td>
</tr>
<tr>
<td>Pseudocyst/ pancreatitis</td>
<td>14</td>
</tr>
<tr>
<td><strong>Total numbers</strong></td>
<td><strong>99</strong></td>
</tr>
</tbody>
</table>
2.12.3. Management and Outcomes

Following the initial imaging, patients either underwent surgery or were managed conservatively. Indications for surgery included the diagnosis of resectable cancer in patients fit for surgery, suspicious features indicating high risk of malignant progression or presence of a mass associated with a pancreatic cystic lesion. The group of patients managed with a conservative ‘wait and watch’ approach included benign cystic lesions (SCA, pseudocyst), branch duct IPMN, main duct IPMN and MCN in patients not fit for an operation. This is summarised in Table 6.

**Table 6: Management of the whole population of suspected pancreatic cystic tumours**

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Total number of patients</th>
<th>Operated</th>
<th>Conservative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main duct IPMN</td>
<td>13</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>Branch duct IPMN</td>
<td>40</td>
<td>4</td>
<td>36</td>
</tr>
<tr>
<td>Mucinous cystic neoplasm</td>
<td>11</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>Serous cystadenoma</td>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>8</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Renal cancer/ Endocrine Ca</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Metastatic cancer</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Benign cyst</td>
<td>6</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Pseudocyst/ pancreatitis</td>
<td>14</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td><strong>Total numbers</strong></td>
<td><strong>99</strong></td>
<td><strong>33</strong></td>
<td><strong>66</strong></td>
</tr>
</tbody>
</table>

Thirty-three patients underwent surgical resection in total. This group included 22 patients with either main duct IPMN, cancer or MCN. In addition, 4 patients with a diagnosis of branch duct IPMN underwent resection (2 for lesions with severe dysplasia/ borderline malignant and 1 each for associated biliary stricture and suspicious imaging). The most common surgical procedures performed were Whipple’s procedure and distal pancreatectomy in 12 patients each with 3 patients undergoing total pancreatectomy and 6 patients undergoing other procedures (enucleation, pseudocyst drainage). In 6 patients, treatment included palliative chemotherapy or symptom control due to progressive unresectable malignancy while 7 patients either declined surgery or were unfit for a major procedure based on their co-morbid status. In the table above, 2 patients with MCN did not undergo surgery. This was due to their underlying comorbid status; with one patient having end stage renal failure and the other undergoing conservative
management of a colovesical fistula due to a poor physiological reserve. Two patients with serous
cystadenoma had surgery. One of these patients presented with duodenal obstruction due to
increasing size of the SCA and the second developed recurrent common bile duct obstruction
and underwent a hepaticojejunostomy. All four patients with primary pancreatic adenocarcinoma
managed conservatively, had chemotherapy. All four were found to have unresectable disease
and in one case a concomitant oesophageal malignancy. In the main duct IPMN group, the four
patients managed conservatively included two patients with other malignancies diagnosed
including one myeloma and one ovarian malignancy; the other two patients declined both EUS
and surgery.

2.12.4. Main duct IPMN

Thirteen patients were diagnosed as having a main duct IPMN. The median age was 68 years at
presentation (range 35-89 years). Nine out of thirteen patients were symptomatic at presentation
with abdominal pain being the most common symptom. Four patients had evidence of
pancreatititis at some time and 2 patients had jaundice. On CT, 6 patients had diffuse disease that
was not restricted to one part of the gland. Seven patients had complex disease on presentation
with 4 of these presenting with a solid component. Mural nodules were observed in 2 patients on
CT and 5 patients on EUS. Eight patients underwent EUS with cytology showing mucinous
neoplasm in 7 and adenocarcinoma in one patient.

Management and Outcomes for main duct IPMN

From the 13 patients, 9 patients underwent operative intervention while 4 patients were managed
conservatively. Of the 4 patients that were managed conservatively, two patients had a
concomitant extra-pancreatic malignancy (1 endometrial cancer and 1 myeloma, both undergoing
treatment). Of the other 2 patients, one was followed up in view of advanced age (89 years) and
the second patient declined intervention. Both these patients showed disease progression on
follow up, with the former developing mural nodules and the latter showing increased main
pancreatic duct dilatation.

Nine patients with main duct IPMN underwent surgery. Intraoperative findings and histology
revealed 6 patients (46%) as having IPMN adenocarcinoma with four having R1 resections, one
had inoperable disease and one had a R0 resection. The high R1 rate could be attributed to several reasons. Two patients had multifocal and diffuse disease while two others had T3 tumours with margin involvement and vascular and perineural invasion. This rate of R1 resection is higher than that previously reported in the literature, with R1 resection rates reported between 9-14% for invasive IPMN (236, 237). The underlying histological subtype of the IPMN plays a role with anaplastic and tubular adenocarcinoma having a worse prognosis than colloid carcinoma arising from an IPMN (238). However, R1 resection rates for pancreatic ductal adenocarcinoma have been variously reported between 20-50%, increasing with more intensive histological workup (239-243). There exists the possibility that in our series, the main duct dilatation observed on CT and EUS could have been secondary to an underlying adenocarcinoma, leading to a misclassification of the primary tumour pathology. However, it must be noted that all the adenocarcinomas were reported as mucinous adenocarcinomas, thereby implying origin from an IPMN. Three patients (23%) had main duct IPMN with dysplasia (ranging between moderate to high grade dysplasia). On follow up, the 5 patients with R1 or inoperable disease (55%) died due to disease progression. Of this subset, one developed a second primary (lung adenocarcinoma) and one had a pancreatic neuroendocrine tumour.

Chemotherapy

All patients with a primary adenocarcinoma with nodal spread and those who underwent surgery with R1 resections underwent systemic chemotherapy.

Survival

Overall survival following resection was analysed. The small number of patients in the series and the follow up period in this study necessarily limited the survival data. The median survival for patients with R1 disease and adenocarcinoma (6 patients in total) was 11 months (range 4 -19 months). This contrasted with 3 patients with R0 resections and a histological diagnosis of moderate dysplasia with IPMN who were alive at the end of the study period.
Group Mortality

The overall mortality rate related to the primary pathology (main duct IPMN) in this group of patients was 38% (5/13). The morbidity was 44% (4/9) related primarily to the underlying physiological status of the operated patients. Complications included acute coronary syndrome in two patients, exacerbation of COPD in one and aspiration pneumonia in another. There were no pancreatic leaks identified. Median length of hospital stay was 12 days (range 8-25). There was disease progression in 2 patients and other pathology adversely impacting management of the IPMN in a further 2 patients. The rate of extra-pancreatic malignancies was significant in this subgroup of patients with 3 out of 9 patients (33%) having either a concomitant or developing a second malignancy. In addition, 2 out of 4 patients demonstrated disease progression when followed up. Despite the small number of patients with main duct IPMN, this demonstrates that this subgroup should be aggressively treated initially and monitored for the development of an extra-pancreatic primary post-operatively.

2.12.5. Pancreatic cystic lesions ≤ 3 cm in size

Forty-five patients had cystic lesions measuring 3 cm or less in size. The median age at presentation was 66 years (range 35-89 years). Patients were followed up for an average of 24.5 months (7-115 months). Twenty-one patients were asymptomatic on presentation (46%) while abdominal pain was the commonest symptom in 26%. Lesions were most common in the head of the pancreas (42%), however, significantly, disease was multifocal in 19 patients (42%) and complex cysts were seen in 11 (26%). The mean cyst size on EUS was 20mm with a mass present in 5 patients (11%). The diagnoses of these lesions are detailed in Table 7. Fifty-three per cent of lesions less than 3 cm in size were diagnosed as branch duct IPMNs. Benign cystic pathology (pseudocyst, simple cyst and SCA) accounted for 17% while main duct IPMN (13%), all cancers (8%) and MCN (6%) accounted for the rest. In this group, 9 patients (20%) underwent a surgical procedure including Whipple’s procedure (4 patients), distal pancreatectomy (2 patients) and total pancreatectomy (3 patients).
Table 7: Management of pancreatic cystic lesions less than 3cm in size

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Total number of patients</th>
<th>Operated</th>
<th>Conservative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main duct IPMN</td>
<td>6 (13%)</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Branch duct IPMN</td>
<td>24 (53%)</td>
<td>1</td>
<td>23</td>
</tr>
<tr>
<td>Mucinous cystic neoplasm</td>
<td>3 (6%)</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Serous cystadenoma</td>
<td>1 (2%)</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>3 (6%)</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Metastatic cancer</td>
<td>1 (2%)</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Benign cyst</td>
<td>3 (6%)</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Pseudocyst/ pancreatitis</td>
<td>4 (8%)</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Total numbers</td>
<td>45</td>
<td>9</td>
<td>36</td>
</tr>
</tbody>
</table>

*Surgical outcomes for lesions less than 3cm*

Mortality for lesions measuring less than 3 cm in diameter was related directly to the pathology for which surgery was carried out. Of the 9 patients operated on, there were no perioperative deaths. Two patients were found to have IPMN related adenocarcinoma on histology and were reported to have R1 resections, both subsequently died of recurrent disease. Histological diagnosis showed features of main duct IPMN. The solitary branch duct IPMN undergoing surgery showed features suggestive of progression to main duct IPMN and underwent an R0 resection.

*Outcomes for patients with lesions less than 3cm managed conservatively*

There were 36 patients who underwent conservative management. These included 23 branch duct IPMNs, 2 patients with main duct IPMN and one case each of adenocarcinoma and metastatic adenocarcinoma (Table 7). There were 6 deaths in this group. Allocating the cause of mortality according to the type of lesion, there were 6 deaths in the branch duct IPMN group. Two patients, including one adenocarcinoma and one metastatic cancer, were alive at the time the study was closed. Causes of mortality in the branch duct IPMN category included 3 deaths due to unrelated medical causes, 1 death due to concomitant metastatic bowel cancer and 2 deaths related to progression of disease with the development of pancreatic adenocarcinoma.
Group Mortality

The overall mortality in this group was 17% (8/45). Of these, 6 were unrelated to the primary pathology. The disease specific mortality was 4% (2/45). However, in view of the anticipated mortality of two additional patients with inoperable adenocarcinoma (primary pancreatic and metastatic) that were alive at the time the study closed, the disease specific mortality in this group of cystic lesions <3cm in size could be revised upwards to 6% (3/45) and the overall mortality adjusted to 20% (9/45).

2.12.6. Branch duct IPMN

Forty patients had cystic lesions classified as branch duct IPMN. The median age at presentation was 66 years (range 35-89 years). Twenty-six patients were female (65%). Patients were followed up for a median of 26 months (5-115 months). Nineteen patients were asymptomatic on presentation (47%). Abdominal pain was the commonest presenting symptom in 15% (6/40). On CT, lesions were most common in the head of the pancreas in 17 patients (42%). Multifocal cystic disease was seen in 14 patients (35%), while single complex cysts were seen in 14 (35%). Four patients (10%) had both complex and multifocal cysts. The cyst size was less than 30mm in 24 patients (60%) and only 19 patients (42%) showed evidence of pancreatic ductal communication on CT. Twenty-three patients underwent EUS. The mean cyst size on EUS was 20mm with single lesions reported in 30% and a mass present in 3 patients (7.5%). Cytology showed adenocarcinoma in 3 patients (6%) and was non-diagnostic in 8 patients (20%).
Management and Outcome for branch duct IPMN

Forty patients were categorized as having branch duct IPMN on preliminary investigations. Of these, 36 patients were managed conservatively and 4 patients underwent surgery.

Surgical Outcome

Four patients underwent operative intervention. These included 1 patient with a diagnosis of BD IPMN with multifocal disease and high Ca19-9 levels, one with an associated mass and two with a rising blood Ca19-9 level. In this context, a rising Ca19-9 associated with high-risk features (such as a mass or multifocal disease), potentially indicated the development of carcinoma and thereby triggered consideration for resection. Post-operative histology revealed IPMN with moderate atypia in 2 patients, a diagnosis of pancreatic pseudo-papillary tumour in the patient with an associated mass, with features of chronic pancreatitis in one patient. There was no operative or disease specific mortality and none of the patients received any chemotherapy.

Outcomes following Conservative Management

Thirty-six patients were managed conservatively. This group was followed up for a median of 31 months (range 5-115). One patient was found to have adenocarcinoma on cytology during initial EUS and was offered chemotherapy, as she was not fit for resection. Seven patients (17.5%) showed a change or progression of disease over a median period of 24 months (range 12 – 36 months). This included 2 patients with expanding cyst size, one with development of additional cysts and 4 patients with onset of progressively more complex cysts with or without a solid component. All patients with a documented change in imaging underwent re-evaluation with EUS or at a MDT meeting. Of the cases that developed complex cysts, two progressed to IPMN adenocarcinoma, one had cytology that displayed a mucinous neoplasm with high-grade dysplasia, one developed a lung mass and one was re-classified as a SCA. Two patients showed a progressive increase in cyst size on surveillance CT scans. EUS cytology was non-diagnostic in one and showed moderate dysplasia in the other. The option of surgical intervention was limited in the patient with moderate dysplasia due to underlying comorbidities, while the second case underwent surgery. The histology was subsequently reported as chronic pancreatitis with cystic change.
Incidence of adenocarcinoma in branch duct IPMN

Overall, at the time of diagnosis, 1 patient was found to have an associated adenocarcinoma on CT and EUS cytology. In addition, a further 2 patients (5.5%) developed adenocarcinoma during routine monitoring. One other patient developed high-grade dysplasia during surveillance. Thus, four patients (11%) were found to have or subsequently develop branch duct IPMN associated severe dysplasia or adenocarcinoma. In the surveillance group, the rate of progression to severe dysplasia/adenocarcinoma was 8.3% (3/36) over the follow up period of 24 months. Furthermore, two more patients showed progressive disease with evidence of moderate dysplasia. Thus the cumulative rate of progression was 13.8% (5/36) in the surveillance group of patients with branch duct IPMN.

Group Mortality

The overall mortality rate in branch duct IPMN was 15% (6/40) with three deaths related to the presence or development of IPMN adenocarcinoma, two deaths were due to the development of extra-pancreatic malignancy and 1 death was due to medical causes. All three patients with IPMN associated carcinoma had chemotherapy after discussion in a multidisciplinary meeting.

Incidence of extra-pancreatic malignancy (EPM)

Three patients out of 53 total IPMN (5.6%) were noted to develop synchronous malignancies during follow up. This included one patient who developed a lung mass in addition to cyst progression, a second patient who developed a metastatic renal cell carcinoma and a third with a concurrent oesophageal cancer.

In this study, the overall occurrence of EPMs was 26% overall among 99 suspected pancreatic cystic neoplasms. We found a total of 23 previous cancers (23%) in twenty-two patients with pancreatic cystic tumours. These cancers included among others three colon cancers, two each of renal cancer, bladder cancer, breast cancer, MEN-1 syndrome and oropharyngeal cancer with one each among prostate, thyroid, salivary, spinal, melanoma, myeloma, endometrial and vulval cancers. One patient had haemangioblastomas (as part of Von Hippel-Lindau disease) and one had a gastric B-cell lymphoma at the time of presentation. One patient was found to have familial
adenomatous polyposis in the group with the concomitant presence of BD IPMN. There was no malignancy involved however.

In the population comprised exclusively of individuals with IPMN, the overall occurrence of EPM was 37.7% (20/53). This included 17 EPM (32%) that were previously diagnosed and 3 that were synchronous. The most common EPMs included breast, bladder and MEN-1 syndrome. This is a significantly high proportion (3%) in comparison to the natural incidence (0.4%) of malignancies in the general population. The UK National cancer intelligence network reports an annual age standardized incidence of 402.8 new cases per 100,000 population (244). The annual 10-year age standardized proportional prevalence of non-melanoma cancer in the UK has been reported as 1501 per 100,000 population (244) and is this is significantly lower than the prevalence of malignancy in our group (37.7%).

2.13. Discussion

This is the largest prospective cohort study within the UK aimed to document the natural history of suspected pancreatic cystic neoplasms over a period of three years. We found that careful classification with a combination of modalities was required to monitor and plan appropriate management. Pancreatic cystic neoplasms are relatively rare, the incidence being variously reported at 2.04 cases per 100,000 persons (95% confidence interval, 1.28–2.80) and the prevalence of 25.9 cases per 100,000 individuals or 1 case in every 3852 individuals (165). In another study, an analysis of more than 24,000 CT and MR scans over an 8-year period demonstrated that only 0.7% of studied scans had pancreatic cystic neoplasms (245). Despite the relative rarity of these lesions, they offer us a valuable insight into the development of pancreatic cancer. More importantly, they offer us a chance to monitor and predict the development of pancreatic cancer at an early stage, thereby improving survival outcomes.

The study group was similar in demographic characteristics in the literature (246, 247). The median age of presentation with pancreatic cystic tumours was 65 years (range 35-89 years) with a female preponderance and an almost even split between symptomatic and asymptomatic patients. We also collected data on the comorbid status of each individual. Over a period of observation prior to the onset of the study, anecdotal data suggested that our patient group was
elderly and as such posed a riskier proposition for major surgery if needed. Therefore, patients were ranked by the degree of organ dysfunction (single/multiple) into mild, moderate and severe categories of impairment and we confirmed our observation that a significant proportion (56%) of patients presented with multisystem dysfunction in addition to the accompanying cystic tumour. This finding may merely be reflective of the age group in which the disease presents, however, it carries significant weight in deciding the management approach in a group where the morbidity and mortality of major surgery would preclude it as an option.

The accurate categorization of these lesions is imperative in order to plan their management and accurately predict their evolution. We attempted a triple modality assessment including radiology (CT or MRI), EUS and cytology to define the characteristics of each individual suspected cystic neoplasm. All patients underwent radiological studies, however, EUS and cytology was only carried out in 70%, typically limited either by anatomical constraints or fitness for the procedure. The number of tests used to precisely categorize a subset of these patients emphasizes the challenge faced in making a conclusive diagnosis. This has been noted previously by other authors (248, 249). CT is the primary modality used in the preliminary classification of these lesions. Several studies have shown that MRI scans provide a greater degree of certainty with regards to ductal communication, mural nodules and septations than CT (250, 251). MRCP should be the investigation of choice for the preliminary characterization of suspected branch duct IPMN, small incidental cysts and cysts with suspicious features. EUS is useful in detecting malignant lesions by delineating mural nodules and surrounding invasion (252, 253). In 99 suspected cystic neoplasms within the study, we found that despite a careful clinical history, imaging studies, EUS and cytology, 20 patients had a benign pathology (5 benign, 15 pancreatitis with pseudocyst). Five of these benign cases (1 foregut cyst and 4 pseudocysts) underwent surgery for suspected malignancy. This is a significant proportion of patients with a purely benign pathology to expose to the risk of major abdominal surgery and signifies the need for multimodality assessment combined with judicious observation in equivocal lesions. A recent study has claimed that testing pancreatic juice for GNAS mutations may help distinguish between MCN and BD-IPMN (254).

In our study group, 8.1% of patients from the complete cohort developed a pancreatic adenocarcinoma. We further evaluated the risk of malignancy in three separate subgroups (MD-
IPMN, BD-IPMN and cystic lesions < 3cm in size) that carry differing risks of progression to malignancy. Features that indicate likelihood of progression to malignancy include cyst size >3 cm, thick walled cyst, mural nodules, MPD dilatation of 10mm and the presence of solid nodules (255-257).

In our set of 13 patients with MD-IPMN, 6 patients were found to have IPMN adenocarcinoma (46%) and three patients (23%) had main duct IPMN with dysplasia (ranging between moderate to high grade dysplasia). During the conservative management of 4 patients with MD-IPMN, disease progression was seen in 2 patients and in 2 patients medical problems prevented the management of IPMN. The frequency of malignancy in main duct IPMN reported in the literature ranges between 36- 100% while the frequency of invasive MD-IPMN ranges from 11-81% ((258-264). Our study had a similar rate of invasive adenocarcinoma as the reported literature. The overall mortality rate related to the primary pathology (invasive adenocarcinoma with main duct IPMN) in this group of patients was 38% (5/13) as 5/6 patients had R1 resections. Also of note, we found that the rate of synchronous extra-pancreatic malignancies was significant in this subgroup, affecting 3 out of 9 patients (33%). This corresponds to previously reported rates of extra-pancreatic malignancy in the literature (265-267). . It has previously been reported that adenocarcinoma in IPMN has a better prognosis (268) however, in the event of stage 2/3 carcinoma, the prognosis is similar to that of primary pancreatic ductal adenocarcinoma (269). In our study, the IPMN carcinomas had an overall mortality rate was 38% as would be expected for primary pancreatic cancer.

We also assessed the management and progression in patients with BD-IPMN. A total of four patients were found to have (or subsequently develop) severe dysplasia or adenocarcinoma (11%) on a median follow up of 31 months. One patient was found to have adenocarcinoma at initial presentation while seven patients (17.5%) showed a change or progression of disease on routine monitoring. Of those 7 cases, 2 (5.5%) progressed to IPMN adenocarcinoma, while one case displayed high-grade dysplasia. The cumulative rate of progression of disease was 13.8% (5/36). In previous studies, the frequency of malignancy in resected BD IPMN has ranged between 7-43% ((261, 270) with the frequency of invasive carcinoma is between 1-35% ((225, 271-273). This is consistent with our study. It has been noted however, that the growth rate of branch duct IPMN is dependent on the rate of growth that the cyst exhibits with cysts growing...
faster than 2 mm/year demonstrating 3- and 5-year cumulative risks of malignancy of 6.4% and 45.5%, whereas cysts growing less than 2 mm/year had corresponding risks of 1.8% and 1.8%, respectively (274). In our group of patients with branch duct IPMN, the median size of the cysts was 20mm (range 7-60mm). Previous studies have found that the incidence of malignant IPMN for cyst sizes of <1 cm was 17% and for cysts <3 cm was 19% (270). In cystic lesions with a size less than 3 cm, the risk was directly related to the underlying pathology. However, BD IPMN cysts <3cm could be safely observed.

The incidence of extra-pancreatic malignancy (EPM) in our study population was examined. It has been previously acknowledged that there is an increased incidence of EPM in IPMN (172). This has been reported to occur at frequencies varying between 10-40% (267, 275-277). The overall prevalence of extra-pancreatic neoplasms in our study among all suspected cystic tumours was 26%.

In the population of patients exclusively comprising of IPMN, the overall occurrence of EPM was 37.7% (20/53). In this group, three patients (15%) were noted to develop synchronous malignancies during follow up including one each of renal, lung and oesophageal cancers. Seventeen EPM (85%) were diagnosed prior to the detection of IPMN with the most common most common EPMs including breast, bladder and MEN-1 syndrome. The prevalence of EPM is higher than previously reported from large population based European studies assessing the prevalence of EPM in IPMN (277, 278). In these studies, the prevalence has been variously reported as between 16.8-23.6% with up to 70-80% of EPM preceding the occurrence of IPMN. In EPMs, colorectal, gastric, lung and breast cancers have been most frequently reported (276, 279). We found a higher rate of breast, bladder and MEN-1 related malignancies with a scattering of other malignancies. One patient in our population had familial adenomatous polyposis (FAP). Some studies have suggested a genetic basis for the EPM associated with IPMN in cases of FAP. Previous case reports have documented the loss of the wild allele of the adenomatous polyposis coli gene in IPMN, causing inactivation of both alleles and a an identical immunohistochemical profile to associated duodenal adenomas thereby indicating that there may be a genetic basis for the development of IPMN in FAP (280, 281). In addition, one study suggested that the transcription of MUC2 might be related to the synchronous EPM seen with IPMN (282). Lastly, studies of patients with IPMN and EPM have suggested that these patients
have a poorer prognosis than those without EPM (172, 175, 265, 275). We did not specifically assess for the prognosis in this subgroup of patients although the notion that these patients have a generalized genetic predisposition towards malignancy warrant further exploration in large scale studies.

**Conclusion**

In conclusion, patients with suspected pancreatic cystic neoplasms should be investigated with at least three modalities to enable accurate characterization. Main duct IPMN carry a significant risk of development of pancreatic cancer while branch duct IPMN carry a lesser risk therefore an observational policy may be adopted in the absence of signs of malignancy. The risk of developing a cancer in cystic lesions less than 3 cm in size is directly related to the underlying pathology. A significant number of these patients have extra-pancreatic malignancies and therefore warrant continued surveillance.
Chapter 3 - Poly A RT-PCR measurements of indicator genes in pancreas cancer and pancreatic cystic tumours

3.1. Introduction

Pancreas cancer is the 9th most common cause of cancer in the UK (1, 283). Although only 8000 cases are diagnosed annually, resection is only possible in 5-15% (2, 3). The poor prognosis associated with pancreas cancer is related to its loco-regional spread prior to diagnosis preventing a curative surgical resection. In addition, the pancreas is difficult to biopsy and chemotherapy and radiotherapy have a limited effectiveness in treating this cancer. This emphasizes the need to develop effective measures to diagnose and treat this cancer. The overall outlook of pancreatic cancer may be improved if patients were diagnosed at an earlier stage or if those patients with lesions known to be pre-malignant could be offered surgery at an earlier stage of the disease.

Thus, current research has focused on identifying the genetic pathways involved in the pathogenesis of pancreas cancer primarily as a means of identifying these cancers at an early stage to facilitate surgical resection. We have previously assessed the feasibility of Poly A PCR as a clinically feasible methodology to test for over-expressed genes in pancreatic cancer. We will aim to further this technique by applying it to a population including both pancreas cancer and pancreatic cystic tumours assessing for the correlation between selected over-expressed genes in pancreatic juice and pancreatic tumour tissue from the same patient group.

3.2. Molecular Pathology of Pancreas Cancer

There is considerable evidence that pancreatic cancer is a genetic disease, in that a series of consistent genetic changes in pancreatic cancer has been shown to occur (32). In addition, support is provided by progressive changes seen in precursor lesions that are consistent with those seen in progression to adenocarcinoma (30), while in familial pancreatic adenocarcinoma, the genetic changes have been well characterized (62).
Pancreas cancer is a disease of inherited and somatic mutations, the progression of which is mediated through a series of progressive, non-random, cascading genetic signals producing alterations in cellular morphology (71). The various associated genetic abnormalities include alterations in chromosome or gene copy number, microsatellite instability, epigenetic silencing, intragenic point mutations, and gene overexpression secondary to increased transcription (284). Genome based studies have been used to identify candidate genes that are involved in the pathogenesis and propagation of the cancer (75, 285, 286). The importance of these candidate genes depends on the frequency with which they are identified in the cancer. Gene expression studies represent epigenetic alterations not detected by sequencing or copy number analysis. It has been shown that two-thirds of pancreatic cancers contained an average of 63 genetic alterations, with genetic alterations seen in 12 core processes (74). In addition to other genetic changes occurring, specific genes may be over-expressed or silenced in pancreatic cancer. Altered gene expression in tumors can result from chromosomal rearrangements, epigenetic silencing, or mutations in genes that are upstream of a target gene. The use of global gene expression technology has identified a multitude of overexpressed genes in pancreatic ductal adenocarcinoma (59, 142, 287, 288). These overexpressed genes may serve as useful either as markers for screening if their gene products are released into the blood or pancreatic juice, or as targets for the development of new therapies and studies to assess the applicability of these over-expressed genes in assisting clinical practice need to be undertaken. These future studies should aim to assess the frequency and reproducibility with which these genes can be measured in the given patient population in clinical practice and the accuracy in assisting with an early diagnosis of the disease.

3.3. Development and use of Poly A PCR

Global cDNA amplification by poly A PCR permits the rapid and economical analysis of a small quantity of cells with very low gene levels. Additionally, multiple genes can be assessed at the same time increasing the likelihood of detection of over-expressed genes. As the cDNA can be amplified when needed, a repository of patient cDNA can be preserved, permitting future clinical use (289). In a previous study, we have shown that it is feasible to use Poly A PCR as a clinical tool in analyzing over-expressed genes in pancreatic juice and tissue by global cDNA amplification using Poly A PCR followed by real time PCR measurement of specific genes (290,
This study aims to follow on with this work, by assessing if a further gene subset derived from a database of collated microarray studies and based on a combination of stromal and epithelial genes may be used as markers of pancreas cancer from samples of pancreatic juice taken from patients who have resections for pancreatic cancer.

3.3.1. Aim of this study

This study aims to assess the clinical utility of over-expressed genes identified in previous studies in pancreatic juice and compare these with the same set of genes in pancreatic tumour tissue. This tests the hypothesis that pancreatic juice may contain pancreatic tumour cells from pancreatic ductal cancer arising from pancreatic ductal cells. It seems reasonable to suppose that these extruded cells may be present in pancreatic juice long before pancreatic cancer is clinically or radiologically visible. Thus over-expressed genes in these pancreatic cancer cells may be detected in pancreatic juice and facilitate an early and curative treatment.

3.4. Patients and Methods

3.4.1. Study Design

The study was designed as a prospective single centre cohort of 2 groups of inter-related patients comprising of individuals with suspected pancreatic cancer and patients with high-risk pancreatic cystic tumors that comprised the study population.

3.4.2. Patients

Patients referred with suspected pancreatic cancer and pancreatic cystic tumors for surgery constituted the study population and were entered into a prospective electronic database over a period of 2 years from November 2009 to December 2011.

A total of 52 consecutive patients were recruited into the study from a total of 57 possible cases. Five patients were excluded due to inoperable lesions. Of the 52 cases, intra-operative pancreatic juice was obtained for 48 patients and freshly dissected pancreatic tumour tissue for 49 patients. This included some samples that were unpaired. This was either due to technical difficulties in
obtaining pancreatic juice intra-operatively or if the dissection of a fresh tumour tissue would disrupt accurate reporting of the cancer as judged by a gastrointestinal pathologist. After all exclusions, the total subset included forty-two paired pancreatic juice and tissue samples as depicted in the figure below. Figure 28.

**Figure 28: Patient recruitment**

The demographic and histological details are documented in Table 8.
<table>
<thead>
<tr>
<th>Case</th>
<th>Sex</th>
<th>Age</th>
<th>Histopathology</th>
<th>Juice</th>
<th>Tissue</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Male</td>
<td>58</td>
<td>IPMN</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>2</td>
<td>Female</td>
<td>52</td>
<td>Pancreatic endocrine tumor</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>3</td>
<td>Female</td>
<td>44</td>
<td>Ampullary adenocarcinoma</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>4</td>
<td>Female</td>
<td>67</td>
<td>Pancreatic Adenocarcinoma</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>5</td>
<td>Male</td>
<td>60</td>
<td>IPMN</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>6</td>
<td>Male</td>
<td>39</td>
<td>Chronic pancreatitis</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>7</td>
<td>Female</td>
<td>61</td>
<td>Pancreatic Adenocarcinoma</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>8</td>
<td>Male</td>
<td>76</td>
<td>IPMN Adenocarcinoma</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>9</td>
<td>Male</td>
<td>62</td>
<td>Pancreatic Adenocarcinoma</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>10</td>
<td>Male</td>
<td>65</td>
<td>Renal Adenocarcinoma</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>11</td>
<td>Female</td>
<td>79</td>
<td>Pancreatic Adenocarcinoma</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>12</td>
<td>Male</td>
<td>59</td>
<td>Pancreatic Adenocarcinoma</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>13</td>
<td>Female</td>
<td>52</td>
<td>Pancreatic endocrine tumor</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>14</td>
<td>Female</td>
<td>63</td>
<td>IPMN</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>15</td>
<td>Male</td>
<td>53</td>
<td>Chronic pancreatitis</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>16</td>
<td>Female</td>
<td>73</td>
<td>IPMN</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>17</td>
<td>Female</td>
<td>41</td>
<td>MCN</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>18</td>
<td>Male</td>
<td>75</td>
<td>Pancreatic Adenocarcinoma</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>19</td>
<td>Male</td>
<td>49</td>
<td>Chronic pancreatitis</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>20</td>
<td>Male</td>
<td>60</td>
<td>Ampullary Adenocarcinoma</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>21</td>
<td>Male</td>
<td>74</td>
<td>Chronic pancreatitis</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>22</td>
<td>Female</td>
<td>78</td>
<td>IPMN</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>23</td>
<td>Male</td>
<td>60</td>
<td>Pancreatic Adenocarcinoma</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>24</td>
<td>Male</td>
<td>79</td>
<td>Duodenal Adenocarcinoma</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>25</td>
<td>Male</td>
<td>75</td>
<td>IPMN Adenocarcinoma</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>26</td>
<td>Female</td>
<td>46</td>
<td>Pancreatic Adenocarcinoma</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>27</td>
<td>Female</td>
<td>64</td>
<td>Pancreatic Adenocarcinoma</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>28</td>
<td>Female</td>
<td>47</td>
<td>Chronic pancreatitis</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>29</td>
<td>Female</td>
<td>54</td>
<td>Pancreatic Adenocarcinoma</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>30</td>
<td>Male</td>
<td>60</td>
<td>GIST</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>
These cases comprised of 46 cases of pancreaticobiliary malignancy and 6 cases of benign pancreatic disease. The median age of the patients was 63 years (range from 27 to 80 years); 25 (48%) cases were female and 27 (52%) were male.

The most common procedure performed was a pancreaticoduodenectomy (44 patients). The procedures performed are shown in Figure 29.
3.4.3. Data collection

Data was collected and recorded on electronic case report datasheets in accordance with trust confidentiality policy. This included:
1) Demographic data: Age, gender and co-morbidities.

2) Clinical Data: type and duration of symptoms, radiological, pathological investigations and their outcomes.

3) Samples collected: pancreatic juice and pancreatic tissue

3.4.4. Sample Storage

Samples were stored in order to permit future evaluation of new biomarkers, providing direct comparisons with the present study.

3.4.5. Ethical Approval

The North-West 9 Research ethics committee approved this study (reference number: 10/H1014/86).

3.4.6. Sample collection

Pancreatic juice and tissue samples were collected from patients undergoing pancreatic surgery at the Manchester Royal Infirmary from the period commencing Dec 2009 to March 2012.

3.4.6.1. Technique for sample collection:

Pancreatic juice samples and matched fresh tissue samples were collected from patients undergoing pancreatic resections using a previously validated methodology (291).

Collection of pancreatic juice

The technique for collection of pancreatic juice during pancreaticoduodenectomy included the placing of a sling under the plane of the planned transection line and lifting the pancreas gently. This controls any local bleeding into the area. The anterior surface of the pancreas was then divided till the pancreatic duct was visualized. If the duct was clearly seen, before transecting it, a 22-gauge needle was used to aspirate pancreatic juice. In the event that the pancreatic duct was of small caliber and not clearly seen, the pancreatic tissue was transected completely and a no. 8
The infant feeding catheter was threaded into the duct. A syringe was then attached to the end of the infant feeding tube and pancreatic juice aspirated. The volume aspirated was then recorded. 250 microliters of pancreatic juice sample was immediately immersed in 750 microliters of Isogen-LS™ (Wako, Tokyo, Japan). Isogen-LS consists of phenol and guanidine thiocyanate inhibiting the RNAses in pancreatic juice.

**Collection of pancreatic tissue**

The tissue specimens were harvested under direct vision within 15 minutes of the tumor being removed. The suspected tumor tissue was biopsied by a pathologist under direct vision, so as to prevent disruption of tissue planes. The identified tumor was cut open and a section of tissue including the stroma taken. The tissue biopsy is then divided into two sections; one of which was preserved as the research biopsy and the second section reported on histology by the pathologist. The research biopsy was weighed so that maximum weight was 30mg or less (according to the parameters stated by the kit used) and then stored in RNAlater™ (Qiagen, Hilden, Germany). The proportion of RNAlater to the tissue section was about 10μL of reagent per milligram of tissue. The reason we chose fresh tissue as opposed to micro-dissected tissue as the preferred method was primarily for two reasons: firstly, the use of fresh tissue correlated with current tissue processing service provisions and secondly, fresh samples permitted the inclusion of stromal tissue in addition to tumor tissue allowing the use of stromal genes to be used in tumour clarification. The significant drawback of this method was the potential for inaccurate sampling of the fresh tissue specimen as it was often difficult to differentiate cut tumour from the surrounding stromal reaction. Freshly cut tumour samples are also limited by the need to preserve tissue planes prior to reporting, in order to ensure correct pathological staging. However, we felt that this necessary limitation of service provision replicated the realities of performing this technique within the bounds of our service. We attempted to overcome this by ensuring a specialist pathologist performed the tumour sampling, reporting of two sets of tissue samples (tumour and research biopsy) and testing for stromal genes. The figure below describes the number of correctly sampled research biopsies compared with the final histological diagnosis (Figure 31).
Figure 31: Histology of the research biopsies compared with that of the reported tumour from the final histological report. Blue represents concurrence between the research biopsies and the final pathological report while red indicates that the research biopsy and the final pathological report did not match (sampling error)

The non-diagnostic research biopsy samples in the figure above represented the discrepancy between that of the actual histology and the biopsy specimen. In all but one case, the samples were reported as normal pancreatic tissue. In a solitary case of a duodenal carcinoma, the research biopsy sample was reported to be a tubulovillous adenoma. Thus, a total of 18 paired pancreatic cancer samples and 8 paired pancreatic cystic tumours were available for analysis. The pancreatic juice and tissue samples were then stored at -80 degrees till further processing.

3.5. Selection of Genes

Compilations of recent gene expression studies have pointed at a considerable number of genes that could convey prognostic and potentially diagnostic information, offering an avenue to an earlier diagnosis of pancreas cancer (78). To have the best chance of determining a correlation between over-expressed genes in pancreatic juice and tumor tissue, a gene set that covered both parenchymal and stromal expression as well as over-expressed genes in precursor lesions (such as pancreatic cystic neoplasms and high-risk chronic pancreatitis) was chosen. In addition, as pancreatic juice is a relatively low yield substrate for analysis, over-expressed genes previously demonstrated in juice were considered.
Earlier papers focused on single gene analysis, but with advances in global gene profiling, large-scale gene expression studies using Serial analysis of Gene Expression (SAGE) or microarray technology have demonstrated a large cohort of over-expressed genes \((74, 75, 78)\).

**Method for the selection of indicator genes**

To select the genes for this study, papers over the past ten years identifying over-expressed genes in pancreatic cancer were searched. A list of these genes was compiled and those genes with the greatest number of hits (multiple published reports of over-expression in pancreas cancer) were shortlisted. This list was then compared with a compilation of over-expressed genes in pancreatic cancer, chronic pancreatitis and precursor lesions \((78)\). The shortlisted candidate genes comprised of genes that were over-expressed more than 4 fold in juice and tissue samples from patients with pancreas cancer and pancreatic cystic tumours. The final selection of 17 genes was made using a combination of the genes over-expressed in a combination of parenchyma, stromal tissue and precursor lesions with attempts to incorporate both pancreas cancer and pancreatic cystic tumours (IPMN). This overlap of genes would hypothetically allow us to assess those genes with the greatest likelihood of success, regardless of areas biopsied. GAPDH was chosen the housekeeping gene. The numbers of genes assessed in this study was restricted by certain factors, including the limitations of cost, availability of substrate and the requirement of over-expression in both pancreatic juice and tumour tissue. Primers and probes were designed for these 18 genes are described below in Table 9.
Table 9: A list of the selected 'indicator' genes

<table>
<thead>
<tr>
<th>No.</th>
<th>Gene</th>
<th>Identifier</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ANXA2</td>
<td>Annexin A2</td>
<td>Calcium regulated membrane binding protein</td>
</tr>
<tr>
<td>2</td>
<td>CENPF</td>
<td>Centromere protein F</td>
<td>Kinetochore function and chromosome segregation in mitosis</td>
</tr>
<tr>
<td>3</td>
<td>MMP7</td>
<td>Matrix Metalloproteinase 7</td>
<td>Degrades extracellular matrix. Activates procollagenase.</td>
</tr>
<tr>
<td>4</td>
<td>MMP11</td>
<td>Matrix Metalloproteinase 11</td>
<td>Role in progression of epithelial malignancies</td>
</tr>
<tr>
<td>5</td>
<td>MSLN</td>
<td>Mesothelin</td>
<td>Role in cellular adhesion and as a cytokine</td>
</tr>
<tr>
<td>6</td>
<td>MUC5AC</td>
<td>Mucin 5AC</td>
<td>Mucus formation</td>
</tr>
<tr>
<td>7</td>
<td>REG3A</td>
<td>Regenerating islet-derived 3 alpha</td>
<td>May be involved in cell proliferation or differentiation</td>
</tr>
<tr>
<td>8</td>
<td>TFF1</td>
<td>Trefoil factor 1</td>
<td>Mucous stabilizer in gastrointestinal tract</td>
</tr>
<tr>
<td>9</td>
<td>SERPINB5</td>
<td>Serpin peptidase inhibitor clade B, member 5</td>
<td>Tumor suppressor, blocks growth, invasion and metastatic properties of mammary tumors</td>
</tr>
<tr>
<td>10</td>
<td>FSCN1</td>
<td>Fascin homolog 1</td>
<td>Organizes actin into bundles, important for the formation of cellular protrusions</td>
</tr>
<tr>
<td>11</td>
<td>CASP3</td>
<td>Caspase 3</td>
<td>Involved in execution of apoptosis</td>
</tr>
<tr>
<td>12</td>
<td>LCN2</td>
<td>Lipocalin 2</td>
<td>Ion trafficking protein involved in apoptosis, immunity and renal development</td>
</tr>
<tr>
<td>13</td>
<td>SPINK1</td>
<td>Serine peptidase inhibitor Kazal type 1</td>
<td>Trypsin inhibitor</td>
</tr>
<tr>
<td>14</td>
<td>S100P</td>
<td>S100 calcium binding protein P</td>
<td>Regulation of cell cycle progression and differentiation</td>
</tr>
<tr>
<td>15</td>
<td>S100A6</td>
<td>S100 calcium binding protein A6</td>
<td>Involved in cell cycle progression and differentiation</td>
</tr>
<tr>
<td>16</td>
<td>CEACAM6</td>
<td>CEA related cell adhesion molecule 6</td>
<td>Non specific tumor marker</td>
</tr>
<tr>
<td>17</td>
<td>NQO1</td>
<td>NAD(P)H dehydrogenase</td>
<td>Serves as a quinone reductase in biosynthetic and detoxification pathways</td>
</tr>
</tbody>
</table>
3.6. Sequencing of primers and probes

Taqman PCR primers and probes were designed for 17 indicator genes and two housekeeping genes (GAPDH and PSMB6). The sequences of these genes were identified through the Ensembl genome browser and primers and probes were designed using PrimerExpress proprietary software. The assays are based on the Taqman 5’ nuclease end for amplification and detection of the target in cDNA samples. The custom Taqman gene expression assay contains two target-specific primers and one Taqman MGB FAM™ dye labeled probe. The sequences of primers and probes are in the tables below (Table 10 & Table 11).

Table 10: List of primer sequences

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANXA2F</td>
<td>CCGTGGTTCACTGCGATTCAAGAAC</td>
</tr>
<tr>
<td>ANXA2R</td>
<td>CATGGAGTCATACACGCGATCA</td>
</tr>
<tr>
<td>CASP3F</td>
<td>CCTGGTATATATTCTTGCGGAA</td>
</tr>
<tr>
<td>CASP3R</td>
<td>GCACAAAGCGACTGGATGAA</td>
</tr>
<tr>
<td>CEACAM6F</td>
<td>CTGCCACCGACGCTCTTAAC</td>
</tr>
<tr>
<td>CEACAM6R</td>
<td>GGAACGTCCATTGATAAACCA</td>
</tr>
<tr>
<td>CENPFF</td>
<td>CCGAGAACCACCCACACGAA</td>
</tr>
<tr>
<td>CENPFR</td>
<td>CAGTCGGACTTCTTCAGGA</td>
</tr>
<tr>
<td>FSCN1F</td>
<td>GAACCTCAGAGCTCTCCTCATGA</td>
</tr>
<tr>
<td>FSCN1R</td>
<td>GCTCCCCGCAGGAACAC</td>
</tr>
<tr>
<td>LCN2F</td>
<td>GAACTTCAATCCGCTTCCTCAAAA</td>
</tr>
<tr>
<td>LCN2R</td>
<td>ACAGGGAAGACGATGTGGTTTT</td>
</tr>
<tr>
<td>MMP7F</td>
<td>GCTGGCTCATGCTTGGC</td>
</tr>
<tr>
<td>MMP7R</td>
<td>TCCTCATCGAAGTGACATCTC</td>
</tr>
<tr>
<td>MMP11F</td>
<td>TCAGGATGCTGTGCGATGCTATG</td>
</tr>
<tr>
<td>MMP11R</td>
<td>TTAGGCTCTCCACTTCCAGTAG</td>
</tr>
<tr>
<td>MSLNF</td>
<td>CGCTCAGTCACGAGATGTGA</td>
</tr>
<tr>
<td>MSLNR</td>
<td>CCGTCAGGCACTTTCATG</td>
</tr>
<tr>
<td>MUC5ACF</td>
<td>GCAGCCAAAGGTCACCACCATTT</td>
</tr>
<tr>
<td>MUC5ACR</td>
<td>GCTGGTCTGGCGGATGA</td>
</tr>
<tr>
<td>Primer</td>
<td>Sequence</td>
</tr>
<tr>
<td>------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>NQO1F</td>
<td>TGGAAGAAACGCCTGGGAGAA</td>
</tr>
<tr>
<td>NQO1R</td>
<td>GAGGGCTGCTTTGGAGCAAAAT</td>
</tr>
<tr>
<td>REG3AF</td>
<td>TCTGGATTTGGGTCCCATGAC</td>
</tr>
<tr>
<td>REG3AR</td>
<td>CTCCCAACCTTCTCCATTGG</td>
</tr>
<tr>
<td>S100A6F</td>
<td>GCAGGATGCTGAAATTTGCAA</td>
</tr>
<tr>
<td>S100A6R</td>
<td>GTTCACCTCCTGGTGCCCTTGTC</td>
</tr>
<tr>
<td>S100PF</td>
<td>CAGCCATGGGCATGATCA</td>
</tr>
<tr>
<td>S100PR</td>
<td>CCCTCGCTGCCCCGAATAT</td>
</tr>
<tr>
<td>SERPINB5F</td>
<td>TGCAAGCAAGGATGAATTGGA</td>
</tr>
<tr>
<td>SERPINB5R</td>
<td>AGTTTTTGTGTGCCTGTATGTAA</td>
</tr>
<tr>
<td>SPINK1F</td>
<td>CTGTTGGACTGATGGAATACTTATC</td>
</tr>
<tr>
<td>SPINK1R</td>
<td>TGGCGTTTCCGTATTTTCA</td>
</tr>
<tr>
<td>TFF1F</td>
<td>TTTCGACGACACCGTTCGTC</td>
</tr>
<tr>
<td>TFF1R</td>
<td>GGAGGGACGTCGATGGTATTAG</td>
</tr>
</tbody>
</table>

Table 11: List of probe sequences

<table>
<thead>
<tr>
<th>Probe</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANXA2P</td>
<td>AGCCCTGTATTTTG</td>
</tr>
<tr>
<td>BIRC5P</td>
<td>AGAGAGGAAACATAAAAAG</td>
</tr>
<tr>
<td>CASP3P</td>
<td>TCAAAAGGATGGCTCCTG</td>
</tr>
<tr>
<td>CEACAM6P</td>
<td>ACCTGCACAGTACTC</td>
</tr>
<tr>
<td>CENPFP</td>
<td>TCCGTTCCAGTCAT</td>
</tr>
<tr>
<td>FSCN1P</td>
<td>CTCATCAACGCACCCAT</td>
</tr>
<tr>
<td>LCN2P</td>
<td>CTCTGGGCTCCTCC</td>
</tr>
<tr>
<td>MMP7P</td>
<td>CTGGGACTGGTCTC</td>
</tr>
<tr>
<td>MMP11P</td>
<td>CTTTCTGCGGCAGCC</td>
</tr>
<tr>
<td>MSLNP</td>
<td>CATGGACTTTGGCCACG</td>
</tr>
<tr>
<td>MUC5ACP</td>
<td>AGACCCCTCAACCTTCT</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Probe</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>NQO1P</td>
<td>ATTTGGGATGAGACACCA</td>
</tr>
<tr>
<td>POSTNP</td>
<td>TTGCAAGCAACAAAAA</td>
</tr>
<tr>
<td>REG3AP</td>
<td>CCACACAGGCGACCC</td>
</tr>
<tr>
<td>S100A6P</td>
<td>CTGATGGGAAGACTTGGGAC</td>
</tr>
<tr>
<td>S100PP</td>
<td>AGACGTCTTTTCCC</td>
</tr>
<tr>
<td>SERPINB5P</td>
<td>TGCTGACCATCCCTT</td>
</tr>
<tr>
<td>SPINK1P</td>
<td>CAATGAATGCGTGTATGTG</td>
</tr>
<tr>
<td>SPP1P</td>
<td>AGTCAGGAACCTTTC</td>
</tr>
<tr>
<td>TFF1P</td>
<td>CCGTGGTCTTCTAT</td>
</tr>
<tr>
<td>GAPDHP</td>
<td>CTGTAGCCAAATTCG</td>
</tr>
<tr>
<td>PSMB6P</td>
<td>CCCGCTCCCCAG</td>
</tr>
</tbody>
</table>
3.7. Materials and Methods (292)

1) Extraction of RNA and global amplification of polyadenylated mRNAs (poly(A) RT PCR)

Isogen-LS (Wako, Osaka, Japan) was used in combination with RNeasy mini Kit (Qiagen, Crawley, UK) for total RNA extraction and purification from pancreatic juice. Total RNA from pancreatic tissue was extracted following the RNeasy kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Samples were then quantified using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA) analysis and RNA integrity was checked by Agilent 2100 Bioanalyzer (Agilent Technologies, Waldbronn, Germany) according to the manufacturer's instructions. The poly(A) PCR method has three steps, namely poly(A) primed reverse transcription to produce a first-strand cDNA for each mRNA; polyadenylation of the first-strand cDNA, rendering it defined at both ends; and polymerase chain reaction using a poly(dT) primer, which anneals to the poly(A) site on the first-strand cDNA. We performed poly(A) PCR on RNA isolated from pancreatic juice and tumour tissue following standard published protocols.

2) TaqMan real-time quantitative PCR

TaqMan™ PCR primers and probes were designed for 17 Indicator Genes that have been shown to be over-expressed in pancreatic cancer including housekeeping genes (GAPDH and PSMB6) using Custom TaqMan® Gene Expression Assays (Applied Biosystems, Foster City, USA). All the PCR primer pairs were designed for the mRNA sequence within 300 base pairs of the Poly A signal of each indicator gene. The real-time PCR reactions were performed in 96 well optical reaction plates with a final volume of 25µl, containing 12.5µl of TaqMan Universal PCR Mastermix (Applied Biosystems, Foster City, USA), 1.25µl of Custom TaqMan® Gene Expression Assays Primers and probes (Applied Biosystems, Foster City, USA), 8.75µl of UltraPure™ DEPC water (Invitrogen, Carlsbad, USA) and 2.5µl of cDNA. All samples were then analyzed using an ABI Prism 7900HT sequence detection system (Applied Biosystems, Foster City, USA).
3) Real-Time PCR

The results were calculated by the ABI Prism 7900HT. The amplification plots that were generated indicated the quantity of fluorescence generated from each reaction at each PCR cycle. From these amplification plots, Ct values (cycle at which a set threshold for the quantity of reporter fluorescent emission is reached) were obtained for each sample. PCR’s were run in triplicates for each sample and Ct averages obtained.

4) Normalization

Following real-time PCR amplification and quantification of the selected genes, this factor was used for the normalization of expression levels of each of the genes measured. Averages were normalized to housekeeping genes following the $2^{ΔΔCt}$ method (293-295).

### 3.8. Results

#### 3.8.1 RNA Extraction

Total RNA was extracted successfully from all 52 samples including 48 samples of pancreatic juice and 49 samples of tissue.

#### 3.8.2. Pancreatic Juice

The median volume of pancreatic juice obtained by aspiration during surgery was 650μL (range 100μL to 6000μL). The median RNA yield quantified by NanoDrop was 4.25ng/μL (range from 0.85 – 234.61ng/μL, standard deviation ± 46.34). The median purity ratio was 1.72 (range from 0.61 to 2.45, standard deviation ±0.4399) for pancreatic juice extractions.

#### 3.8.3 Pancreatic Tissue

The median RNA yield as quantified by NanoDrop was 354.56ng/μL (range 3.2ng/μL to 4071.13ng/μL, standard deviation ±929.11). The median purity ratio 260/280 was 2.05 (range from 1.64 to 2.56, standard deviation ± 0.124) for total RNA of pancreatic tissue.
3.8.4. Poly A PCR

Poly A cDNA was generated from the mRNA extracted for all the paired samples. The reliability of the methodology was verified by establishing the cDNA quantity using gel electrophoresis. GAPDH gene cDNA PCR amplification was used to display specific cDNA.

3.8.5. Real-time PCR

Gene expression analysis of the 18 indicator genes was performed in 42 directly paired samples from a total of 52 cases with pancreatic juice available for 48 patients and tissue for 49 patients. Ct values were obtained for each gene in each paired sample type.

3.8.5.1. Gene expression levels in all sample types

No gene expression was seen in the negative control samples. Samples were then normalized for the housekeeping gene in each sample. The figure below illustrates the median gene expression levels (median ΔCt values in logarithmic scale) of all the genes assessed. (Figure 32).

Figure 32: This figure illustrates the median gene expression values (Ct values in log scale, after normalisation for GAPDH) in the entire cohort of pancreatic juice and tissue samples.
The median gene expression values after normalization to the housekeeper gene for all sample types demonstrates corresponding expression levels in both pancreatic juice and pancreatic tissue. All the genes with the exception of SPINK1 were over-expressed. Levels of gene expression in pancreatic juice matched that of tissue in all but 2 genes, LCN2 and TFF1. In fact, gene expression in pancreatic juice was higher than that of tissue in eight of the genes assessed. All but four genes showed a greater than 5-fold expression in tissue samples, indicating that the gene selection was appropriate despite the heterogeneous nature of the group.

3.8.5.2. Gene expression levels in pancreas cancer samples

Following gene expression analysis, the sample types were compared separately based on tissue histology subtype (pancreas cancer and pancreatic cystic tumours). The median gene expression values for paired pancreatic juice and tissue samples from only patients with pancreas cancer are shown in the figure below. (Figure 33).

![Figure 33](image)

**Figure 33:** This graph describes the median gene expression values (Ct values in log scale after normalisation for GAPDH) in paired pancreatic juice and tissue samples in pancreas cancer.

The graph illustrates that the overall gene expression levels in both pancreatic juice and tissue samples was broadly similar with all but three genes showing a greater than 5-fold expression in
pancreas cancer tumour tissues. In general, the overall levels of expression in pancreas juice and tumour tissue appeared to correspond. Levels of gene expression in pancreatic juice samples matched those of the tissue samples in pancreas cancer, with significantly lower expression seen in TFF1 and LCN2, with LCN2 appearing under-expressed in pancreatic juice. In a previous study comparing levels of LCN2 in pancreatic juice, a differential expression had also been noted, with levels of LCN2 in juice not corresponding with those from tumour tissue (290). Lower levels of expression in juice were also seen in the genes S100A6 and S100P, corresponding to their overall expression as seen in all sample types. The only feature of note was that the relative gene expression level in the pancreatic juice samples compared with that of the tumour tissue was slightly lower in the pancreas cancer group compared with those from the entire group.

3.8.5.3. Gene expression levels in the IPMN samples

The median gene expression values for paired pancreatic juice and tissue samples from only patients with IPMN are shown in the figure below. (Figure 34).

Figure 34: The figure illustrates the median gene expression values in paired pancreatic juice and tissue samples from patients with IPMN

In the IPMN group, gene expression levels were again broadly similar for both the tissue types. There was greater expression of the mucin gene MUC5AC in pancreatic juice compared to
tumour tissue as would be expected with mucinous tumours. Gene expression levels for LCN2 and SPINK1 were consistently low across all the samples. SPINK1 was the only gene appearing as under-expressed in pancreatic juice while showing poor tissue expression. Expression of the TFF1 gene in pancreatic juice was also consistently low across the samples. Overall, expression levels in IPMN tumour tissue were higher than that of pancreatic juice in nine of the seventeen genes, this could possibly be a reflection of the fact that six of the eight samples were benign IPMNs.

3.8.5.4. Comparison of gene expression levels between pancreatic juice and pancreatic tissue samples in pancreatic cancer

In order to test the preliminary hypothesis that genes over-expressed in pancreatic tumour tissue in patients with pancreas cancer would be over-expressed in pancreatic juice, a Spearman’s correlation was performed in these paired samples in patients who had undergone surgery for pancreas cancer.

The correlation coefficients are shown below at Table 12.
Table 12: Spearman’s correlation coefficients in paired samples from patients with pancreas cancer

<table>
<thead>
<tr>
<th>Gene</th>
<th>Correlation Coefficient</th>
<th>Sig (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANXA2</td>
<td>-0.42</td>
<td>0.886</td>
</tr>
<tr>
<td>CASP3</td>
<td>-0.275</td>
<td>0.342</td>
</tr>
<tr>
<td>CEACAM6</td>
<td>0.046</td>
<td>0.876</td>
</tr>
<tr>
<td>CENPF</td>
<td>-0.22</td>
<td>0.45</td>
</tr>
<tr>
<td>FSCN1</td>
<td>0.464</td>
<td>0.095</td>
</tr>
<tr>
<td>LCN2</td>
<td>0.222</td>
<td>0.446</td>
</tr>
<tr>
<td>MMP7</td>
<td>-0.9</td>
<td>0.759</td>
</tr>
<tr>
<td>MMP11</td>
<td>0.029</td>
<td>0.923</td>
</tr>
<tr>
<td>MUC5AC</td>
<td>0.226</td>
<td>0.436</td>
</tr>
<tr>
<td>MSLN</td>
<td>0.679</td>
<td>0.008</td>
</tr>
<tr>
<td>NQO1</td>
<td>0.125</td>
<td>0.67</td>
</tr>
<tr>
<td>REG3A</td>
<td>0.121</td>
<td>0.681</td>
</tr>
<tr>
<td>S100A6</td>
<td>-0.323</td>
<td>0.26</td>
</tr>
<tr>
<td>S100P</td>
<td>-0.275</td>
<td>0.342</td>
</tr>
<tr>
<td>SERPINB5</td>
<td>0.305</td>
<td>0.288</td>
</tr>
<tr>
<td>SPINK1</td>
<td>0.112</td>
<td>0.706</td>
</tr>
<tr>
<td>TFF1</td>
<td>0.108</td>
<td>0.714</td>
</tr>
</tbody>
</table>

The expression levels between paired pancreatic juice and tissue samples in pancreas cancer revealed p values less than 0.05 in only one gene, MSLN. MSLN is a gene that plays a role in cell adhesion and encodes a pre-protein that divides into two protein products and has previously been demonstrated to be over-expressed in pancreas cancer.

The significance of this result depends on it being as a result of multiple comparisons, and therefore to correct this, a Bonferroni correction was used. With no correction the chance of finding one or more significant differences was 0.1276 (12.76%). To adjust with no correlation for an overall alpha level of 0.008, the Bonferroni adjustment for each test indicated lowering the significance level to 0.0004706. This implies that the result was probably not significant.
3.8.5.5. Comparison of gene expression levels between paired samples in pancreatic cystic tumours

We also tested for genes that were over-expressed in pancreatic cystic tumour tissue in patients undergoing surgery for suspected malignancy would be over-expressed in pancreatic juice. A Spearman's correlation was performed in these paired samples in patients and is depicted below Table 13.

Table 13: Comparison of gene expression levels between paired samples in pancreatic cystic tumours

<table>
<thead>
<tr>
<th>Gene</th>
<th>Correlation Coefficient</th>
<th>Sig (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANXA2</td>
<td>0.543</td>
<td>0.266</td>
</tr>
<tr>
<td>CASP3</td>
<td>-0.314</td>
<td>0.544</td>
</tr>
<tr>
<td>CEACAM6</td>
<td>0.257</td>
<td>0.653</td>
</tr>
<tr>
<td>CENPF</td>
<td>0.2</td>
<td>0.704</td>
</tr>
<tr>
<td>FSCN1</td>
<td>-0.257</td>
<td>0.623</td>
</tr>
<tr>
<td>LCN2</td>
<td>0.714</td>
<td>0.111</td>
</tr>
<tr>
<td>MMP7</td>
<td>0.886</td>
<td>0.019</td>
</tr>
<tr>
<td>MMP11</td>
<td>0.2</td>
<td>0.704</td>
</tr>
<tr>
<td>MUC5AC</td>
<td>0.086</td>
<td>0.872</td>
</tr>
<tr>
<td>MSLN</td>
<td>0.6</td>
<td>0.208</td>
</tr>
<tr>
<td>NQO1</td>
<td>-0.029</td>
<td>0.957</td>
</tr>
<tr>
<td>REG3A</td>
<td>-0.486</td>
<td>0.329</td>
</tr>
<tr>
<td>S100A6</td>
<td>0.314</td>
<td>0.544</td>
</tr>
<tr>
<td>S100P</td>
<td>0.486</td>
<td>0.329</td>
</tr>
<tr>
<td>SERPINB5</td>
<td>0.257</td>
<td>0.623</td>
</tr>
<tr>
<td>SPINK1</td>
<td>-0.371</td>
<td>0.468</td>
</tr>
<tr>
<td>TFF1</td>
<td>-0.314</td>
<td>0.544</td>
</tr>
</tbody>
</table>

The expression levels between paired pancreatic juice and tissue samples in pancreatic cystic tumours revealed p values less than 0.05 in only one gene, MMP7. The MMP7 gene encodes for matrix metalloproteases. These are zinc dependent endo-peptidases that are the major proteases
involved in ECM degradation. MMPs are capable of degrading a wide range of extracellular molecules and a number of bioactive molecules. MMPs are implicated in destruction of the basement membrane with implications for tumour metastasis. The significance of this result depends on it being as a result of multiple comparisons, and therefore to correct this, a Bonferroni correction was used. With no correction the chance of finding one or more significant differences was 0.1571 (15.71%). A Bonferroni adjustment with no correlation for an overall alpha level of 0.01 indicated lowering the significance level to 0.00058. This implies that the result was probably not significant.

3.9. Discussion

The progressive identification of gene signatures through the use of gene expression technology has increased pressure to utilize these advances in clinical practice to improve outcomes in cancer. This has particular emphasis in patients with difficult-to-diagnose diseases like pancreas cancer. This also applies to precursors of pancreas cancer with a high risk; such as pancreatic cystic tumours, more specifically, intraductal papillary mucinous neoplasms and mucinous cystic neoplasms. Studies have looked at identifying markers in blood samples to facilitate diagnosis with limited success (296, 297). In this study pancreatic juice was specifically chosen based on the premise that malignant cells shed from the primary tumour into pancreatic juice could be used to identify over-expressed genes within the primary tumour thereby providing evidence of malignant change. In addition, pancreatic juice is available during endoscopic assessment of pancreatic lesions in a significant percentage of patients and is a feasible target sample.

The subject population included patients who underwent pancreatic resections for suspected pancreas cancer, pancreatic cystic lesions or peri-ampullary cancer. However, results were analyzed for only pancreas cancer and pancreatic cystic tumours as the indicator genes chosen were based on over-expressed genes within this subgroup and to minimize confounding results. The aetiology of pancreas cancer originates through two pathways, either as a progression from pancreatic intra-epithelial neoplasia or via a secondary pathway involving the development of cystic lesions that gradually progress to cancer (42). This study aimed to assess if there was a correlation between over-expressed genes in pancreatic juice and tumour in patients with pancreas cancer and pancreatic cystic tumours.
3.9.1. Methodology

The collection of pancreatic juice and tissue samples was an essential prerequisite for this study and there were limitations associated with the technique of collection. Our group has previously assessed and validated this technique (290). All attempts were made to ensure that pancreatic juice samples were free of blood at collection, which can act as a confounder substrate. Nonetheless, it cannot be completely discounted that microscopic volumes of blood from the opened pancreatic duct could have been mixed with the collected specimen. In addition, the volume of pancreatic juice required depends on a degree of pancreatic ductal obstruction, which may not always be present. We were able to obtain sufficient samples volumes through the use of a small calibre infant feeding tube left to drain freely for a few minutes. This did carry the drawback that the tube may have bypassed the tumour or that shed cells may not have been collected. The pancreatic juice was necessarily collected from the duct distal to the tumour due to intraoperative technical considerations. Consideration must be given to the possibility that, as the pancreatic juice obtained was downstream of the tumour tissue, there may not have been sufficient tumour cells shed to enable their detection. Another confounder is the presence of nucleases within pancreatic juice. Pancreatic juice is a difficult substrate for the measurement of mRNA due to the presence of powerful nucleases. To minimise nuclease-mediated destruction of mRNA prior to storage, the collected pancreatic juice was immediately stored in a nuclease inhibitory solution and then put on ice. However, RNA degradation could possibly occur during the operative process with the lag time for degradation beginning at the onset of pancreatic dissection and manipulation right through to sample collection. The RNA was quantified using UV absorption measured by a spectrophotometer. Ideally, scanning spectrophotometry should be used as this makes it possible to also identify possible sources of contamination. One of the problems using conventional spectrophotometers is that the cuvettes are large, making it difficult to measure low concentrations of RNA without loosing an unacceptable fraction of the sometimes precious and valuable RNA sample. In this study, the RNA absorption spectra were assessed for each sample to ensure maximal purity.

The collection of pancreatic tumour tissue was limited by our proposed use of this technique as a clinical tool. We elected to utilize fresh tumour tissue as our comparator sample in order to facilitate use in clinical practice based on current processing protocols. It would perhaps have
been easier to use micro-dissected tissue samples, which would have enhanced the accuracy of
detection of our indicator genes, however, this would have significantly added to both processing
times, cost and deviated from current tissue processing. Our gene selection was also based on
genes over-expressed in a combination of parenchymal, stromal and precursor lesions allowing
for a wider pick-up of genes involved in pancreatic carcinogenesis though we had to limit the
number to 18 genes due to cost constraints. We appreciate that there are a wide variety of over-
expressed genes implicated in pancreas cancer, however, not all are expressed in pancreatic
juice with most studies focused on microarrays on tissue samples (74, 78). However, genes were
chosen based on their expression in both substrates, with documented high expression levels
and based on our previous analysis with the highest likelihood of correlation. Pancreatic juice has
high levels of RNAse’s (298), with difficulties in maintaining standardization and RNA integrity.
We used an RNAse inhibitor that our group had previously validated in comparison with other
methodologies (292). The purity of RNA directly correlates with the accuracy of gene expression
and we measured this using A260/280 ratio using Nanodrop software (Nanodrop technologies,
Delaware, USA). The RNA quality was assessed using the RNA integrity number (Agilent
technologies), based on studies demonstrating that good quality RNA is essential prior to
quantitative RT PCR. We chose house keeping genes (GAPDH, PSMB6) previously been shown
to have consistent results based on work our group has previously performed with this
methodology (292).

3.9.2. Choice of methodology (Poly A PCR)

The cDNA analysis of discrete selected marker genes, previously identified by DNA microarray
studies, permits the ability to identify low levels of multiple genes in substrates with very small cell
numbers. In addition, it is relatively inexpensive as compared to large-scale microarray studies.
Limitations of microarrays include relatively low detection thresholds for individual genes (299).
The technique of Poly A PCR permits a rapid assessment (over 48 hours), provisionally enabling
its use in clinical settings. Also, once Poly A cDNA is available, it can essentially be stored
permanently; with implications for further testing in the event of future genetic advances.
3.9.3. Gene expression between pancreatic juice and tumour tissue in pancreas cancer

We compared our set of 18 marker genes to paired samples obtained in the same patient with pancreas cancer undergoing surgery. In previous work, our group looked at mRNA expression levels in pancreatic tissue and juice samples comparing a range of peri-ampullary cancers versus benign samples (291). Peri-ampullary malignancies are known to originate from the same endoderm and have been shown to share behavioral patterns (300, 301). Previous work has demonstrated that despite non-significant results, the expression levels of several genes (EGFR, ANXA1, MSLN and PLAT) showed p values less than 0.05 prior to correction for false discovery rate based on the use of multiple samples. This study aimed to clarify this relationship further by specifically looking at patients with pancreas cancer. We chose genes based on an analysis of papers looking specifically at genes that would be over-expressed in pancreatic tissue, juice and in precursor lesions. This allowed us to use the same gene set in our second group (those with pancreatic cystic tumours) and assess if these genes could be used in the assessment of this patient subset. In this study, the overall gene expression levels in the entire sample cohort including pancreas cancer, other cancer types and IPMN and chronic pancreatitis showed that the overall gene expression profiles matched between pancreatic juice and pancreatic tissue. There was some difference in the levels of genes in juice and tissue between the entire cohort as compared to those from pancreas cancer and IPMN. In part, the variety of tissue types could account for this. Some genes, notably, LCN2 and TFF1 showed poor expression in the juice samples across the cohort, while SPINK1 was under-expressed in the group. This corroborates with previous findings published comparing levels of LCN2 and SPINK1 between juice and tissue (290). Levels of expression of the remaining cohort of genes in pancreatic tissue and juice are also similar to those previously published in a compendium of expressed genes (78).

The gene expression levels in the pancreas cancer cohort showed that in eight of the genes, expression levels in pancreatic juice were higher than those of tissue. In the solitary under-expressed gene (SPINK1), levels of the gene in juice were also higher than that of tissue. Levels of S100P and S100A6 are both members of a family of calcium binding protein genes and levels of the genes often correlate in vitro (302). Low levels of expression of both these genes were observed in the pancreatic juice samples in the pancreas cancer group. In general, the relative concordance between gene levels in juice and tumour tissue meant that these genes could be
used to assess for a correlation between the two substrates. This study demonstrated that one gene out of seventeen; MSLN (mesothelin) was significantly over-expressed in both pancreatic juice and tissue samples in the cohort of patients with pancreatic cancer, though it did not meet the cutoff set by the Bonferroni correction. Previously, we had noted over-expression of mesothelin in hepatobiliary cancers compared to benign paired samples. Mesothelin is a gene previously well documented to be over-expressed in pancreas cancer, mesotheliomas and ovarian cancer. The gene encodes a precursor protein that encodes for two products, megakaryocyte-potentiating factor and mesothelin. Megakaryocyte-potentiation factor can stimulate colony formation in bone marrow megakaryocytes, while mesothelin is a cell-surface protein that may function as a cell adhesion protein (303). Despite this the exact role of MSLN in carcinogenesis is still unknown.

3.9.4. Gene expression between pancreatic juice and tumour tissue in pancreatic cystic tumours

Gene expression levels in the IPMN group show similar levels of gene expression within the pancreatic juice and tissue. We compared our set of 18 marker genes to paired pancreatic juice and tumour samples obtained in the same patient group with pancreatic cystic tumours undergoing operative intervention. To date, no studies that have compared the gene expression levels between the sample types to assess for indicative biomarkers. We demonstrated that one gene out of seventeen; MMP-7 was significantly over-expressed in both pancreatic juice and tissue samples in the cohort of patients with pancreatic cystic tumours. On correction for multiple comparisons, this result did not meet the level of significance. However, MMP-7 has previously been suggested as a biomarker for both pancreatic cancer and IPMN in pancreatic juice samples (304, 305). Thus, this result may proffer some indication as to the possibility of identifying biomarkers in pancreatic juice in patients with pancreatic cystic tumours but further studies would be required to validate this.

3.10. Conclusion

Our preliminary hypothesis was based on the precept that pancreatic juice shares close contact with pancreatic ductal adenocarcinoma cells leading to cell shedding into the juice. This would
permit the measurement of over-expressed genes within these shed cells to facilitate a
correlation with the underlying cancer. The low yield of statistically significant results comparing
the over-expression of this carefully chosen set of genes in pancreas cancer may be related to
several issues. The caveat being that there is no guarantee that pancreatic juice from cancer
contains the shed cells. This would necessarily be dependent on the sampling technique.
Additionally, RNA in pancreatic juice could as well originate from ductal epithelium as from
stromal cells, normal pancreas or an inflammatory response to the tumour. A limitation of our
technique was that pancreatic juice was harvested from intra-operatively cut pancreas, thereby
introducing an element of increased risk of non tumoural RNA within the sample. Pancreatic juice
itself has powerful endonucleases and we are uncertain as to how this may affect certain
transcripts in a sample type with low cellularity. We have tried to minimize any effect of the
RNAse by using a previously validated RNAse inhibitor in our study. Lastly, surgery itself may
play a role in activation of RNAse activity.

A further enquiry based on the efficacy of the chosen subset of genes could be raised. It would be
ideal to have explored the samples by testing a larger subset of genes but we were exposed to
limitations of cost and resources. As the focus of this study was to assess the feasibility of using
this technique as a clinical tool, it was felt that our gene selection would be sufficient for purpose.

Based on previous work and this study, there is no clear evidence that there is a consistent
detection of over-expressed genes in pancreatic juice compared with that of pancreatic cancer
tissue. A previous study demonstrated non-significant over-expression of ANXA2 in pancreas
cancer within these paired sample types and we have shown MSLN to be over-expressed in
pancreatic juice in pancreas cancer. Despite these positive results, and in view of the restrictions
due to properties of the substrate (RNAse, poor cellular material) and uncertainty of the origin of
the mRNA isolated, we cannot justify the regular use of pancreatic juice as a routine predictor of
pancreas cancer. Pancreatic juice continues to be an encouraging substrate for investigation,
however, the issues highlighted above need to be resolved prior to justifying its use in clinical
practice.
Chapter 4 - Metabolomic Studies in Pancreatic Cancer

4.1. Introduction

Metabolomics is the comprehensive analysis of the low molecular weight molecules that are intermediary or end products of metabolism (306). Metabolites are generally regarded as being less than 1500 Daltons in weight. The goal of metabolic analysis is to extract, identify and quantify all the metabolites in a given biological sample. The metabolome of a cell is the collection of downstream products of gene transcription, translation and post-translational protein modification and may be said to represent the true cellular phenotype at a given time. Thus, the biochemical phenotype of an organism is the final result of genetic and environmental interactions and fluctuations in metabolites represent dynamic alterations in biochemical pathways, in response to physiological change.

Metabolomic studies can be used to investigate the relationship between changes in the endogenous metabolite concentration versus that of the disease. Factors such as age, sex, body mass index, alcohol, smoking and dietary factors influence the metabolite composition in any individual and therefore any comparison between normal and diseased individuals requires standardization of these elements. Metabolic analysis comparison with adjusted normal controls may also assist in identifying metabolic pathways altered in the progression to cancer. Therefore, metabolites are suitable to be studied as possible markers during an altered metabolic state induced by diseases such as cancer. They have been variously studied in breast (307), ovarian (308) and prostate cancer (309) and their potential use in pancreatic cancer (310) is being explored.

This chapter will review the current literature on the use of metabolomics as a tool in identifying possible biomarkers in pancreas cancer and evaluate its potential in a prospective cohort of patients with pancreas cancer using liquid chromatography mass spectrometry and discuss future directions and applications in clinical practice. The aims of this chapter are threefold: 1. Review the current literature on the use of metabolomics as a tool in identifying biomarkers in pancreas cancer. 2. Evaluate the potential of this technique in a cohort of patients with pancreas cancer,
ampullary cancer and endocrine cancer and 3. Identify applications in clinical practice and discuss future directions.

4.2. Approaches to investigating the metabolome

4.2.1 Background

It is important to understand the complex relationships integrating the changes occurring from the genetic command, through the protein products, to the resultant metabolic interactions that underpin cellular function. This provides us with a holistic picture of the mechanics behind the external phenotype and is important in enhancing our understanding of cellular dysfunction. In addition, this has potential applications in diseases involving metabolic alterations, cancer diagnosis, pharmaceutical development and understanding of cellular biochemical processes.

It is estimated that the human metabolome comprises of 2700 metabolites (311) compared with the 31,897 genes and one million different proteins estimated to result from gene expression, alternative splicing and post-translational modifications (312, 313). This makes measuring the metabolome and integrating metabolomic information with mRNA and protein expression more attractive in order to visualize a comprehensive snapshot of the controlling machinery of an organism.

4.2.2. Overview of laboratory techniques

A vast difference in the chemical properties of metabolites makes it difficult to assess the entire range of the metabolome using a single analytical method. Therefore, the metabolome is investigated using several different approaches (314).

These include:

1) Metabolomic fingerprinting: This technique uses pattern recognition to provide a biochemical snapshot of a biological tissue or fluid. Variations in the global metabolite composition are detected, measured and correlated with disease-induced changes. This technique functions as a global, high throughput and rapid analysis to classify samples. It can also be used as a
screening tool to discriminate between biological samples of differing origin such as case/control or diseased versus healthy.

2) Metabolic profiling: This is the quantification of a group of related compounds or metabolites within a given metabolic pathway, for example the quantification of lipid pathways in lipidomics (315).

3) Metabolic target analysis: This technique aims to quantify a small number of known metabolites, often related to a specific metabolic reaction (316).

4) Metabolic footprinting: Analysis of the metabolites secreted or excreted by an organism. This includes the environmental and growth substances produced if the organism is being cultured (317).

4.3. Review of metabolomic studies in pancreas cancer

To date, seventeen studies have performed a metabolomic analysis on pancreatic cancer (310, 318-332). Studies were variously performed on serum (321-326) plasma (320) saliva (310) tissues (318, 319), bile (327) and urine (328, 329) from patients with pancreatic cancer. Four studies have focused on animal models (330-333). These are summarized in Table 14.

Table 14: Summary of metabolomic studies in pancreatic cancer

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample</th>
<th>Method</th>
<th>Samples</th>
<th>Metabolites</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kaur (318)</td>
<td>Tissue</td>
<td>UPLC-MS</td>
<td>5 cancers + normal tissue</td>
<td><strong>Increased:</strong> Taurine <strong>Decreased:</strong> succinate, maleate, uridine, malic acid, 5-UMP, 5-AMP</td>
</tr>
<tr>
<td>Cho (319)</td>
<td>In-vivo tissue</td>
<td>1H NMR</td>
<td>21 cancer 15 CP</td>
<td><strong>Increased</strong> lipids in PDAC versus chronic pancreatitis</td>
</tr>
<tr>
<td>Urayama (320)</td>
<td>Plasma</td>
<td>GC-TOF-MS, RP LC/ESI-MS, HILIC/ESI-MS</td>
<td>5 cancer 5 Benign</td>
<td><strong>Increased:</strong> Arachidonic acid, Lyso PC, phosphocholine, phosphoethanolamine, taurocholic acid, cholyglycine <strong>Decreased:</strong> Glutamine, lysine, phenylalanine</td>
</tr>
<tr>
<td>Study</td>
<td>Sample</td>
<td>Method</td>
<td>Samples</td>
<td>Metabolites</td>
</tr>
<tr>
<td>---------------</td>
<td>--------</td>
<td>------------------</td>
<td>------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Ouyang (321)</td>
<td>Serum</td>
<td>1H NMR</td>
<td>17 cancer 23</td>
<td>Increased: Isoleucine, leucine, creatine, triglyceride</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>controls</td>
<td>Decreased: 3-hydroxybutyrate, 3-hydroxyisovalerate, lactate, trimethylamine-N-oxide</td>
</tr>
<tr>
<td>Tesiram (322)</td>
<td>Serum</td>
<td>1H NMR</td>
<td>14 cancer 12</td>
<td>Increased: Total choline, taurine, glucose, unidentified complex matter</td>
</tr>
<tr>
<td>Bathe (323)</td>
<td>Serum</td>
<td>1H NMR</td>
<td>43 cancer 41</td>
<td>Increased: Glutamate, acetone, 3-hydroxybutyrate</td>
</tr>
<tr>
<td>Nishiumi (324)</td>
<td>Serum</td>
<td>GC-MS</td>
<td>20 cancer 9</td>
<td>Increased: lactic acid, thiodiglycolic acid, aconitic acid, 7-hydroxooctanoic acid, aspargine, homogentisic acid, N-acetyl tyrosine</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>controls</td>
<td>Decreased: L-glycine, stearic acid, octanoic acid, glyceric acid, uric acid, lauric acid, decanoic acid, margaric acid, myristic acid</td>
</tr>
<tr>
<td>Leichtle (325)</td>
<td>Serum</td>
<td>Tandem mass</td>
<td>40 cancer 40</td>
<td>Combined amino acid metabolite panel comprising 22 out of 26 amino acids showed superior selectivity when assessed against Ca19-9 for discrimination between the three groups.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>spectrometry</td>
<td>controls 23</td>
<td></td>
</tr>
<tr>
<td>Ritchie (326)</td>
<td>Serum</td>
<td>FI-FTCR-MS</td>
<td>40 cancer 40</td>
<td>Decreased: serum metabolite levels of five systems in cancers compared to controls including 36-carbon ultra long-chain fatty acids, multiple choline-related systems including PC, L-PC and sphingomyelins, as well as vinyl ether containing plasmalogen ethanolamines.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FI-MS/MS</td>
<td>controls 23</td>
<td>Markers: ROC-AUCs based on FI-MS/MS of selected markers from each system ranged between 0.93 ±0.03 and 0.97 ±0.02. Biomarker PC-594 (an ultra long-chain fatty acid), was further validated</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>IPMN 14</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>control 1000</td>
<td></td>
</tr>
<tr>
<td>Sugimoto (310)</td>
<td>Saliva</td>
<td>CE-TOF MS</td>
<td>18 cancer 87</td>
<td>Increased: Leucine, isoleucine, tryptophan, valine, glutamic acid, phenylalanine, glutamine and aspartic acid.</td>
</tr>
<tr>
<td>Bezabeh (327)</td>
<td>Bile</td>
<td>1 H NMR 32P NMR</td>
<td>4 cancer 13</td>
<td>Increased: D-glucuronic acid</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>other</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Sample</td>
<td>Method</td>
<td>Samples</td>
<td>Metabolites</td>
</tr>
<tr>
<td>---------------</td>
<td>------------</td>
<td>-----------------</td>
<td>--------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| Davis (328)   | Urine      | 1H NMR          | 32 cancer 32 controls 25 benign       | **Increased:** Acetone, Hypoxanthine, O-acetylcarnitine, Dimethylamine, Choline, 1-Methylnicotinamide, Threonine, Fucose, Cis-Aconitate, 4-Pyridoxate, Trimethylamine-N-oxide, Aminobutyrate, Tryptophan, Trans-aconitate, 4-Hydroxyphenylacetate, Taurine  
**Decreased:** Trigonelline |
| Napoli (329)  | Urine      | 1H NMR          | Male only ages 62 +/- 6 33 cancers 54 controls | **Increased:** acetoacetate, leucine, glucose, 2-phenylacetamide, and some acetylated compounds.  
**Decreased:** citrate, creatinine, glycine, hippurate, 3-hydroxyisovalerate, trigonelline, and an unknown signal at 3.72 ppm |
| Fang (330)    | Animal tissue | 1H NMR        | 5 cancer 5 CP 10 controls            | **Increased:** leucine, isoleucine, taurine, valine, lactate, alanine, lipid content.  
**Decreased:** phosphocholine, glycerophosphocholine, betaine |
| Kaplan (331)  | Perfused cells | 32 P MRS      | 24 hamsters                          | **Increased** precursors of phospholipid synthesis: phosphomonoesters |
|               | Solid tissue | 32 P MRS       | 60 rats                              | **Increased** phosphomonoesters and phosphodiesters |
|               | Solid tissue | 1H MRS         |                                      | **Increased:** Taurine, lactate, alanine, glutamic acid, glycine  
**Decreased:** glutamine, No significant difference in phospholipid precursors |
| He (332)      | Nude mouse xenograft model using SW 1990 | 1H NMR and PCA analysis | Control nude mice SW1990 cancer mice Cancer mice following radiotherapy | **Increased in PC:** levels of choline, taurine, alanine, isoleucine, leucine, valine, lactate, glutamic acid  
**Decreased:** phosphocholine, glycerophosphocholine, betaine  
**Decreased post radiotherapy:** levels of choline and betaine  
**Increased post radiotherapy:** Acetic acid |
| Yabushita (333) | Transgenic rat model | GC MS | 6 cancers 4 controls | **Decreased in both tissue and serum:** Palmitelaidic acid |

Legend 1: NMR- nuclear magnetic resonance spectroscopy; UPLC-MS- Ultrahigh performance liquid chromatography mass spectrometry; GC-TOF-MS - gas chromatography - time of flight-mass spectroscopy; RP-reversed phase; ESI - electrospray ionisation mass; CE- capillary
electrophoresis; PCA – principal component analysis; FI- FTICR- flow injection, Fourier transform ion cyclotron resonance MS; HILIC – hydrophilic interaction chromatography.

4.3.1. Profiling of pancreatic tumor tissue

4.3.1.1 Study using UPLC/ESI-MS

Kaur et al (318) reported the metabolomic profiling of matched human pancreatic tumor and normal tissue in five patient samples using ultra performance liquid chromatography (UPLC) coupled with electrospray ionization mass spectrometry (ESI-MS). They performed a multivariate data analysis with the identification of metabolites through an accurate mass based database search followed by marker validation using tandem mass spectrometry and relative quantitation. On multivariate analysis, the significantly altered metabolites in pancreatic tissue, as compared to normal tissue, were found to be succinate, taurine, maleate, uridine, glutathione, nicotinamide adenine dinucleotide (NAD), uridine monophosphate (UMP), adenosine monophosphate (AMP) and UDP-N-acetyl-glucosamine. The authors reflected that the down-regulation of citric acid cycle intermediates succinate and malate may have an overall impact on the energy metabolism of the cell, while lower levels of uridine, 5-Uridine monophosphate (5-UMP), and 5-adenosine monophosphate (5-AMP) could reflect the rapid turnover of these nucleotides in the tumor tissue. In addition, low levels of glutathione and NAD were also seen in tumor tissue relative to normal tissue. Taurine was found to be elevated in tumor tissues. These changes in tumor tissue may reflect an increase in cellular metabolic activity and could possibly be used to enhance our understanding of the cellular mechanics in pancreatic cancer cells.

4.3.1.2. In-vivo NMR analysis: to differentiate between pancreatic cancer and chronic pancreatitis.

In the second study on pancreatic tumour tissues, Cho and colleagues (319) aimed to identify the differences between the in vivo proton magnetic resonance spectroscopy (1H-MRS) features of chronic focal pancreatitis and pancreatic carcinoma. The 1H-MR spectra from 36 human pancreases were evaluated in vivo and included 15 cases of chronic focal pancreatitis and 21 cases of pancreatic carcinoma. All cases were confirmed histologically after surgery. The ratios of the peak area (P) of all peaks at 1.6–4.1 ppm to that of lipid peaks (0.9–1.6 ppm) (P [1.6–4.1
ppm)/P [0.9–1.6 ppm]) in the chronic focal pancreatitis and pancreatic carcinoma groups were evaluated, and the results were compared.

They found that in vivo 1H-MR spectra showed significantly fewer lipids in chronic focal pancreatitis than in pancreatic carcinoma. The ratio of P (1.6–4.1 ppm)/P (0.9–1.6 ppm) in chronic focal pancreatitis was significantly higher than that in pancreatic carcinoma because of a decreased peak area of lipids. Using a value of 2.5 as positive for pancreatic cancer, the sensitivity and the specificity for pancreatic cancer was 100% and 53.3%, respectively.

4.3.2. Plasma profiling

In the only study that used plasma as a substrate for metabolite extraction, Urayama et al (320) performed mass spectrometry based metabolic profiling of plasma samples from five histologically proven cases of early-stage pancreatic cancer and five noncancerous cases [3 chronic pancreatitis, 2 normal] to seek novel metabolic biomarkers of pancreatic cancer. A combination of techniques was used to permit the analysis of various metabolites with different polarity and molecular weight range. They generated datasets using gas chromatography/time-of-flight mass spectrometry (GC/TOF-MS), reversed-phase liquid chromatography/electrospray ionization mass spectrometry (RP-LC/ESI-MS) and hydrophilic interaction chromatography (HILIC)-LC/ESI-MS based techniques. They found that selective amino acids, bile acids, and polar lipids were detected with increased or decreased levels in pancreatic cancer samples compared to controls. Feature selection to select the best cancer predictors demonstrated that certain lipid levels were significantly elevated in this group of patients. It has been suggested that the growth-promoting effect of lipids on pancreatic cancer cells occurs at multiple levels including transduction of signals induced by hormones, structural role on cell membranes, and as energy supply (334) The candidate biomarkers included amino acids such as N-methylalanine, lysine, glutamine, phenylalanine, fatty acids such as arachidonic acid, lipids such as lysoPC (18:2), PC (34:2), PE (26:0), and bile acids such as tauroursodeoxycholic acid, taurocholic acid, deoxycholylglycine and cholyglycine. A low level of glutathione found in this study could reflect the fact that tumors show increased glycolysis and glutaminolysis.
4.3.3. Serum profiling

4.3.3.1. Studies with NMR generated profiles.

Three studies (321-323) used 1H nuclear magnetic resonance (NMR) generated profiles to characterize pancreatic cancer. Ouyang et al (321) used 1H NMR metabolomic profiles to discriminate pancreatic cancer patients from healthy controls using serum samples in seventeen pancreatic cancer patients and twenty-three controls. The metabolic analysis revealed significantly lower levels of 3-hydroxybutyrate, 3-hydroxyisovalerate, lactate, and trimethylamine-N-oxide as well as significant higher levels of isoleucine, triglyceride, leucine, and creatinine in the serum from pancreatic cancer patients compared to that of healthy controls. They extrapolated that the subtle differences in metabolite profiles in serum of pancreatic cancer patients and that of healthy subjects could be identified by NMR-based metabolomics and exploited as metabolic markers for the early detection of pancreatic cancer. There are a few criticisms of this study. The tumour stage and pathology were not detailed, precluding an accurate assessment of the cancers. The study population characteristics were not described, leading no information to determine if there were any confounders such as sex, body mass index, diabetes and jaundice. In addition, despite showing that there were significant differences in the metabolite profiles between serum samples and pancreas cancer, there was no evidence to suggest that this may actually help in the early diagnosis of pancreas cancer. At best, this study suggests a panel of metabolite biomarkers may help distinguish between established pancreas cancer and healthy patients. Tesiram et al (322) also used 1H NMR and combined it with spin system alignment-total correlation spectroscopy (TOCSY) and heteronuclear multiple/single quantum coherence (HMQC or HSQC) spectroscopy in synchrony to determine the metabolic profiles and serum levels of 14 pancreatic cancer patients and 12 controls. They found significantly higher levels of total choline, taurine and glucose plus triglycerides in cancers versus control samples. Species that could not be individually identified (unresolved complex matter) were also observed to be significantly different between the two groups. Bathe et al (323) quantified 58 metabolites from the serum samples of 56 pancreatic cancer patients compared to 43 controls using 1H NMR. Twenty-four of these metabolites were analyzed using PCA and supervised orthogonal partial discriminants analysis in a subgroup of 43 cancers and 41 controls and were seen to have a visible separation between the groups. Elevated levels of glutamate,
acetone and 3-hydroxybutyrate were strongly related to the disease process. Elevated levels of phenylalanine, mannose, formate and glucose were associated with age and disease. The AUROC area was 0.8372 indicating excellent predictive ability. This study used patients with benign pancreatic and hepatobiliary disease as controls and as such this could conceivably confound the results.

4.3.3.2. Studies with MS generated profiles

Three studies (324-326) also evaluated the used of MS based metabolomic techniques in studying pancreas cancer. In the first study, Nishiumi and colleagues (324) evaluated the differences between 20 pancreatic cancer patients with locally advanced or metastatic disease and 9 controls using gas chromatography mass spectrometry (GC/MS). They detected a total of 60 metabolites in serum, with a significant difference in the levels of 18 of the 60 metabolites in pancreatic cancer patients compared to the controls. The levels of lactic acid, thiodiglycolic acid, 7-hydroxoctanoic acid, asparagine, aconitic acid, homogentisic acid, and N-acetyltyrosine were significantly increased in the pancreatic cancer patients while levels of L-glycine, urea, octanoic acid, glyceric acid, decanoic acid, lauric acid, myristic acid, palmitic acid, uric acid, margaric acid, and stearic acid were significantly decreased. Further analysis revealed that the variations in the levels of lactic acid, aconitic acid, thiodiglycolic acid, urea, octanoic acid, glyceric acid, uric acid, and stearic acid largely contributed to the observed separation of the pancreatic cancer patients and healthy volunteers. The differences between different stages of pancreatic cancer were also assessed and the authors suggested that it was possible to discriminate between the Stage III and Stage IV groups. In this study, despite the attempt of the authors to sub-classify the metabolite profile according to stage of cancer progression, it is uncertain what impact this would have in identifying pancreatic cancer at an earlier stage.

In the second study, Leichtle et al (325) collected serum samples from 40 patients with pancreatic carcinoma, 40 controls, and 23 pancreatitis patients and generated amino acid profiles by routine mass-spectrometry. They compared these amino acid profiles in an intrinsic three-class bioinformatic approach and evaluated their selectivity using a multiple marker panel combined with the conventional tumor marker CA 19-9. They found 22 of 26 amino acids altered in at least one out of ten possible comparisons. The selectivity of single amino acid concentrations was
assessed in all disease states simultaneously via the volume under ROC surface (VUS).

Combined models of Ca19-9 and amino acids were also evaluated using PCA. This study used four means of assessment including the preliminary selection by calculation of VUS values of the single amino acid concentrations, by generating a binomial logistic regression model using a Bayesian multinomial logit model averaging procedure to generate multi-marker panels (including CA 19-9), analyze these generated predictors using VUS, and lastly by determining the non-inferiority and superiority of the panel. They found that the panel predictors had a high VUS and postulated that multiple metabolite panels would show an improved selectivity compared to Ca19-9 alone. This study provides a comprehensive framework for the development of biomarker panels and subsequent biomarker validation. Drawbacks could include a degree of bias due to minimizing inherent variability and a loss of selectivity due to strict pre-analytical modeling.

In the third study, Ritchie et al (326) used high-resolution, flow-injection Fourier transform ion cyclotron resonance mass spectrometry (FI-FTICR-MS) to generate comprehensive metabolomic profiles from the serum of 40 patients with pancreas cancer and 50 controls. The authors used three statistical models to analyze the data. These included PCA to compress the data and assess variance, followed by hierarchical cluster analysis to group related metabolic systems, and lastly, random forest to classify and cross-validate the strongest metabolic discriminators. FI-FTICR-MS metabolomic analysis showed significant reductions in the serum levels of metabolites belonging to five systems in cancer patients compared to controls. The metabolic systems included 36-carbon ultra long-chain fatty acids, multiple choline-related systems including phosphatidylcholines, lysophosphatidylcholines and sphingomyelins, as well as vinyl ether-containing plasmalogen ethanolamines. Targeted flow-injection tandem mass spectrometry (FI-MS/MS) assays for specific metabolic systems were developed and used to validate the FI-FTICR-MS results. The results confirmed the involvement of three altered metabolic systems in the serum of cancer patients: ultra long-chain fatty acids, numerous choline-containing glycerophospholipids, and vinyl ether-containing ethanolamine phospholipids called plasmalogens. Although most of the individual metabolites alone showed a significant reduction in PC patient serum, the strongest discriminator based on multiple statistical criteria was PC-594 (p=9.9E-14). The ROC-AUCs based on FI-MS/MS of selected markers from each system ranged between 0.93 ±0.03 and 0.97 ±0.02. An ultra long chain fatty acid PC 594 was further validated in a low risk population with a sensitivity of 87% in a population with a predicted pretest risk of 10%.
4.3.4. Profiling using saliva samples

Sugimoto et al. (310) conducted a comprehensive metabolite analysis of saliva samples obtained from 215 individuals (69 oral, 18 pancreatic and 30 breast cancer patients, 11 periodontal disease patients and 87 healthy controls) using capillary electrophoresis time-of-flight mass spectrometry (CE-TOF-MS). They identified 57 principal metabolites that can be used to accurately predict the probability of being affected by each individual disease. They used independent multiple linear regression models to discriminate healthy individuals and each disease cohort using a stepwise variable selection method (backward procedure to eliminate non-predictive peaks with a threshold of p<0.10 to construct the predictive models. The models were evaluated by tenfold cross-validation. PCAs developed individually for the control and each disease group showed partial discriminative separation of the subjects. Eight amino acid metabolites (leucine with isoleucine, tryptophan, valine, glutamic acid, phenylalanine, glutamine, and aspartic acid) were identified as pancreatic cancer-specific markers. They postulated that quantitative information for these metabolites and their combinations could enable us to predict disease susceptibility. The drawbacks of this study included a relatively small sample size for pancreatic cancer samples, the methodology of collection of saliva samples has not been validated with respect to time of collection, influence of dietary factors, smoking and need for a better matched controls and cases.

4.3.5. Profiling using bile samples

One study aimed to identify the levels of d-glucuronic acid in bile samples obtained from four pancreatic cancer patients using one-dimensional 1H NMR and two-dimensional 1H−1H correlated spectroscopy experiments (327). This group observed that d-glucuronic acid was highly elevated in the bile samples obtained from pancreatic cancer patients whereas it was either absent or negligible in control and chronic pancreatitis patients. Bile is a relatively difficult to access fluid and as such can only be obtained through invasive means during the investigation of pancreatobiliary disease. As such, the use of bile as a source of metabolites to diagnose early pancreatic cancer would entail significant risk to the patient and preclude its use as a screening tool.
4.3.6. Profiling using urine samples

Urine represents an ideal bio-fluid for NMR analysis because it contains the highest number of water-soluble metabolites, requires minimal sample preparation, is non-invasive and can be sampled frequently. Clinically, urinalysis can be used to screen large populations, however, metabolomic analysis is limited by high variability in phenotype, lifestyle, diet and other environmental factors (335, 336). Davis et al (328) performed a global metabolite analysis of urine samples from 32 patients with pancreatic ductal cancer, 32 controls and 25 patients with benign pancreatic disease using 1H NMR spectroscopy. They identified 22 metabolites that showed significantly different levels of variation between the cancers and the controls. There was also a clear difference between the cancers and the urinary signatures from patients with benign pancreatic pathology. The metabolites that were significantly increased included acetone, hypoxanthine, O-acetylcarnitine, dimethylamine, choline, 1-methylnicotinamide, threonine, fucose, cis-aconitate, 4-pyridoxate, trimethylamine-N-oxide, aminobutyrate, tryptophan, trans-aconitate, 4-hydroxyphenylacetate and taurine. Trigonelline was decreased in the urinary cancer samples.

Napoli et al (329) used NMR spectroscopy was applied to investigate the urine metabolome of thirty-three patients with pancreatic cancer to detect altered metabolic profiles in comparison with 54 matched controls. Statistically significant differences were found between urine metabolomic profiles of pancreatic cancer and controls. The levels of acetoacetate, leucine, glucose, 2-phenylacetamide, and some acetylated compounds were increased, while citrate, creatinine, glycine, hippurate, 3 hydroxyisovalerate, trigonelline, and an unknown signal at 3.72 ppm were decreased in pancreas cancer compared to controls.

4.4. Animal studies

4.4.1 Sprague-Dawley rat model

Using animal models, Fang et al (330) investigated the metabolic profiles of Sprague–Dawley rat pancreases by high-resolution magic angle spinning proton magnetic resonance spectroscopy (HR-MAS 1H NMR) to discriminate pancreatic cancer from chronic pancreatitis. Intact pancreatic
tissue samples were obtained from Sprague–Dawley rats with histologically proven pancreatic cancer, chronic pancreatitis, and two sets of matched controls for each group. Two 1H NMR experiments, single-pulse and Carr–Purcell–Meiboom–Gill (CPMG), were carried out separately. Increases in phosphocholine and glycerophosphocholine levels and decreases in leucine, isoleucine, valine, lactate and alanine levels were observed in chronic pancreatitis, whereas the opposite trends were observed in pancreatic cancer. Increasing taurine and decreasing betaine were found both in chronic pancreatitis and in pancreatic cancer. Additionally, the lipid content in pancreatic cancer was higher than that in chronic pancreatitis. PCA was carried out for the single-pulse and CPMG 1H NMR spectra respectively, to visualize separation among the samples and to extract characteristic metabolites of pancreatic cancer and chronic pancreatitis. Decreased phosphocholine and glycerophosphocholine were suggested as unique metabolite indicators of pancreatic cancer and that this combination had the potential to improve early diagnosis of pancreatic cancer in clinical settings.

4.4.2 Transgenic pancreatic cancer rat model with expression of human K-rasG12V oncogene and induction is regulated by the Cre/lox system

Yabushita et al (333) evaluated pancreatic cancer tumour serum and tumour tissue were assessed using non-targeted and targeted GC-MS in a previously validated rat pancreas cancer model (337) in which expression of human K-rasG12V oncogene and induction is regulated by the Cre/lox system. Three metabolites showed correspondence between the pancreatic tumour tissue and serum. These included palmitelaidic acid, arachidonic acid and 2(1H)-pyrimidinone. None of the other discriminatory metabolites had had corresponding fluctuations in both the serum and pancreatic tissue. Comparison of the pancreatic tumour tissue and serum found that only palmitoleic acid were significantly different between cancer samples and controls while 2(1H)-pyrimidinone and arachidonic acid while significant in pancreatic tissue were not significantly different in serum samples. In the serum, aminomalonic acid, 2-aminoethyl dihydrogen phosphate, citrate and lanthionine were significantly increased. However, none of these metabolites was significantly increased in pancreatic tissues while aminomalonic acid was significantly decreased. The fatty acids palmitelaidic acid, linoleic acid, octadecanoic acid, myristic acid, arachidonic acid and hexadecanoic acid were significantly decreased in pancreatic tissues. Of these, palmitelaidic acid (significantly) and arachidonic acid (non significant) were
decreased in serum.

4.4.3 Nude mouse xenograft model using injected pancreatic cancer cell SW1990

He et al (332) used 1H NMR to determine the metabolic changes before and after radiotherapy in a nude mouse xenograft pancreatic cancer model. Compared to normal nude mouse pancreas, the levels of choline, taurine, alanine, isoleucine, leucine, valine, lactate, and glutamic acid of the pancreatic cancer group were increased, while phosphocholine, glycerophosphocholine, and betaine were decreased. The ratio of phosphocholine to creatine, and glycerophosphocholine to creatine showed noticeable decrease in the pancreatic cancer group. Evaluation of the tissue metabolic profile after treatment with three different radiation doses showed no significant change in metabolites in the 1H NMR spectra.

4.4.4. Perfused pancreatic cancer cell lines and induced pancreatic cancer

In the second study using animal models, Kaplan and colleagues (331) assessed the role of nuclear magnetic resonance spectroscopy (MRS) in pancreatic cancer diagnosis and its treatment in three models of pancreatic neoplasms. These models were perfused MIA PaCa-2 human pancreatic cancer cells, subcutaneous implanted pancreatic tumors in hamsters, and pancreatic tumors induced in situ in rats by direct application of the carcinogen 7,12-dimethyl benzanthracene. These were studied by phosphorous (31P), sodium (23Na) and proton (1H) spectra. Their NMRS data showed that 31P MRS demonstrated a difference in the lipid profiles of mice with induced pancreas cancer as opposed to normal tissue, though this was not significant. They found no 23Na MRS characteristics of pancreatic cancer and the 1H MRS showed that while taurine and lactate were elevated in pancreatic tumors compared to normal pancreas, creatine and glutamate were higher in the intact organs.

4.5. Factors that influence metabolite levels

There are primarily two forms of metabolomic studies, targeted and untargeted analysis. A targeted analysis is directed to a defined hypothesis and to identify specific metabolites or those belonging to a specific system. On the other hand, untargeted profiling of metabolites provides information regarding thousands of metabolites in a given state. This metabolomic profiling is
done using mainly two high profiling methods – mass spectrometry and nuclear magnetic resonance spectroscopy. There are differences between the two methods. NMR spectroscopy has the advantage of being quantitative, nonselective and highly reproducible. Samples are easily prepped and can be several analyses can be performed on the same sample. Its drawback is its low sensitivity (338, 339). Mass spectrometry on the other hand is very sensitive and can analyse thousands of metabolites in a given sample, but requires intensive sample preparation and the reproducibility can vary between protocols. To enhance the range of metabolite detection, chromatography or capillary electrophoresis can be used, particularly when dealing with complex sample types. The use of chromatography helps in identifying different metabolites based on their physical properties and chemical nature. Gas chromatography is used to analyse volatile compounds while liquid chromatography provides a greater range of metabolite detection but can be used to identify a greater range of metabolites. The human metabolome project was launched in 2004 to clarify and validate previously recorded metabolites and develop a database for recording new metabolites, classify them and identify the conditions they are associated with (340). The project analysed 149 serum samples from healthy volunteers and patients with cardiovascular disease using a 5 different profiling methods and identified a large number of metabolites. However, no metabolite was identified by all five methodologies. In general, the use of UPLC –MS has been advocated as the best metabolomic investigation for plasma and serum and has been recommended as the preliminary comprehensive profiling method. This variation in metabolite pick-up could be attributed to various factors including the separation method, instrument sensitivity, the type of analysis (targeted vs. nontargeted) and the stability, solubility, and volatility of compounds (341). In clinical practice, there are significant variations based around patient and environmental factors, which can influence metabolite levels. Various factors including age, sex (342), body mass index (particularly a rapid change such as cachexia or obesity) (343), diabetes (344), jaundice (345), smoking (346, 347), alcohol (348), nutritional status (including fasting) (349), renal failure, thyroid status (350), hyperlipidaemia (351) have been shown to influence the metabolomic profile at a given time. The exact influence that any of these elements, separately or in a combination, exert on the metabolite profile of an individual (particularly one that has an underlying pathology) at any given time is difficult to estimate, given the lack of studies assessing the role that they play. There is also evidence that the time of sampling may play a role in the metabolite profiles of patients with the body's circadian rhythm exerting control over metabolite levels (352, 353). Studies to detail the role of a diurnal variation
in metabolic profiles have not been undertaken, with understandable confounders such as meals and hormonal variations.

The limitations of the current metabolomic methodologies include the lack of a distinct method for a comprehensive evaluation of the metabolome, a lack of standardization of processing and data extraction and the lack of a comprehensive metabolite database to corroborate provisional biomarkers. Further research is still desirable before advocating an optimal method for blood metabolite analysis that has clinical utility.

4.6. Conclusions

Due to the varied nature of samples and methods used in different histological types of pancreatic cancer in this series, it is difficult to correlate between the metabolomic profiles of patients with pancreatic adenocarcinoma. The metabolite profiles identified in the studies above cover certain basic features of cellular metabolism, from which it is difficult to attribute the degree of change due to the cancer itself. Clearly, there is considerable scope for further research in this subgroup of patients. Avenues for further exploration would include; studies with larger sample sizes, comparison of profiles in multiple sample types in the same set of patients, more studies including plasma or serum samples and larger series using a single method of sample separation and detection.

Despite this, certain features stand out between these studies:

1. Metabolites involved in cellular respiration:

Raised levels of lactate and lower levels of substrates in the citric acid cycle could represent the increased metabolic status within the neoplastic cells. The elevated lactate and decreased glutamate levels in pancreatic tumors as compared to the normal pancreas possibly illustrates the differences in energy production with increased glycolysis in tumor tissue compared to oxidative phosphorylation in normal tissue. This however, is not specific to pancreatic cancer and can be seen in many other cancers. NAD, 5-AMP and 5-UMP are also indicators of the redox potential within the cell and are involved in cellular respiration. Metabolites such as lactate, acetoacetate and beta hydroxyl butyrate are directly related to the cellular redox potential. Carcinogenesis has
demonstrated an increased utilization of glycolytic and glutaminolytic pathways with a consequent decrease in glutamine levels. This has also been seen in several of the studies detailed above.

2. **Lipid content of pancreatic cancer tissue:**

High tissue lipid content has been put forward as an identifier of pancreatic cancer as opposed to chronic pancreatitis. In vitro studies have demonstrated that lipids are required for pancreatic cancer cell proliferation. This growth-promoting effect of lipids on pancreatic cancer cells influences hormone signal transduction, cell membrane structure, and cellular energy supply (334). An increase in levels of arachidonic acid may reflect a state of chronic inflammation. Chronic inflammation induced by cytokines has been shown to contribute to carcinogenesis (354). Omega-6 fatty acids like arachidonic acid are associated with an increase in the production of pro-inflammatory enzymes and cytokines including cyclooxygenase-2, tumor necrosis factor and IL-1. TNF-alpha is also a potent activator of nuclear transcription factor (NF-kB) and has been shown to be over-expressed in pancreatic cancer (219).

3. **Bile acids and taurine:**

Taurine is a cytosolic amino acid involved in the conjugation of bile acids, membrane stabilization, regulation of osmolality and is an anti-oxidant involved in modulation of calcium signaling. Pancreatic taurine production occurs via the cysteine sulfinic acid pathway. Raised taurine levels have also been seen in transplanted CAPAN-1 pancreatic cell lines in nude mice (355). In addition, Urayama et al (320) also demonstrated a rise in levels of bile acids (taurocholic acid, deoxycholyglycine, cholyglycine) in this set of plasma samples in patients with pancreatic cancer.

4. **D-glucuronic acid**

D-Glucuronic acid is synthesized in the liver from UDP glucose, conjugated to bilirubin, forming bilirubin diglucuronide and is excreted into the bile. It is involved in a number of key detoxification pathways. The levels of d-glucuronic acid conjugates in the bile of controls or patients with biliary obstruction are almost undetectable by 1H NMR spectroscopy. The authors propose that the deconjugation of bilirubin diglucuronide present in the bile may account for the presence of this substrate in patients with pancreatic cancer.
5. **Isoleucine, leucine and 3-hydroxyisovalerate**

Leucine is a branched chain amino acid that is involved in cell growth. The catabolism of leucine produces 3-Hydroxyisovalerate as a substrate. Cho and colleagues (319) showed that NF-κB activity regulates the expression of leucine zipper tumor suppressor 2 (lzts2) in human adipose tissue-derived mesenchymal stem cells. Leucine zipper tumor suppressor 2 is frequently lost in a variety of human tumors and is a putative tumor suppressor (356). It represses the transactivation of beta catenin, and affects beta catenin localization by interacting with it. Modulation of lzts 2 expression by lentiviral techniques is shown to affect the proliferation of breast, colon, prostate and glioma cancer cell lines (357). Therefore theoretically, a decrease in serum 3-hydroxyisovalerate levels may be involved in pancreatic cancer.

In conclusion, there is a paucity of metabolomic studies in patients with pancreatic cancer and no studies to date in patients with pancreatic cystic tumors. There is considerable scope to expand our knowledge about the metabolomic profiles in this difficult subset of cancer patients especially given their poor prognosis based on a delayed diagnosis. Current studies have provided a step towards identifying metabolomic markers in pancreatic cancer but require larger sample sizes and biomarker validation studies.

### 4.7. Aims

We aim to generate and compare the metabolic profiles for pancreatic cancer and associated ampullary and endocrine cancers as a preliminary study to identify metabolic profiles in serum samples.

This study will seek to identify metabolites differentiating these conditions based on the null hypothesis that there is no difference in the metabolite profiles between the three conditions.

The preliminary goal was to perform an untargeted analysis on 39 consecutive serum samples of pancreas cancer, ampullary cancer and endocrine cancer collected from patients over a period extending from September 2011 to September 2012 and compared with matched controls obtained from the HUSERMET database. The HUSERMET database is collaboration between the University of Manchester, GlaxoSmithKline and Astra Zeneca. It aims to characterise the
human metabolome and establish typical metabolites and their ranges observed for each technology including GCMS, LCMS, FTIR and NMR. It also aims to identify the disease-independent changes that occur as a function of diet, gender, age, therapeutic interventions and lifestyle and identify biomarkers for Alzheimer’s disease and ovarian cancer.

4.8. Patients

4.8.1. Demographics

A total of 39 consecutive patients were recruited into the study over a period extending from September 2011 to September 2012. Three patients were excluded due to concerns about capacity to consent. Serum was available for all patients. We found it difficult to recruit healthy volunteers matching the age, sex and BMI status of the patient population. One of the main reasons for this was the cachexia experienced by almost all the patients at the time of diagnosis. In addition, new onset diabetes, jaundice and comorbidities in the patient group could not be matched. Ideally, the control group would have matched the patient population in terms of parameters such as age, sex, BMI, diabetes, to minimize artifacts in the profiles being compared. The average BMI in our patient population was 21 (range 19 – 26). The demographic and histological details are documented in Table 15.

Table 15: Patient demographics

<table>
<thead>
<tr>
<th>No</th>
<th>Case</th>
<th>Age</th>
<th>Sex</th>
<th>Management</th>
<th>Diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>P014</td>
<td>73</td>
<td>M</td>
<td>Operated</td>
<td>Pancreatic cancer</td>
</tr>
<tr>
<td>2</td>
<td>P016</td>
<td>77</td>
<td>M</td>
<td>Operated</td>
<td>Pancreatic cancer</td>
</tr>
<tr>
<td>3</td>
<td>P017</td>
<td>74</td>
<td>F</td>
<td>Operated</td>
<td>Pancreatic cancer</td>
</tr>
<tr>
<td>4</td>
<td>P018</td>
<td>59</td>
<td>M</td>
<td>Operated</td>
<td>Pancreatic cancer</td>
</tr>
<tr>
<td>5</td>
<td>P024</td>
<td>60</td>
<td>M</td>
<td>Inoperable</td>
<td>Pancreatic cancer</td>
</tr>
<tr>
<td>6</td>
<td>P025</td>
<td>59</td>
<td>F</td>
<td>Operated</td>
<td>Pancreatic cancer</td>
</tr>
<tr>
<td>7</td>
<td>P027</td>
<td>54</td>
<td>M</td>
<td>Inoperable</td>
<td>Pancreatic cancer</td>
</tr>
<tr>
<td>8</td>
<td>P028</td>
<td>77</td>
<td>M</td>
<td>Operated</td>
<td>Pancreatic cancer</td>
</tr>
<tr>
<td>9</td>
<td>P030</td>
<td>69</td>
<td>F</td>
<td>Operated</td>
<td>Pancreatic cancer</td>
</tr>
<tr>
<td>10</td>
<td>P035</td>
<td>63</td>
<td>M</td>
<td>Inoperable</td>
<td>Pancreatic cancer</td>
</tr>
<tr>
<td>No</td>
<td>Case</td>
<td>Age</td>
<td>Sex</td>
<td>Management</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>----</td>
<td>------</td>
<td>-----</td>
<td>-----</td>
<td>------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>11</td>
<td>P037</td>
<td>68</td>
<td>M</td>
<td>Operated</td>
<td>Pancreatic cancer</td>
</tr>
<tr>
<td>12</td>
<td>P038</td>
<td>82</td>
<td>M</td>
<td>Inoperable</td>
<td>Pancreatic cancer</td>
</tr>
<tr>
<td>13</td>
<td>P039</td>
<td>59</td>
<td>M</td>
<td>Operated</td>
<td>Pancreatic cancer</td>
</tr>
<tr>
<td>14</td>
<td>P044</td>
<td>60</td>
<td>M</td>
<td>Inoperable</td>
<td>Pancreatic cancer</td>
</tr>
<tr>
<td>15</td>
<td>P045</td>
<td>73</td>
<td>F</td>
<td>Inoperable</td>
<td>Pancreatic cancer</td>
</tr>
<tr>
<td>16</td>
<td>P048</td>
<td>70</td>
<td>F</td>
<td>Inoperable</td>
<td>Pancreatic cancer</td>
</tr>
<tr>
<td>17</td>
<td>P049</td>
<td>63</td>
<td>M</td>
<td>Operated</td>
<td>Pancreatic cancer</td>
</tr>
<tr>
<td>18</td>
<td>P053</td>
<td>78</td>
<td>F</td>
<td>Operated</td>
<td>Pancreatic cancer</td>
</tr>
<tr>
<td>19</td>
<td>P054</td>
<td>42</td>
<td>M</td>
<td>Inoperable</td>
<td>Pancreatic cancer</td>
</tr>
<tr>
<td>20</td>
<td>P015</td>
<td>49</td>
<td>M</td>
<td>Conservative</td>
<td>Endocrine tumour</td>
</tr>
<tr>
<td>21</td>
<td>P047</td>
<td>36</td>
<td>M</td>
<td>Operated</td>
<td>Endocrine tumour</td>
</tr>
<tr>
<td>22</td>
<td>P052</td>
<td>58</td>
<td>F</td>
<td>Operated</td>
<td>Endocrine cancer</td>
</tr>
<tr>
<td>23</td>
<td>P013</td>
<td>63</td>
<td>F</td>
<td>Operated</td>
<td>Endocrine cancer</td>
</tr>
<tr>
<td>24</td>
<td>P034</td>
<td>57</td>
<td>M</td>
<td>Operated</td>
<td>Endocrine cancer</td>
</tr>
<tr>
<td>25</td>
<td>P019</td>
<td>80</td>
<td>F</td>
<td>Operated</td>
<td>Ampullary cancer</td>
</tr>
<tr>
<td>26</td>
<td>P023</td>
<td>76</td>
<td>M</td>
<td>Operated</td>
<td>Ampullary cancer</td>
</tr>
<tr>
<td>27</td>
<td>P026</td>
<td>80</td>
<td>M</td>
<td>Operated</td>
<td>Ampullary cancer</td>
</tr>
<tr>
<td>28</td>
<td>P029</td>
<td>72</td>
<td>M</td>
<td>Operated</td>
<td>Ampullary cancer</td>
</tr>
<tr>
<td>29</td>
<td>P031</td>
<td>61</td>
<td>M</td>
<td>Operated</td>
<td>Ampullary cancer</td>
</tr>
<tr>
<td>30</td>
<td>P033</td>
<td>72</td>
<td>F</td>
<td>Operated</td>
<td>Ampullary cancer</td>
</tr>
<tr>
<td>31</td>
<td>P043</td>
<td>76</td>
<td>M</td>
<td>Operated</td>
<td>Ampullary cancer</td>
</tr>
<tr>
<td>32</td>
<td>P046</td>
<td>69</td>
<td>M</td>
<td>Operated</td>
<td>Ampullary cancer</td>
</tr>
<tr>
<td>33</td>
<td>P050</td>
<td>77</td>
<td>M</td>
<td>Inoperable</td>
<td>Ampullary cancer</td>
</tr>
<tr>
<td>34</td>
<td>P011</td>
<td>67</td>
<td>F</td>
<td>Operated</td>
<td>Cholangiocarcinoma</td>
</tr>
<tr>
<td>35</td>
<td>P032</td>
<td>69</td>
<td>M</td>
<td>Operated</td>
<td>Pancreatic cystic tumour</td>
</tr>
<tr>
<td>36</td>
<td>P021</td>
<td>47</td>
<td>M</td>
<td>Operated</td>
<td>Chronic Pancreatitis</td>
</tr>
<tr>
<td>37</td>
<td>P022</td>
<td>27</td>
<td>M</td>
<td>Operated</td>
<td>Chronic Pancreatitis</td>
</tr>
<tr>
<td>38</td>
<td>P020</td>
<td>65</td>
<td>F</td>
<td></td>
<td>Control</td>
</tr>
<tr>
<td>39</td>
<td>P040</td>
<td>40</td>
<td>F</td>
<td></td>
<td>Control</td>
</tr>
</tbody>
</table>
The recruitment and sampling for this study was undertaken after the previous studies, hence the patient populations from whom the samples were collected (including both the pancreas cancer and cystic tumour group), do not match those of the previous studies in this thesis.

**4.8.2. Distribution of Pathology**

This series comprises of 19 cases of pancreatic adenocarcinoma. In addition, there were 5 cases of endocrine tumors and 9 ampullary cancers. Two patients had chronic pancreatitis and 2 controls were collected. There was one case each of pancreatic cystic neoplasm and cholangiocarcinoma. This is shown in Figure 35.

![Figure 35: Relative proportions of sample histology](image)

The median age of the patients was 67 years (range from 27 to 80 years); 13 (33%) cases were female and 26 (67%) were male.

**4.9. Protocols**

**4.9.1. Serum Collection**

Serum collection was performed on fasting patients, just prior to surgery. The sample was allowed to clot for at least one hour at ~4 °C and were spun within 6 hours. The serum blood collection tube was centrifuged at ~2500g for 15 minutes at ~4 °C. The blood collection tube cap was carefully removed to avoid disturbing the pellet or splashing of the blood. A pipette with
appropriate filter tip was used to transfer at least six 0.5mL and two 0.2mL aliquots of serum into labeled cryovials. Once these eight aliquots were made, the remainder of the sample was split into two aliquots of at least 0.5mL. Care was taken to avoid taking any red blood cells over into the serum. Sub-Aliquots were filled to a volume of 500µl each.

4.9.2. Cryovial Storage

The cryovials were then placed into labeled and barcoded cryoboxes (10x10) filled from the left hand corner (A1), with aliquots being filled left to right and top to bottom, and placed immediately into -80ºC storage. New aliquots were placed in a manner so as to minimise the thawing effect on the samples already frozen.

4.9.3. LC-TOF-MS analysis

Ultra high performance liquid chromatography mass spectrometry (UHPLC-MS) analysis was carried out on an Accela UHPLC autosampler system coupled to an electrospray LTQ-Orbitrap XL hybrid mass spectrometry system (ThermoFisher, Bremen, Germany). Analysis was carried out in both positive and negative ESI modes whilst each run was completely randomised to negate for any bias. A gradient type UHPLC method was used during each run as is previously described (358, 359). 10 µL of the extract was injected onto a Hypersil GOLD UHPLC C\textsubscript{18} column (length 100mm, diameter 2.1 mm, particle size 1.9 µm, Thermo-Fisher Ltd. Hemel Hempsted, UK) held at a constant temperature of 50ºC whilst a solvent flow rate of 400 µL/min\textsuperscript{-1} was used to drive the chromatographic separation.

Xcaliber software (Thermo-Fisher Ltd. Hemel Hempsted, U.K.) was used as the operating system for the Thermo LTQ-Orbitrap XL MS system following the method described (359).

Data processing was initiated by the conversion of the standard UHPLC raw files in to the universal NetCDF format via the software conversion tool within Xcaliber. Subsequently, in-house peak deconvolution software containing the XCMS algorithm ([http://massspec.scripps.edu/xcms/xcms.php](http://massspec.scripps.edu/xcms/xcms.php)) was used for pick picking as described previously (359, 360). The output from this system resulted in a MS Excel based data matrix of mass spectral features with related accurate m/z and retention time pairs. Data from the internally
pooled QC samples was then used to align for instrument drift and quality control \textit{via} application of an in-house Matlab script \textsuperscript{(360)}. The data matrix was also signal corrected to remove peaks that crossed the 20% relative standard deviation (RSD) threshold within QC samples across the analytical run. Normalisation of each peak within the samples was achieved using the mean peak area whilst putative identification of all features were performed applying the PUTMEDID-LCMS set of workflows as previously described \textsuperscript{(361)}. Ambiguity arising from the same \textit{m/z} ratio can lie within lipid identification due to differing points of unsaturation and multiple identifications. Multiple adducts of the same lipid can also occur due to the presence of different charged (composite) species (i.e. protonated and sodiated ions).

\textbf{4.9.4. Data Processing}

A set of reference spectra was compiled for a list of serum metabolites generated and all subsequently generated metabolites were searched against this reference table. Where possible, peak identities were assigned on the basis of mass spectra similarity to NIST library entries or a metabolite library generated at the University of Manchester. Once the reference table was generated, the supporting information for each peak was specified.

\textbf{4.9.5. Statistical Analysis}

Both univariate and multivariate analysis were performed on the ion ratio data sets. Univariate analysis (ANOVA) was analyzed with Microsoft Excel 2003 and multivariate analysis was performed with Matlab R 2008a (Math Works Inc. MA). Unsupervised multivariate analysis was performed using principal components analysis (PCA). Principal components analysis (PCA) is a statistical method used to identify patterns in data and expressing the data in a manner that permits identification of their similarities and differences. The use of PCA can convert a possibly correlated dataset into a set of linearly uncorrelated variables. Also, once patterns within the dataset have been identified, the data can be compressed by reducing the dimensions into principal components, without loss of information. The PCA results are discussed in terms of component scores (the value at which the transformed variable corresponds to the data point) and the loading (the weight by which the original variable must be multiplied to get the component score).
4.10. Results

The 39 sample types were distributed as follows (as shown in Figure 35):

- Pancreatic cancer (PC): 19
- Ampullary cancer (AC): 9
- Endocrine cancer (EC): 5
- Chronic pancreatitis (CP): 2
- Cholangiocarcinoma (CC): 1
- Pancreatic cystic neoplasm (CY): 1
- Control samples: 2

These samples were matched for age, sex and body mass index (BMI) from a set of 21 controls obtained from the HUSERMET (previously analyzed metabolomic set collected from healthy individuals) database stored at the metabolomics laboratory at the University of Manchester.

All the resected endocrine cancers were well-differentiated tumours histologically. They were graded according to their mitotic count and Ki67 proliferative index with all four operated tumours demonstrating a Ki67 index of less than 3%. Endocrine cancers and ampullary cancers have different pathophysiological features as compared to pancreas cancer and could possibly have a different metabolic profile. No previous studies have been carried out to assess these differences.

4.10.1. Case only PCA analysis

PC scores are the derived composite scores computed for each observation based on the amount of weighted variation for each principal component. The PCA scores for the metabolites in both positive and negative ion modes were extracted from the serum samples. Positive ion modes are found in Table 16 and negative ion modes are found in Table 17.
### Table 16: LCMS ESI + PCA results

<table>
<thead>
<tr>
<th>Idx</th>
<th>ID</th>
<th>A</th>
<th>Sex</th>
<th>BMI</th>
<th>Plan</th>
<th>D</th>
<th>Abr</th>
<th>PC1_score</th>
<th>PC2_score</th>
</tr>
</thead>
<tbody>
<tr>
<td>39</td>
<td>P054</td>
<td>42</td>
<td>M</td>
<td>N</td>
<td>Inoperable</td>
<td>PC</td>
<td>P</td>
<td>-47.2838531</td>
<td>2.31893095</td>
</tr>
<tr>
<td>1</td>
<td>P011</td>
<td>67</td>
<td>F</td>
<td>N</td>
<td>Operated</td>
<td>CCa</td>
<td>B</td>
<td>7.387702999</td>
<td>-8.831585996</td>
</tr>
<tr>
<td>5</td>
<td>P016</td>
<td>77</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>PC</td>
<td>P</td>
<td>-8.091662657</td>
<td>0.350391088</td>
</tr>
<tr>
<td>2</td>
<td>P013</td>
<td>63</td>
<td>F</td>
<td>N</td>
<td>Operated</td>
<td>EC</td>
<td>E</td>
<td>-4.561269032</td>
<td>-18.04412425</td>
</tr>
<tr>
<td>27</td>
<td>P039</td>
<td>59</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>PC</td>
<td>P</td>
<td>-6.4056651</td>
<td>15.37758602</td>
</tr>
<tr>
<td>4</td>
<td>P015</td>
<td>49</td>
<td>M</td>
<td>N</td>
<td>Conserv</td>
<td>EC</td>
<td>E</td>
<td>34.77396066</td>
<td>1.834730221</td>
</tr>
<tr>
<td>31</td>
<td>P045</td>
<td>73</td>
<td>F</td>
<td>N</td>
<td>Inoperable</td>
<td>PC</td>
<td>P</td>
<td>-14.38010101</td>
<td>-31.12006096</td>
</tr>
<tr>
<td>8</td>
<td>P019</td>
<td>80</td>
<td>F</td>
<td>N</td>
<td>Operated</td>
<td>AC</td>
<td>A</td>
<td>10.66422846</td>
<td>25.94631423</td>
</tr>
<tr>
<td>30</td>
<td>P044</td>
<td>60</td>
<td>M</td>
<td>N</td>
<td>Inoperable</td>
<td>PC</td>
<td>P</td>
<td>-27.07787516</td>
<td>1.115789235</td>
</tr>
<tr>
<td>10</td>
<td>P021</td>
<td>47</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>CP</td>
<td>CP</td>
<td>12.82226493</td>
<td>-24.44096846</td>
</tr>
<tr>
<td>6</td>
<td>P017</td>
<td>74</td>
<td>F</td>
<td>N</td>
<td>Operated</td>
<td>PC</td>
<td>P</td>
<td>43.61121946</td>
<td>8.72343482</td>
</tr>
<tr>
<td>9</td>
<td>P020</td>
<td>65</td>
<td>F</td>
<td>N</td>
<td>n/a</td>
<td>C2</td>
<td>C2</td>
<td>-56.51745488</td>
<td>-8.924139822</td>
</tr>
<tr>
<td>7</td>
<td>P018</td>
<td>59</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>PC</td>
<td>P</td>
<td>-3.375629308</td>
<td>-25.61191119</td>
</tr>
<tr>
<td>11</td>
<td>P022</td>
<td>27</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>CP</td>
<td>CP</td>
<td>-3.761094694</td>
<td>18.22779265</td>
</tr>
<tr>
<td>26</td>
<td>P038</td>
<td>82</td>
<td>M</td>
<td>N</td>
<td>Inoperable</td>
<td>PC</td>
<td>P</td>
<td>49.42555605</td>
<td>35.8423828</td>
</tr>
<tr>
<td>12</td>
<td>P023</td>
<td>76</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>AC</td>
<td>A</td>
<td>-3.882960551</td>
<td>5.357126338</td>
</tr>
<tr>
<td>16</td>
<td>P027</td>
<td>54</td>
<td>M</td>
<td>N</td>
<td>Inoperable</td>
<td>PC</td>
<td>P</td>
<td>25.76390085</td>
<td>-9.09854021</td>
</tr>
<tr>
<td>15</td>
<td>P026</td>
<td>80</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>AC</td>
<td>A</td>
<td>-30.73111271</td>
<td>-2.220429715</td>
</tr>
<tr>
<td>18</td>
<td>P029</td>
<td>72</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>AC</td>
<td>A</td>
<td>-8.295921593</td>
<td>-0.446207022</td>
</tr>
<tr>
<td>38</td>
<td>P053</td>
<td>78</td>
<td>F</td>
<td>N</td>
<td>Operated</td>
<td>PC</td>
<td>P</td>
<td>36.99088287</td>
<td>-22.56243155</td>
</tr>
<tr>
<td>20</td>
<td>P031</td>
<td>61</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>AC</td>
<td>A</td>
<td>7.104170038</td>
<td>15.53860047</td>
</tr>
<tr>
<td>24</td>
<td>P035</td>
<td>63</td>
<td>M</td>
<td>N</td>
<td>Inoperable</td>
<td>PC</td>
<td>P</td>
<td>64.45603932</td>
<td>17.77419188</td>
</tr>
<tr>
<td>21</td>
<td>P032</td>
<td>69</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>Cyst</td>
<td>CY</td>
<td>28.46182724</td>
<td>-12.73197902</td>
</tr>
<tr>
<td>13</td>
<td>P024</td>
<td>60</td>
<td>M</td>
<td>N</td>
<td>Inoperable</td>
<td>PC</td>
<td>P</td>
<td>16.7432482</td>
<td>-25.51088196</td>
</tr>
<tr>
<td>22</td>
<td>P033</td>
<td>72</td>
<td>F</td>
<td>N</td>
<td>Operated</td>
<td>AC</td>
<td>A</td>
<td>58.25871273</td>
<td>29.43370151</td>
</tr>
<tr>
<td>34</td>
<td>P048</td>
<td>70</td>
<td>F</td>
<td>N</td>
<td>Inoperable</td>
<td>PC</td>
<td>P</td>
<td>-27.62594845</td>
<td>-14.70100705</td>
</tr>
<tr>
<td>29</td>
<td>P043</td>
<td>76</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>AC</td>
<td>A</td>
<td>16.11478967</td>
<td>2.538288671</td>
</tr>
<tr>
<td>14</td>
<td>P025</td>
<td>59</td>
<td>F</td>
<td>N</td>
<td>Operated</td>
<td>PC</td>
<td>P</td>
<td>-32.194322</td>
<td>27.18737767</td>
</tr>
<tr>
<td>32</td>
<td>P046</td>
<td>69</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>AC</td>
<td>A</td>
<td>-12.53783467</td>
<td>-0.136004148</td>
</tr>
<tr>
<td>19</td>
<td>P030</td>
<td>69</td>
<td>F</td>
<td>N</td>
<td>Operated</td>
<td>PC</td>
<td>P</td>
<td>-29.4624634</td>
<td>-28.91607674</td>
</tr>
<tr>
<td>3</td>
<td>P014</td>
<td>73</td>
<td>M</td>
<td>N</td>
<td>Operable</td>
<td>PC</td>
<td>P</td>
<td>-3.394504707</td>
<td>24.82245578</td>
</tr>
<tr>
<td>23</td>
<td>P034</td>
<td>57</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>EC</td>
<td>E</td>
<td>-9.257361079</td>
<td>6.403311829</td>
</tr>
<tr>
<td>17</td>
<td>P028</td>
<td>77</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>PC</td>
<td>P</td>
<td>-19.22787294</td>
<td>-5.090515044</td>
</tr>
<tr>
<td>35</td>
<td>P049</td>
<td>63</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>PC</td>
<td>P</td>
<td>-56.19309115</td>
<td>41.14776541</td>
</tr>
<tr>
<td>33</td>
<td>P047</td>
<td>36</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>EC</td>
<td>E</td>
<td>3.80522097</td>
<td>-25.06899776</td>
</tr>
<tr>
<td>40</td>
<td>P055</td>
<td>76</td>
<td>M</td>
<td>N</td>
<td>Inoperable</td>
<td>PC</td>
<td>P</td>
<td>-19.47028353</td>
<td>-1.799605112</td>
</tr>
<tr>
<td>37</td>
<td>P052</td>
<td>58</td>
<td>F</td>
<td>N</td>
<td>Operated</td>
<td>EC</td>
<td>E</td>
<td>10.63497289</td>
<td>-23.06323352</td>
</tr>
<tr>
<td>36</td>
<td>P051</td>
<td>75</td>
<td>F</td>
<td>N</td>
<td>Operated</td>
<td>PC</td>
<td>P</td>
<td>12.56103134</td>
<td>-47.03626878</td>
</tr>
<tr>
<td>25</td>
<td>P037</td>
<td>68</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>PC</td>
<td>P</td>
<td>-20.36766515</td>
<td>55.41203099</td>
</tr>
<tr>
<td>Idx</td>
<td>ID</td>
<td>A</td>
<td>Sex</td>
<td>BMI</td>
<td>Plan</td>
<td>D</td>
<td>Abr</td>
<td>PC1 score</td>
<td>PC2 score</td>
</tr>
<tr>
<td>-----</td>
<td>----</td>
<td>----</td>
<td>------</td>
<td>-----</td>
<td>--------</td>
<td>----</td>
<td>-----</td>
<td>-----------</td>
<td>-----------</td>
</tr>
<tr>
<td>39</td>
<td>P054</td>
<td>42</td>
<td>M</td>
<td>N</td>
<td>Inoperable</td>
<td>PC</td>
<td>P</td>
<td>-41.7172641</td>
<td>-2.399174484</td>
</tr>
<tr>
<td>1</td>
<td>P011</td>
<td>67</td>
<td>F</td>
<td>N</td>
<td>Operated</td>
<td>CC</td>
<td>B</td>
<td>14.65857728</td>
<td>-1.243093567</td>
</tr>
<tr>
<td>5</td>
<td>P016</td>
<td>77</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>PC</td>
<td>P</td>
<td>-15.15487798</td>
<td>-4.613489307</td>
</tr>
<tr>
<td>2</td>
<td>P013</td>
<td>63</td>
<td>F</td>
<td>N</td>
<td>Operated</td>
<td>EC</td>
<td>E</td>
<td>6.622670933</td>
<td>28.36650043</td>
</tr>
<tr>
<td>27</td>
<td>P039</td>
<td>59</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>PC</td>
<td>P</td>
<td>18.23138693</td>
<td>4.801096549</td>
</tr>
<tr>
<td>4</td>
<td>P015</td>
<td>49</td>
<td>M</td>
<td>N</td>
<td>Conservat</td>
<td>EC</td>
<td>E</td>
<td>7.986133589</td>
<td>37.10911589</td>
</tr>
<tr>
<td>31</td>
<td>P045</td>
<td>73</td>
<td>F</td>
<td>N</td>
<td>Inoperable</td>
<td>PC</td>
<td>P</td>
<td>-5.769342221</td>
<td>17.22413763</td>
</tr>
<tr>
<td>8</td>
<td>P019</td>
<td>80</td>
<td>F</td>
<td>N</td>
<td>Operated</td>
<td>AC</td>
<td>A</td>
<td>13.3358928</td>
<td>-11.4812839</td>
</tr>
<tr>
<td>30</td>
<td>P044</td>
<td>60</td>
<td>M</td>
<td>N</td>
<td>Inoperable</td>
<td>PC</td>
<td>P</td>
<td>-57.55796271</td>
<td>-23.47425872</td>
</tr>
<tr>
<td>6</td>
<td>P017</td>
<td>74</td>
<td>F</td>
<td>N</td>
<td>Operated</td>
<td>PC</td>
<td>P</td>
<td>29.06977677</td>
<td>-21.85536862</td>
</tr>
<tr>
<td>9</td>
<td>P020</td>
<td>65</td>
<td>F</td>
<td>N</td>
<td>n/a</td>
<td>C2</td>
<td>C2</td>
<td>-4.48863391</td>
<td>37.10346019</td>
</tr>
<tr>
<td>7</td>
<td>P018</td>
<td>59</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>PC</td>
<td>P</td>
<td>-7.005263961</td>
<td>30.36942839</td>
</tr>
<tr>
<td>11</td>
<td>P022</td>
<td>27</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>CP</td>
<td>CP</td>
<td>-4.945821149</td>
<td>2.542560869</td>
</tr>
<tr>
<td>26</td>
<td>P038</td>
<td>82</td>
<td>M</td>
<td>N</td>
<td>Inoperable</td>
<td>PC</td>
<td>P</td>
<td>1.523887522</td>
<td>-21.5679058</td>
</tr>
<tr>
<td>12</td>
<td>P023</td>
<td>76</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>AC</td>
<td>A</td>
<td>24.12600259</td>
<td>9.527972508</td>
</tr>
<tr>
<td>16</td>
<td>P027</td>
<td>54</td>
<td>M</td>
<td>N</td>
<td>Inoperable</td>
<td>PC</td>
<td>P</td>
<td>5.589488826</td>
<td>17.99141168</td>
</tr>
<tr>
<td>15</td>
<td>P026</td>
<td>80</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>AC</td>
<td>A</td>
<td>-26.9085833</td>
<td>-1.75045969</td>
</tr>
<tr>
<td>18</td>
<td>P029</td>
<td>72</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>AC</td>
<td>A</td>
<td>-20.10120307</td>
<td>0.762505963</td>
</tr>
<tr>
<td>38</td>
<td>P053</td>
<td>78</td>
<td>F</td>
<td>N</td>
<td>Operated</td>
<td>PC</td>
<td>P</td>
<td>12.32121939</td>
<td>-4.82282667</td>
</tr>
<tr>
<td>20</td>
<td>P031</td>
<td>61</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>AC</td>
<td>A</td>
<td>17.65563035</td>
<td>-13.8666478</td>
</tr>
<tr>
<td>24</td>
<td>P035</td>
<td>63</td>
<td>M</td>
<td>N</td>
<td>Inoperable</td>
<td>PC</td>
<td>P</td>
<td>15.78936822</td>
<td>-31.76899825</td>
</tr>
<tr>
<td>21</td>
<td>P032</td>
<td>69</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>CY</td>
<td>CY</td>
<td>22.58348573</td>
<td>-8.762528374</td>
</tr>
<tr>
<td>13</td>
<td>P024</td>
<td>60</td>
<td>M</td>
<td>N</td>
<td>Inoperable</td>
<td>PC</td>
<td>P</td>
<td>22.95864774</td>
<td>11.31956737</td>
</tr>
<tr>
<td>22</td>
<td>P033</td>
<td>72</td>
<td>F</td>
<td>N</td>
<td>Operated</td>
<td>AC</td>
<td>A</td>
<td>32.84859163</td>
<td>-21.39235642</td>
</tr>
<tr>
<td>34</td>
<td>P048</td>
<td>70</td>
<td>F</td>
<td>N</td>
<td>Inoperable</td>
<td>PC</td>
<td>P</td>
<td>-2.48558693</td>
<td>-3.378785514</td>
</tr>
<tr>
<td>29</td>
<td>P043</td>
<td>76</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>AC</td>
<td>A</td>
<td>-11.2166376</td>
<td>0.243680974</td>
</tr>
<tr>
<td>14</td>
<td>P025</td>
<td>59</td>
<td>F</td>
<td>N</td>
<td>Operated</td>
<td>PC</td>
<td>P</td>
<td>-7.604790228</td>
<td>-17.94689969</td>
</tr>
<tr>
<td>32</td>
<td>P046</td>
<td>69</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>AC</td>
<td>A</td>
<td>-21.91526534</td>
<td>-11.50293023</td>
</tr>
<tr>
<td>19</td>
<td>P030</td>
<td>69</td>
<td>F</td>
<td>N</td>
<td>Operated</td>
<td>PC</td>
<td>P</td>
<td>-7.900932444</td>
<td>13.94864977</td>
</tr>
<tr>
<td>3</td>
<td>P014</td>
<td>73</td>
<td>M</td>
<td>N</td>
<td>Operable</td>
<td>PC</td>
<td>P</td>
<td>31.77289379</td>
<td>1.80360523</td>
</tr>
<tr>
<td>23</td>
<td>P034</td>
<td>57</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>EC</td>
<td>E</td>
<td>-6.198140142</td>
<td>14.01499134</td>
</tr>
<tr>
<td>17</td>
<td>P028</td>
<td>77</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>PC</td>
<td>P</td>
<td>-15.19673845</td>
<td>9.289194675</td>
</tr>
<tr>
<td>35</td>
<td>P049</td>
<td>63</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>PC</td>
<td>P</td>
<td>1.546843868</td>
<td>-12.647758</td>
</tr>
<tr>
<td>33</td>
<td>P047</td>
<td>36</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>EC</td>
<td>E</td>
<td>-7.709371999</td>
<td>11.32544462</td>
</tr>
<tr>
<td>40</td>
<td>P055</td>
<td>76</td>
<td>M</td>
<td>N</td>
<td>Inoperable</td>
<td>PC</td>
<td>P</td>
<td>-51.53192274</td>
<td>-16.0573066</td>
</tr>
<tr>
<td>37</td>
<td>P052</td>
<td>58</td>
<td>F</td>
<td>N</td>
<td>Operated</td>
<td>EC</td>
<td>E</td>
<td>27.2666262</td>
<td>-7.772803749</td>
</tr>
<tr>
<td>36</td>
<td>P051</td>
<td>75</td>
<td>F</td>
<td>N</td>
<td>Operated</td>
<td>PC</td>
<td>P</td>
<td>21.97799571</td>
<td>2.245730784</td>
</tr>
<tr>
<td>25</td>
<td>P037</td>
<td>68</td>
<td>M</td>
<td>N</td>
<td>Operated</td>
<td>PC</td>
<td>P</td>
<td>-2.456776526</td>
<td>-11.95952326</td>
</tr>
<tr>
<td>39</td>
<td>P054</td>
<td>42</td>
<td>M</td>
<td>N</td>
<td>Inoperable</td>
<td>PC</td>
<td>P</td>
<td>-41.7172641</td>
<td>-2.399174484</td>
</tr>
</tbody>
</table>
4.10.2. Principal Component Analysis Plots

The principal components isolated were then graphically displayed in PCA plots. Of the two control samples collected in the study, one of the samples showed significant degradation and could not be used in analysis. The PCA plots below shows that the samples collected from the patient group were dissimilar from the HUSERMET controls. This was consistent between the two ion modes (Figure 36) (Figure 37). The control sample collected with the tumour samples did not show the same dissociation as the HUSERMET controls implying that this separation could largely be due to differences in sample collection methods rather than real metabolomic differences.

Figure 36: PCA plot: Negative ion mode showing disparity between cases and controls

Figure 37: PCA: positive ion mode showing disparity between cases and controls
On studying the chromatograms, it appeared that there were also differences in the chromatogram peaks with a proportion of peaks present in the cases that were absent in the controls (and vice versa). On LCMS ion modes for the individual tumours, there was less variation between the components isolated from individuals with the tumour types and the control sample from the study group. As Figure 38 demonstrates, there was less intergroup variation between the individual cancer sample types. However, the gap remained significant to preclude an accurate assessment of metabolomic profiles.

Figure 38: LCMS ESI positive PCA plots for individual cancer types versus the controls showing separation between the two groups. P - pancreas cancer, A - ampullary cancer, E - endocrine cancer
Figure 39: LCMS ESI negative PCA plots for individual cancer types versus the controls showing separation between the two groups.

P- pancreas cancer, A - ampullary cancer, E - endocrine cancer

When the datasets were reduced to those peaks that were present in at least two-thirds of both cases and controls, results were more credible but there was still a degree of separation between the HUSERMET controls and the cases making interpretation between the two groups variable (Figure 40), (Figure 41).
Figure 40: LCMS ESI positive PCA plots for samples sharing two-thirds of peaks between the individual cancers and controls.

P - pancreas cancer, A - ampullary cancer, E - Endocrine cancer

Figure 41: LCMS ESI negative PCA plots for samples with two-thirds of peaks shared between individual cancers and controls.

P - pancreas cancer, A - ampullary cancer, E - endocrine cancer
The graphs above indicate that there is less disparity between the HUSERMET control samples and the study samples when looking at samples with $2/3$rd of shared peaks, than when including all the sample peaks. However, there still appears to be a degree of separation between the sample types. This disparity between the cases and controls on PCA does not permit an accurate interpretation of the data in the case study set, compared with that within the normal control set. This unforeseen limitation prevents the identification of metabolomic markers that distinguish between the groups. As the separation between the controls and cases was very clear, we attempted to find the differences between 3 cancer types, pancreatic cancer (P), ampullary cancer (A) and endocrine cancer (E). The component data demonstrates that the primary reason for this separation between the cases and the controls is not the sampling or the methodology but more likely due to differences in the patient group. This implied that we could assess the variation between components between the groups for any obvious differences.

On excluding the HUSERMET controls, the PCA graphs varied when limited to case samples only. There was a clear separation between cases seen in the negative mode that was not observed in the positive mode. It appeared that negative mode data had a factor in it, which split the case samples into two groups. This separation could be seen when PCA was performed on the full data with controls and quality control samples were included and such separation became more prominent when the PCA was performed exclusively on the case samples. An attempt was made to match the grouping with the metadata but no correlation could be identified between the groups. As this split was not observed in the positive mode this lead to the premise that the result was probably related to the ionization mode alone. In addition, it was difficult to discriminate three cancer types using the negative mode data due to this separation. The one-way ANOVA test identified a few variables with $p$-values less than 0.01 but none of those variables could be considered as "significant" at a false discovery rate (FDR) of 0.1.

The PCA split in the negative mode is represented below with the PCA analysis in the positive mode for comparison (Figure 42).
Figure 42: The split in PCA components became more obvious when the graph was restricted to case samples only, with a clear demarcation seen in the negative mode. P – pancreas cancer, A – ampullary cancer, E – endocrine cancer

On the other hand, there was no clear split observed in the PCA positive ion mode (Figure 43).

Figure 43: No corresponding split was observed in the PCA plots in the positive mode
4.10.3. PC-DFA analysis and results

In general, the use of PCA aims to summarize the overall variability between components, including both inter-group variation and intra-group variation. To assess the relationships between different clusters, discriminant factor analysis (DFA) can be used to emphasize inter-group variability while neglecting intra-group variation. This multivariate technique can be used on these PCA components to allow the best discrimination of the components between the groups. In addition, the use of PC-DFA ensures that the dataset is less than the total number of variables and that the data sample is uncorrelated allowing an interpretation of variation caused by differences in variable characteristics between groups. We used PC-DFA analysis on the case samples only to assess for variation between the three large groups (pancreas cancer, ampullary cancer and endocrine cancer). On the PC-DFA loadings chart, the identified significant variables were highlighted and tested for significance using the FDR. The variables were also assessed for agreement with ANOVA testing. The PC-DFA did show some separation between the three classes but there was little agreement between the PC-DFA loadings and the significant variables in the negative mode identified by the ANOVA test (Figure 44). The negative mode loading plot variables did not show a significant difference on FDR testing and univariate analysis. A possible reason for this could be that the model was over-fitted. It is worth noting that the PC-DFA pattern in the negative mode was similar to that in the positive mode and it may be harsh to conclude that the PC-DFA result in the negative mode was completely over-fitted.
In the positive mode, variables were identified as significant at the level of false discovery rate (FDR) <= 0.1 and using one-way ANOVA test. An interesting observation was that the ANOVA test seemed to favour those variables that discriminated endocrine tumours (E) from the others as shown to the top right of the graph (Figure 45). Also, it appears that the PC-DFA pattern in the positive mode appears similar to that in the negative mode, thereby implying that the negative mode plot was not completely over-fitted.
The PCA loadings scatter plot is used to investigate the relationships between the individual variables. The loadings of a given principal component represent the relative extent to which the original variables influence the PC and thereby express the correlation coefficients between the original variables and the computed PC. A high correlation between PC1 and a variable indicates that the variable is associated with the direction of the maximum amount of variation in the dataset. More than one variable might have a high correlation with PC1. A strong correlation between a variable and PC2 indicates that the variable is responsible for the next largest variation in the data perpendicular to PC1 and so on. The loadings plot depicts the significant variables located at the extremes of the scatter plot accounting for the most variation between components. This is represented in the positive and negative scatter plots below with blue dots indicating the components and outlying variables represented with a pink circle around the blue dots.

In the negative mode, the variables highlighted with the red circles had a p value of less than 0.001. However, none of these variables had a false discovery rate (FDR) lower than 0.1. Also, there was little agreement between the PC-DFA loadings and the ANOVA test suggesting that the PC-DFA model was over-fitted (Figure 46). In the positive mode, variables highlighted with red
circles were identified as significant at the level of false discovery rate (FDR) ≤0.1 and using one-way ANOVA test. The ANOVA test seemed to favour those variables that discriminated endocrine tumours (E) from the others as shown to the top right of the graph (Figure 47).

Figure 46: PC-DFA loadings plot, negative mode with significant components highlighted with pink circles

Figure 47: PC-DFA loadings plot, positive mode with significant components highlighted with pink circles
4.10.4. PCA- Varimax Rotated Scores Plot

To determine if there was indeed a cause for the variation between the tumour groups in the positive mode, a PCA-Varimax rotated scores plot was performed on the variables in the positive mode. The varimax rotation aims to further analyze the PCA results with the goal of making the pattern of loadings clearer. A varimax rotation analysis assumes that the factors in the analysis are uncorrelated and clusters the PCA loadings into several groups and attempts to maximise the influence of one particular factor while minimise those from the others. The results observed in PC-DFA plot can be reproduced via a varimax rotation from the PCA loadings. This shows there was indeed a factor contributing to the separation between the 3 classes. The figure demonstrates that the endocrine cancers varied significantly from those of pancreatic and ampullary cancers (Figure 48). When PCA varimax rotated scores were plotted a ‘factor’ that could account for the significant variables in the endocrine tumour group was observed.

![PCA-Varimax rotated scores plot,Positive mode](image)

Figure 48: PC-DFA varimax rotation plot indicating a separation between the three classes

The Varimax rotation plots demonstrate that the endocrine cancers had a significantly different variation from those of pancreatic and ampullary cancers. Biologically, this would seem consistent, given that both ampullary and pancreatic cancers share the same characteristics in the type of cancer (adenocarcinoma), mode of presentation and spread. In contrast, endocrine
neoplasms originate from a different cell type and behave significantly differently. This could account for the metabolomic variation seen on the PC-DFA loadings and appear to be consistent on the varimax rotation plots. Also, the metabolomic profile shows that components in the positive ion mode account for the differences between the cancer types and we explore this in metabolite plots to assess for significance between the metabolites in the three cancer subtypes.

4.10.5. Variables of interest

To assess the significant metabolites further, box and whisker plots were created for each significant variable (both positive and negative).

The positive mode data generated a total of 39 variables that were identified as "significant" at the level of FDR<0.1 with a good agreement between the PC-DFA loadings and the univariate test. The results of PC-DFA were reproduced, in an unsupervised manner, using the Varimax rotation which confirming that the pattern discovered by PC-DFA was present.

In contrast, the presence of an unknown a factor in the negative ion mode split the case samples into two groups, with progressive separation seen in PCA performed on the full data with control and quality control samples compared with PCA on the case samples alone. No correlation between the group and the metadata and lack of a split in the positive ionization model pointed to an analytical cause (likely over-fitting of the model) accounting for an inability to discriminate between the three cancer types using the negative mode data.

4.10.5.1. Positive mode results

A total of 39 significant variables were identified after statistical analysis. These are shown in the table below with the m/z values and level of significance.
<table>
<thead>
<tr>
<th>Peak Align</th>
<th>m/z</th>
<th>RT</th>
<th>MPA</th>
<th>P_A (zero) p val</th>
<th>P_E (zero) p val</th>
<th>A_E (zero) p val</th>
<th>P_A (NaN) p val</th>
<th>P_E (NaN) p val</th>
<th>A_E (NaN) p val</th>
<th>Metabolites</th>
</tr>
</thead>
<tbody>
<tr>
<td>619</td>
<td>200.8592557</td>
<td>39.20895</td>
<td>1388583.536</td>
<td>0.3413</td>
<td>0.0009</td>
<td>0.3413</td>
<td>0.0009</td>
<td>0.0031</td>
<td></td>
<td></td>
</tr>
<tr>
<td>619</td>
<td>226.1799957</td>
<td>551.114</td>
<td>143817.1874</td>
<td>0.1904</td>
<td>0.6641</td>
<td>0.1818</td>
<td>NaN</td>
<td>NaN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>619</td>
<td>254.6470342</td>
<td>707.3493</td>
<td>143817.1874</td>
<td>0.1904</td>
<td>0.6641</td>
<td>0.1818</td>
<td>NaN</td>
<td>NaN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>619</td>
<td>1107</td>
<td>254.6470342</td>
<td>707.3493</td>
<td>0.1904</td>
<td>0.6641</td>
<td>0.1818</td>
<td>NaN</td>
<td>NaN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>619</td>
<td>1145</td>
<td>254.6470342</td>
<td>707.3493</td>
<td>0.1904</td>
<td>0.6641</td>
<td>0.1818</td>
<td>NaN</td>
<td>NaN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>619</td>
<td>1810</td>
<td>254.6470342</td>
<td>707.3493</td>
<td>0.1904</td>
<td>0.6641</td>
<td>0.1818</td>
<td>NaN</td>
<td>NaN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>619</td>
<td>1877</td>
<td>254.6470342</td>
<td>707.3493</td>
<td>0.1904</td>
<td>0.6641</td>
<td>0.1818</td>
<td>NaN</td>
<td>NaN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>619</td>
<td>1892</td>
<td>254.6470342</td>
<td>707.3493</td>
<td>0.1904</td>
<td>0.6641</td>
<td>0.1818</td>
<td>NaN</td>
<td>NaN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>619</td>
<td>1854</td>
<td>254.6470342</td>
<td>707.3493</td>
<td>0.1904</td>
<td>0.6641</td>
<td>0.1818</td>
<td>NaN</td>
<td>NaN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>619</td>
<td>1810</td>
<td>254.6470342</td>
<td>707.3493</td>
<td>0.1904</td>
<td>0.6641</td>
<td>0.1818</td>
<td>NaN</td>
<td>NaN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>619</td>
<td>2145</td>
<td>254.6470342</td>
<td>707.3493</td>
<td>0.1904</td>
<td>0.6641</td>
<td>0.1818</td>
<td>NaN</td>
<td>NaN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>619</td>
<td>2456</td>
<td>254.6470342</td>
<td>707.3493</td>
<td>0.1904</td>
<td>0.6641</td>
<td>0.1818</td>
<td>NaN</td>
<td>NaN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>619</td>
<td>2491</td>
<td>254.6470342</td>
<td>707.3493</td>
<td>0.1904</td>
<td>0.6641</td>
<td>0.1818</td>
<td>NaN</td>
<td>NaN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>619</td>
<td>2536</td>
<td>254.6470342</td>
<td>707.3493</td>
<td>0.1904</td>
<td>0.6641</td>
<td>0.1818</td>
<td>NaN</td>
<td>NaN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>619</td>
<td>2586</td>
<td>254.6470342</td>
<td>707.3493</td>
<td>0.1904</td>
<td>0.6641</td>
<td>0.1818</td>
<td>NaN</td>
<td>NaN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>619</td>
<td>2663</td>
<td>254.6470342</td>
<td>707.3493</td>
<td>0.1904</td>
<td>0.6641</td>
<td>0.1818</td>
<td>NaN</td>
<td>NaN</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 18 - Significant positive ion variables

<table>
<thead>
<tr>
<th>Peak Align</th>
<th>m/z</th>
<th>RT</th>
<th>MPA</th>
<th>P_A (zero) p val</th>
<th>P_E (zero) p val</th>
<th>A_E (zero) p val</th>
<th>P_A (NaN) p val</th>
<th>P_E (NaN) p val</th>
<th>A_E (NaN) p val</th>
<th>Metabolites</th>
</tr>
</thead>
<tbody>
<tr>
<td>619</td>
<td>200.8592557</td>
<td>39.20895</td>
<td>1388583.536</td>
<td>0.3413</td>
<td>0.0009</td>
<td>0.3413</td>
<td>0.0009</td>
<td>0.0031</td>
<td></td>
<td></td>
</tr>
<tr>
<td>619</td>
<td>226.1799957</td>
<td>551.114</td>
<td>348795.7803</td>
<td>0.0004</td>
<td>0.8964</td>
<td>0.0015</td>
<td>0.0004</td>
<td>0.8964</td>
<td></td>
<td></td>
</tr>
<tr>
<td>619</td>
<td>254.6470342</td>
<td>707.3493</td>
<td>143817.1874</td>
<td>0.1904</td>
<td>0.6641</td>
<td>0.1818</td>
<td>NaN</td>
<td>NaN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>619</td>
<td>2586</td>
<td>254.6470342</td>
<td>707.3493</td>
<td>0.1904</td>
<td>0.6641</td>
<td>0.1818</td>
<td>NaN</td>
<td>NaN</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 18 - Significant positive ion variables

<table>
<thead>
<tr>
<th>Peak Align</th>
<th>m/z</th>
<th>RT</th>
<th>MPA</th>
<th>P_A (zero) p val</th>
<th>P_E (zero) p val</th>
<th>A_E (zero) p val</th>
<th>P_A (NaN) p val</th>
<th>P_E (NaN) p val</th>
<th>A_E (NaN) p val</th>
<th>Metabolites</th>
</tr>
</thead>
<tbody>
<tr>
<td>619</td>
<td>200.8592557</td>
<td>39.20895</td>
<td>1388583.536</td>
<td>0.3413</td>
<td>0.0009</td>
<td>0.3413</td>
<td>0.0009</td>
<td>0.0031</td>
<td></td>
<td></td>
</tr>
<tr>
<td>619</td>
<td>226.1799957</td>
<td>551.114</td>
<td>348795.7803</td>
<td>0.0004</td>
<td>0.8964</td>
<td>0.0015</td>
<td>0.0004</td>
<td>0.8964</td>
<td></td>
<td></td>
</tr>
<tr>
<td>619</td>
<td>254.6470342</td>
<td>707.3493</td>
<td>143817.1874</td>
<td>0.1904</td>
<td>0.6641</td>
<td>0.1818</td>
<td>NaN</td>
<td>NaN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>619</td>
<td>2586</td>
<td>254.6470342</td>
<td>707.3493</td>
<td>0.1904</td>
<td>0.6641</td>
<td>0.1818</td>
<td>NaN</td>
<td>NaN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peak Align</td>
<td>m/z</td>
<td>RT</td>
<td>MPA</td>
<td>P_A (zero) p val</td>
<td>P_E (zero) p val</td>
<td>A_E (zero) p val</td>
<td>P_A (NaN) p val</td>
<td>P_E (NaN) p val</td>
<td>A_E (NaN) p val</td>
<td>Metabolites</td>
</tr>
<tr>
<td>-----------</td>
<td>----------</td>
<td>--------</td>
<td>---------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>3280</td>
<td>430.6965414</td>
<td>39.54725</td>
<td>1244005.885</td>
<td>0.1242</td>
<td>0.0017</td>
<td>0.0450</td>
<td>0.1242</td>
<td>0.0017</td>
<td>0.0450</td>
<td>Isotope</td>
</tr>
<tr>
<td>3308</td>
<td>432.6934956</td>
<td>39.5509</td>
<td>1191636.65</td>
<td>0.1500</td>
<td>0.0017</td>
<td>0.9526</td>
<td>0.2941</td>
<td>0.0051</td>
<td>0.0295</td>
<td>Isotope</td>
</tr>
<tr>
<td>3345</td>
<td>434.690633</td>
<td>39.5074</td>
<td>590272.1347</td>
<td>0.13668</td>
<td>0.0017</td>
<td>0.0295</td>
<td>0.1366</td>
<td>0.0017</td>
<td>0.0295</td>
<td>Isotope</td>
</tr>
<tr>
<td>3401</td>
<td>438.7282698</td>
<td>37.99565</td>
<td>888076.3964</td>
<td>0.25151</td>
<td>0.0017</td>
<td>0.0295</td>
<td>0.2515</td>
<td>0.0017</td>
<td>0.0295</td>
<td>Multiply Chlorinated adducts - not identified</td>
</tr>
<tr>
<td>3431</td>
<td>440.7252919</td>
<td>37.9994</td>
<td>1781868.352</td>
<td>0.17964</td>
<td>0.0027</td>
<td>0.0450</td>
<td>0.1796</td>
<td>0.0027</td>
<td>0.0450</td>
<td>Isotope</td>
</tr>
<tr>
<td>3476</td>
<td>444.9160621</td>
<td>39.1144</td>
<td>153231.4581</td>
<td>0.1795</td>
<td>0.0509</td>
<td>0.0015</td>
<td>0.0481</td>
<td>0.0227</td>
<td>0.0015</td>
<td>Unmatched Adduct</td>
</tr>
</tbody>
</table>
| 3923      | 477.2987922 | 740.3249 | 2808619.901 | 0.0232          | 0.1038          | 0.0062          | 0.0232          | 0.1038          | 0.0062          | 13'-carboxy-alpha-tocotrienol;"24a,24b-epoxy-23-tetradehydr
-24a,24b-dihomo-1alpha,25-dihydroxyvitamin D3 / 24a,24b-epoxy-23-tetradehydr
24a,24b-dihomo-1alpha,25-dihydroxycholecalfierol;Ubiquinone Q4;","27-nor-5b-
cholestane-3a,7a,12a,24,25-pentol","N'-5Z,8Z,11Z,14Z-
eicosatetraenoyl-N'-diethyl-
<p>| 4079      | 490.6521595 | 39.5555 | 622706.2633  | 0.1366          | 0.0014          | 0.0108          | 0.1366          | 0.0014          | 0.0108          | Isotope                                                                 |
| 4170      | 507.1113889 | 74.10364 | 1159552.633 | 0.2319          | 0.0017          | 0.0450          | 0.4207          | 0.0051          | 0.0450          | Multiply Chlorinated adducts - not identified                              |
| 4200      | 498.6837992 | 37.95915 | 988294.3851  | 0.2319          | 0.0017          | 0.0450          | 0.2319          | 0.0017          | 0.0450          | Isotope                                                                 |</p>
<table>
<thead>
<tr>
<th>Peak Align</th>
<th>m/z</th>
<th>RT</th>
<th>MPA</th>
<th>P_A (zero) p val</th>
<th>P_E (zero) p val</th>
<th>A_E (zero) p val</th>
<th>P_A (NaN) p val</th>
<th>P_E (NaN) p val</th>
<th>A_E (NaN) p val</th>
<th>Metabolites</th>
</tr>
</thead>
<tbody>
<tr>
<td>4206</td>
<td>499.0327959</td>
<td>43.8009</td>
<td>169465.7577</td>
<td>0.0232</td>
<td>0.0591</td>
<td>0.0062</td>
<td>0.1366</td>
<td>0.0111</td>
<td>0.0062</td>
<td>Bindarit (USAN/INN); Ditazole (INN); Dolasetron; Oxyphenbutazone; Glyoctamide</td>
</tr>
<tr>
<td>4227</td>
<td>500.6809774</td>
<td>37.9489</td>
<td>927203.7405</td>
<td>0.2133</td>
<td>0.0011</td>
<td>0.0186</td>
<td>0.2133</td>
<td>0.0011</td>
<td>0.0186</td>
<td>Isotope</td>
</tr>
<tr>
<td>4408</td>
<td>514.2877194</td>
<td>726.5711</td>
<td>4746257.328</td>
<td>0.0078</td>
<td>0.0160</td>
<td>0.0031</td>
<td>0.0078</td>
<td>0.0160</td>
<td>0.0031</td>
<td>12-O-2Z, 4E-Octadienoyl-4-deoxyphorbol 13-acetate; Cyclobuxine D</td>
</tr>
<tr>
<td>4843</td>
<td>546.6138134</td>
<td>39.5274</td>
<td>277224.6623</td>
<td>0.7142</td>
<td>0.0014</td>
<td>0.0186</td>
<td>NaN</td>
<td>0.0042</td>
<td>0</td>
<td>Multiply Chlorinated adducts - not identified</td>
</tr>
<tr>
<td>4873</td>
<td>548.6108985</td>
<td>39.5563</td>
<td>359239.0036</td>
<td>0.1794</td>
<td>0.0017</td>
<td>0.0108</td>
<td>0.8643</td>
<td>0.0317</td>
<td>0.0108</td>
<td>Isotope</td>
</tr>
<tr>
<td>4996</td>
<td>556.6426491</td>
<td>37.9143</td>
<td>573690.5708</td>
<td>0.2515</td>
<td>0.0017</td>
<td>0.0186</td>
<td>0.2515</td>
<td>0.0017</td>
<td>0.0186</td>
<td>Isotope</td>
</tr>
<tr>
<td>5025</td>
<td>558.639677</td>
<td>37.91925</td>
<td>647501.088</td>
<td>0.1959</td>
<td>0.0014</td>
<td>0.0450</td>
<td>0.1959</td>
<td>0.0014</td>
<td>0.0450</td>
<td>Isotope</td>
</tr>
<tr>
<td>5056</td>
<td>560.6367712</td>
<td>37.87845</td>
<td>394168.3703</td>
<td>0.1368</td>
<td>0.0022</td>
<td>0.0186</td>
<td>0.1366</td>
<td>0.0022</td>
<td>0.0186</td>
<td>Isotope</td>
</tr>
<tr>
<td>5754</td>
<td>616.598446</td>
<td>37.59805</td>
<td>427246.9442</td>
<td>0.5747</td>
<td>0.0017</td>
<td>0.0108</td>
<td>0.5747</td>
<td>0.0017</td>
<td>0.0108</td>
<td>Isotope</td>
</tr>
<tr>
<td>5482</td>
<td>595.3491498</td>
<td>594.9567</td>
<td>1182556.913</td>
<td>0.6406</td>
<td>0.6476</td>
<td>0.9681</td>
<td>NaN</td>
<td>NaN</td>
<td>1</td>
<td>L-Urobilin, Aplaviroc, Lupeol acetate</td>
</tr>
<tr>
<td>6575</td>
<td>684.5858633</td>
<td>37.4427</td>
<td>342412.2537</td>
<td>0.8643</td>
<td>0.0007</td>
<td>0.0015</td>
<td>0.8643</td>
<td>0.0007</td>
<td>0.0015</td>
<td>Isotope</td>
</tr>
<tr>
<td>7547</td>
<td>775.9853404</td>
<td>743.7402</td>
<td>1281728.007</td>
<td>0.0264</td>
<td>0.6027</td>
<td>0.3543</td>
<td>0.0264</td>
<td>0.6027</td>
<td>0.3543</td>
<td>Multiply Charged ion - not identified</td>
</tr>
</tbody>
</table>
4.10.5.2. Significant metabolites based on cancer type

Box and whisker plots have been used to graphically depict the distribution of the variable concerned between the three main cancer groups. Some variables have been catalogued below to demonstrate the differences between the three tumour types. As can be seen below, there are a large number of metabolites that have not been sufficiently classified. Metabolites vary from each other based on their physical and chemical properties; examples include molecular weight, size, volatility and their hydrophilic or hydrophobic nature. Once detected on metabolic profiling (GC-MS), a limiting factor is often identification of the metabolite. Conversion of the raw chemical data to a biological metabolite requires the registration of the chemical nature of the compound to match with biological compounds previously detected and recorded in a database. Although there is progress being made in the development of these spectral databases, unknown compounds predominate in studies of biological substrates.

In the section below, various examples of metabolites found to be significantly different in each of the three cancers are represented. It can be observed that often no clear biological metabolite has been characterized for the composition of the chemical compound detected. The purpose of the box and whisker plots is to emphasise the differences between the tumour types and to demonstrate the feasibility of detecting metabolomic differences between them. The metabolites are divided according to the cancer types as opposed to the chemical nature of the compounds (positive or negative ion).

1. Metabolites significantly altered in endocrine cancer

In addition to the number of unclassified metabolites (salt complex/ unidentified metabolites), several other metabolites including organic acids, oxido-reductases and the hexose monophosphate pathway were identified in this group. No obvious pattern of predominance from any specific molecular pathway could be identified. In view of the fact that this study attempted an untargeted or profiling approach, if we had identified a specific pathway or biological role dominating the analysis, consideration could have been given to using a more targeted approach (for example: lipidomics to identify changes in lipid metabolites) towards the identification of the metabolites.
1. **Variable No 619**
   Chemical composition: Na, K, Cl, HCOOH complex (unidentified metabolite).

2. **Variable ID 1145**
   Chemical composition: Na, K, Cl, HCOOH complex (unidentified metabolite)

3. **Variable ID 336 (negative ion)**
   Chemical composition: 2-Hydroxyethanesulfonate.
   Family: Organic acids. Location: cytoplasm
4. Variable ID 189 (negative ion)

Chemical composition: 2-Deoxy-D-gluconate; 2-deoxygluconate;

Biological role: Pentose phosphate pathway

2. Metabolites significantly different in ampullary cancer

In the group of metabolites found to be significant in ampullary cancer, we found metabolites playing roles in cellular oxidation, fatty acid metabolism and Vitamin D metabolism. Examples of these metabolite classes are displayed below.

1. Variable ID 827

Chemical composition: 3-hydroxy-9-apo-beta-caroten-9-one

Metabolite class: Fatty acids and conjugates.
2. Variable ID 3923

Chemical composition: 13'-carboxy-alpha-tocotrienol

Metabolite class: Vitamin D2 and derivatives.

3. Variable ID 194

Chemical composition: 2-Deoxy-D-gluconate

Biological role: oxidoreductases

It must be emphasised that despite significant variables from these three classes being identified, the data in the absence of controls, cannot be extrapolated into clinically applicable markers. At best, this raises the focus of enquiry into these metabolite classes in ampullary cancer. Given that these metabolite classes are actively involved in normal physiology, the significance of the results needs clarification with adequate controls.
4.11. Discussion

We incorporated a series of patients with pancreatic, ampullary and endocrine cancers to assess the potential for assessment of the metabolomic profile in these conditions. This is the first study of its type to assess the metabolomic profile of ampullary and endocrine cancers. Clinically, there is a need to integrate genetic, proteomic and metabolomic techniques in the diagnosis and management of pancreatic cancer due to the inevitable delay in diagnosis that prevents early treatment. Furthermore, as surgery is the only therapeutic modality shown to reduce mortality in peri-ampullary tumours, a delay in diagnosis prevents a substantial proportion of patients from receiving curative resections. This research project explored the potential use of LC-MS in serum from patients with this category of tumours. Previous studies have used NMR spectroscopy and GC-MS to characterize the metabolome of pancreas cancer however there are no reports of the use of LC-MS in serum samples to study pancreatic cancer.

4.11.1. Patients and methodology

We used LC-MS to screen serum samples from this patient population with the preliminary aim of identifying significant metabolites. The patient group represented all patients with diagnosed pancreatic cancer, other peri-ampullary cancers (ampullary cancer, cholangiocarcinoma and pancreatic endocrine tumours) and one patient with a pancreatic cystic tumour. We also included two patients with chronic pancreatitis. A common feature in the patient population was the presence of a low to normal body mass index related possibly to the catabolic activity of the disease process. The study group also included patients regardless of tumour stage and operability. All patients were sampled prior to any intervention to minimize influence in metabolomic outcomes. The tumours were grouped into three categories based on the number of patients, (pancreas cancer, ampullary tumours and endocrine tumours). The three other patient groups (chronic pancreatitis, cholangiocarcinoma and pancreatic cystic tumour) were not taken into account in the analysis because of a small number of patients. The average BMI in our group was 21 (range 18-26), though it must be emphasized that most patients with newly diagnosed pancreatic cancer will have lost a proportion of their body weight prior to the diagnosis. The transitional loss of weight rather than the BMI itself is felt to be a factor influencing metabolomic profiles in these patients. The study initially planned to recruit control samples from a set of age,
body mass index and gender matched controls, however, due to a inability to recruit a suitable number ‘normal’ matched individuals (with the primary reason being a lack of BMI and age matching), we resorted to using age, sex and BMI controls from the HUSERMET study. All samples were collected and transported to the lab in ice with the aim of inhibiting metabolism within the sample. All samples were stored at -80 degrees and this is felt sufficient to completely preserve the metabolic status of the sample (362). Sample degradation is a valid concern with stored samples. Samples generally processed within 3 months are accepted as being satisfactory for analysis. All our collected samples satisfied this criterion. The HUSERMET control samples were stored samples and had previously been collected and stored from a period of 5 years previously. The degradation of sample quality over that period of time is unknown and could potentially affect the results. In one study, the short-term storage of blood, hemolysis, and short-term storage of non-cooled plasma resulted in statistically significant increases of 4% to 19% and decreases of 8% to 12% of the metabolites detected (363). To add to the complexity, substrates such as cerebrospinal fluid or urine may show no sample degradation when stored (364, 365). The HUSERMET (Human Serum metabolome in health and disease) project aimed to develop methods to characterize the human metabolome and standardize standard operating procedures (SOP’s) for research. The project was a 4-year collaboration between the University of Manchester, Astra Zeneca and Glaxo SmithKline and recruited several thousand volunteers to provide serum samples. We selected matched controls for the study subjects from the database and used them to perform our analysis.

4.11.2. Comparison between the cases and controls

Our results showed a significant disjunction between the patient population and the HUSERMET controls. This has impeded a reliable interpretation of the metabolite profiles within the group. The recruitment of control samples from matched individuals with no pathology was the primary stumbling block preventing the concomitant collection of ‘normal’ and ‘disease’ samples. The HUSERMET controls, despite being matched, have shown a variation in metabolite peaks. We studied this disparity further by analyzing two-thirds of the peaks that the controls and the cases shared. The graphs showed a closer relationship between the two, however, not enough to permit the data to be used together. The cause for this incongruence is not clear, but is perhaps related to metabolite degradation in the stored control samples. Another reason for this discrepancy may
be related to the affected patient demographic. Pancreas cancer produces a catabolic metabolic state with a subset of patients having symptoms such as jaundice and diabetes that alter the metabolic profile. Most patients will have a low to normal body mass index as well. Controls for the study would also require to be matched to exclude other metabolic derangements that could confound results. Another feature that could perhaps be exploited is the progression in understanding the genetic framework of pancreas cancer. Genetic changes can be measured through their downstream products and vice versa. This could enable a targeted analysis to identify metabolite differences between controls and cancers. In addition, a subgroup analysis of pancreas cancer metabolic data, based on a tumour stage progression, particularly early stage pancreas cancer, may assist in identifying metabolites that distinguish the cases and controls. This could perhaps minimize the metabolic effects of advanced pancreatic cancer and represent a truer difference between a normal and disease metabolome. Our study included an ‘all stage’ approach to patients with pancreas cancer and peri-ampullary malignancy and the metabolite snapshot produced included a composite of both early (operable) and locally advanced and metastatic (inoperable) cancer. This could confound interpretation by the introduction of additional metabolite variables that may not be specific to pancreas cancer.

So, despite an interesting collection of data, the disparity in peaks between the two groups precludes reliable interpretation of the dataset and therefore, the study results cannot be exported to the development of provisional metabolic biomarkers. It does however; open avenues for further assessment using a fresh set of matched cases and collected controls. This could possibly remove the discrepancy produced by the use of database controls and permit a closer examination of the preliminary findings from this study.

4.11.3. Inter-group analysis

We performed an inter-group analysis using PC-DFA to assess the metabolomic variation between the three tumour types. The analysis has shown up some interesting outcomes with a difference in positive mode metabolites between the endocrine group and the peri-ampullary tumours (pancreatic and ampullary). This is an interesting finding given that pancreatic and ampullary tumours behave in a similar manner clinically, with endocrine tumours manifesting with a different clinical profile. The endocrine tumours were all non-functioning tumours and therefore
were not expected to have a significantly different profile to the peri-ampullary cancers. The negative ionization mode metabolites were not significant on evaluating the false discovery rate perhaps due to the tightness of the analytical model used in the study. There was also a factor involved that also discriminated between the components on PCA in the negative ionization mode. This was difficult to account for with no clear cause other than the ionization mode, as there was no such split seen in components from the positive mode.

A greater sample size may allow us to clarify if this is a true difference or related to the analytical mode. Furthermore, the collection of repeat samples from the same groups would facilitate a better correlation of metabolite profiles between the groups. The results demonstrate that assessing the metabolome of endocrine tumours could be potentially useful in identifying markers to differentiate this subclass of pancreatic cancer from ductal adenocarcinoma. It remains to be seen if the negative ion metabolites play a significant role as well and further studies will be needed to clarify this.

4.11.4. Significant metabolites

We found significant metabolites in the positive ion mode as seen in the graphs above. A significant number of these metabolites have not been clearly identified. They include a number of isotopes, salt complexes, fatty acid complex and metabolites from glucose metabolism. The negative ion metabolites that were found to be significant initially also included members of the fatty acid complex and glucose metabolism. It is encouraging to find that there are metabolites that differentiate between the classes, however, to characterize these metabolites further controls would be required. Previous studies have shown alterations in glucose metabolism, lipid content, lysine, isoleucine and d-glucuronic acid metabolism (320, 322, 324, 327). The identified metabolites from our study concur with those above, however, as previously stated, these results would need further validation based on concurrently collected controls to clarify the metabolite pathways involved.

In conclusion, the logical progression following studies of the genome and metabolome in pancreas cancer is the development of markers to aid in facilitating an early diagnosis and assist in follow up of patients. Currently, despite the wealth of genetic information available, it is clear
there is some way to go before we can utilize this knowledge in clinical management of these
patients. In addition, the metabolic end product of these genetic changes is yet to be explored. An
identification of the end product of the genetic command is the next logical step to gain an
appreciation of the altered cellular process and is accomplished by studying the metabolites
generated in the cell. Preliminary metabolomic studies are encouraging, however, larger studies
incorporating a variety of sample types need to be performed. The future objective of further
studies would be to explore and develop clinical markers that could be applied to clinical practice
in this group of patients. The likelihood is that a panel of markers incorporating both genetic and
metabolic signatures could fulfill this need.
Conclusions and Future work

A registry of pancreatic cystic tumours

Pancreatic cystic tumours are being detected with increasing frequency. This is in part due to improved and more frequent use of imaging studies (366-369), but also because we are living longer lives. In the UK, pancreatic cancer accounted for 5% of all cancer deaths in men and 6% of all cancer deaths in women in 2012 (1). The detection of pancreas cancer at an earlier stage could potentially lead to early treatment and better survival. Pancreatic cystic tumours, in particular IPMNs, are recognized precursors of pancreas cancer and now account for between 10-20% of all pancreatic resections at high volume centres (169, 182, 370). The early identification and treatment of high-risk patients with pancreatic cystic lesions likely to progress to pancreas cancer is essential. Cystic lesions with low to moderate risk of progression need to be accurately classified and assessed sequentially to assess progression. There is insufficient data regarding the conservative management of both low to medium risk cystic lesions, with certain centres opting for more aggressive resectional policy (371, 372). Surgery in this group of elderly patients carries significant risks. If the risk of progression of the lesion is lower than the operative morbidity and mortality of major pancreatic surgery, then surgical intervention can be targeted towards those patients that actually need it, rather than an ‘all comers’ approach. There is a lack of clarity on issues such as which features are accurate markers of disease progression, how rapidly do these lesions progress, what interval of follow up is ideal and is early surgery appropriate given the increasing risk of operative intervention in progressively older patients (373, 374). In order to identify the correctly classify suspected pancreatic cystic tumours and determine their natural history, we developed a database of patients presenting with pancreatic cystic tumours. The aim was to define the clinical presentation and investigation of these patients using a management plan to accurately characterize these lesions. In addition, we aimed to chronicle the rate of progression to adenocarcinoma in patients with IPMN. Most clinicians agree that main duct IPMN should be treated as malignant lesions and warrant surgery (225, 375). The goal of surgery is to perform a complete resection (R0) of the IPMN and intraoperative frozen sections have been recommended to assist with this (376). Despite this, the presence of discontinuous lesions may lead to the development of adenocarcinoma in the remnant pancreas (182). In our
study, the overall population of suspected cystic tumours showed an incidence of adenocarcinoma of 8%; however, the likelihood of cancer changed significantly if the lesions were categorized based on pathology. In our group, main duct IPMN showed a cumulative risk of 46% with evidence of progression of disease in a further 23% clearly corroborating the necessity for surgery in all MD-IPMN to prevent cancer developing or progressing within these lesions. The associated mortality in MD-IPMN is related to the underlying adenocarcinoma and was 38% in our group occurring primarily in patients with adenocarcinoma. We experienced a high rate of R1 resections (5/6) in the group of operated MD IPMNs with adenocarcinoma. We did not identify a cause that could account for this. With the increasing resection of MD IPMN, consideration needs to be given to the remnant pancreas. MD IPMN has been proposed as being multicentric, increasing the risk of malignancy in the remnant pancreas (182, 377). A possible means of identifying disease in the remnant is using frozen section; however, controversy exists regarding the extent of resection to which these patients should be subjected. However, the morbidity and mortality that a total pancreatectomy entails has tempered the surgical management of the residual pancreas. Future work in the surgical management of main duct IPMN could include the following: 1) What should the resection limit be when dysplasia is seen on the frozen section? 2) How frequently should the remnant pancreas be assessed with imaging? 3) Is there a role for a prophylactic total pancreatectomy and lymphadenectomy in this group? Also, in those patients deemed unsuitable for surgery, questions exist on the role of chemotherapy, the role of surveillance and rate of progression to cancer.

Branch duct IPMN particularly those with a size less than 3 cm present a dilemma in terms of optimal management given the relatively low risk of malignancy. The database focused on these lesions to assess their progression, rate to surgery and risk of malignancy over the period of this study. The follow up of these lesions is decided by clinical judgment based on the age of the patient, comorbidities, risk of progression to cancer and patient preference, with little evidence in the literature to support a specific type or frequency of investigation in this group. The proposed annual rate of malignancy has been proposed to be about 2-3% in branch duct IPMN (274, 378). We demonstrated that the demographics of our population group were consistent with previously published studies and used a management algorithm based on international guidelines current at the time of the study to allocate patients into the various categories (224). We found that despite
the number of imaging modalities used to identify the lesions, we found a false positive rate of
20% of benign lesions using just CT scans alone, making it imperative that suspected cystic
neoplasms are investigated by a combination of CT, MRCP and EUS. Given that occasionally CT
scans will not distinguish clearly between branch duct IPMN and MCN, MRCP should be used
more often to delineate the nature of these lesions. In lesions of a size less than 3cm, a
conservative approach in the absence of suspicious signs could be warranted (379). However,
there are increasing reports of branch duct IPMN smaller than 3cm can harbor malignancy in up
to 20% of cases (380, 381). Future work regarding surgery for branch duct IPMN should aim to
include the following: 1) Is cyst size alone a good surrogate for predicting the risk of malignancy
or should all branch duct IPMN be scored using a number of radiological and cytological markers
2) Should we operate on younger patients with BD-IPMN less than 3cm 3) What is the clinical
course and rate of progression to cancer of lesions that are larger than 3cm in size and can we
safely observe them given the elderly population in whom they occur 4) And lastly, can these
lesions be removed using non-anatomical resections to minimize the operative morbidity. The
incidence of adenocarcinoma in branch duct IPMN was 11% in our group with disease
progression seen 13.8%. This compares with findings from several cohort studies that suggest
that disease progression is observed in between 10-40% of patients over 5-years with the risk of
progression to adenocarcinoma of about 20% (224, 373, 374, 382). In cystic lesions with a size
less than 3 cm, the risk was directly related to the underlying pathology. In resected BD IPMN, a
proportion of lesions maybe multifocal with a risk of cancer in the remnant pancreas and these
patients should also be followed up. The differentiation of BD-IPMN from other lesions is
important in view of the relatively low risk of cancer and need for surveillance. Future studies will
need to focus on the rate of growth of BD-IPMN cysts and follow up intervals and duration.
However, in our study, the BD IPMN cysts <3cm underwent active surveillance with 4 out of 36
patients (11%) developing high-risk features over 24 months. Admittedly, this period of follow up
may not be enough to accurately identify progression in all cases, but suggests that with
adequate monitoring, BD IPMN are amenable to conservative management. Another feature we
found in our patient group was the high incidence and prevalence of extra-pancreatic
malignancies with 37.7% of patients having a malignancy at some time pre or post diagnosis.
Extra-pancreatic malignancies occur with considerable frequency in IPMN (383, 384). In general,
80% of these EPM precede the development of IPMN, however, surveillance of IPMN should
include close monitoring to assess for their development. Given the high rate of EPM, future research should focus on a common genetic focus that generates this predisposition.

Aside from the improvements made in clinical diagnosis and management of IPMNs, important strides have been made in identifying the molecular pathology of IPMNs. Cyst fluid from pancreatic cysts is easily obtainable via EUS and can be used to differentiate in the diagnosis of BD IPMN and MCN. CEA levels in cyst fluid of greater than 192-200 ng/ml are approximately 80% specific for mucinous cysts (385, 386). The presence of Kras in cyst fluid has been shown to increase the likelihood of the cyst being mucinous (387, 388). Recently, the detection of a combination of Kras and GNAS mutations in cyst fluid have been shown to increase the likelihood of a mucinous cyst (200). Future work will need to focus on applying molecular techniques to these lesions both to understand the manner in which they progress to pancreatic cancer but also to treat them. The development of a national database would help in clarifying the natural history of these lesions. Large scale, multicentre studies focused on follow up, molecular diagnosis and prognostic factors in this increasing patient cohort would help clarify and standardize management of these lesions in the UK.

**Poly A RT-PCR measurement of indicator genes in pancreas cancer and pancreatic cystic tumours**

There is increasing recognition that the earlier diagnosis of pancreas cancer is the best way to improve outcomes from the disease. Despite improvements in surgery and postoperative care including chemotherapy, mortality from pancreas cancer has remained stubbornly high (389). As reviewed earlier in this thesis, huge strides are being taken in establishing the molecular profile of pancreas cancer with the aim of identifying a ‘molecular signature’ of the disease that can be applied to clinical practice. The translation of genetic research findings to use at the bedside is imperative to facilitate the early diagnosis of this disease. There has been research into using molecular biomarkers in various substrates such as serum and urine. In this thesis, the focus is on a means of using pancreatic juice as a surrogate marker for the identification of pancreas cancer using a set of genes previously identified as being over-expressed in pancreatic juice in these patients. Our preliminary hypothesis was based on the precept that pancreatic juice shares
close contact with pancreatic ductal adenocarcinoma cells leading to cell shedding into the juice. This would permit the measurement of over-expressed genes within these cells shed into pancreatic juice by the cancer. By studying the expression of a set of genes in pancreatic juice and the corresponding cancer tissue, we could infer the utility of pancreatic juice as a surrogate marker substrate if a positive correlation was found. In this study, we evaluated the correlation between over-expressed genes in pancreatic juice and tumour in patients with pancreas cancer and pancreatic cystic tumours. We compared our set of 18 marker genes to paired samples obtained in the same patient with pancreas cancer undergoing surgery. We demonstrated that one gene out of seventeen; MSLN (mesothelin) was significantly over-expressed in both pancreatic juice and tissue samples in the cohort of patients with pancreatic cancer, though it did not meet the cutoff set by the Bonferroni correction. Similarly, in the IPMN group, a single gene MMP7 showed an initial correlation between the pancreatic juice samples and the tumour tissue, which was not corroborated on assessing for significance based on testing for multiple samples. The low yield of statistically significant results comparing the over-expression of this carefully chosen set of genes in pancreas cancer may be related to several limitations with the study. 1) Firstly, there was no guarantee that pancreatic juice from cancer contained the shed cells and would necessarily be dependent on the sampling technique. The intraoperative sampling of pancreatic juice may not be ideal as the juice sample is collected distal to the tumour location. 2) RNA in pancreatic juice could as well originate from ductal epithelium as from stromal cells, normal pancreas or an inflammatory response to the tumour. The pancreatic juice was harvested from intra-operatively cut pancreas, thereby introducing an element of increased risk of non-tumoural RNA within the sample. 3) Activated endonucleases in pancreatic juice digest mRNA within the juice, with uncertainty regarding the quantity of degradation of mRNA occurring during pancreatic manipulation and prior to sampling. In an already pauci-cellular sample type, this may affect certain transcripts in an unpredictable manner. 4) A further enquiry based on the efficacy of the chosen subset of genes could be raised. It would be ideal to have explored the samples by testing a larger subset of genes but we were exposed to limitations of cost and resources. As the focus of this study was to assess the feasibility of using this technique as a clinical tool, it was felt that our gene selection would be sufficient for purpose.

Despite this, the results we have shown are encouraging, in that notwithstanding the issues in
sampling, it is possible to assess for a correlation between pancreatic juice and tissue. A previous study had demonstrated non-significant over-expression of ANXA2 in pancreas cancer within these paired sample types (290) and this study has shown MSLN and MMP7 to be non-significantly over-expressed in pancreatic juice in pancreas cancer and pancreatic cystic tumours respectively. Based on previous work and this study, there is no clear evidence that there is a consistent detection of over-expressed genes in pancreatic juice compared with that of pancreatic cancer tissue. In view of restrictions due to substrate properties (RNAse, poor cellular material) and uncertainty of mRNA origin, it is difficult to justify the routine use of pancreatic juice as a surrogate marker substrate for pancreas cancer and questions remain about whether pancreatic juice is the ideal substrate to carry forward these advances. Nonetheless, there are encouraging studies with pancreatic juice that may help determine its future role in facilitating the earlier diagnosis of the disease by translating these advances in molecular pathology to the bedside.

Certainly, substrates that are more easily obtained such as serum, plasma or stool should also be actively considered (390-393). However, due to its proximity to the actual tumour tissue and the relative frequency of access during investigations of pancreatic or biliary disease, pancreatic juice continues to remain an interesting substrate for investigation. Future investigation with regards to pancreatic juice may need to concentrate on several fronts. These include 1) Determining the utility of microRNA (miRNA) as candidate biomarkers. MicroRNA are relatively stable in pancreatic juice and can be easily isolated. Different varieties of miRNA exist in different pancreatic diseases and more than 20 miRNAs have been identified in pancreas cancer (394, 395). Pancreatic juice has also been investigated as a substrate for miRNA with encouraging results and future research could focus on candidates like miRNA-21 and miRNA-155 (396, 397). 2) The DNA methylation status of mucin profiles in pancreatic juice in pancreatic neoplasms has been shown to correspond with the expression profiles in pancreatic tissue (398). Analyzing the mucin profiles in pancreatic juice and pancreatic cystic neoplasms may help identify changes in the mucin profile during progression to cancer and would form an interesting basis in the investigation of precursors like IPMN progressing to invasive disease. 3) Using a panel of biomarker candidates such as neuronal pentraxin-2 gene (399), neutrophil gelatinase-associated lipocalin (NGAL), macrophage inhibitory cytokine-1 (MIC-1) that have also been suggested to be elevated in pancreatic juice in pancreas cancer (400). 3) Lastly, consideration should be given to the use of GNAS and KRas mutational assays in pancreatic juice in patients with mucinous
pancreatic cystic tumours to increase the accuracy of prediction of benign mucinous versus non-mucinous cysts as shown in several studies (401-403). Alone or in combination with expression of mucin profiles, this may help improve the accuracy of diagnosis of mucinous cysts as well as potentially identify lesions at risk of progression.

In summary, there are encouraging avenues of exploration in the detection of early pancreas cancer using pancreas juice as a substrate for markers albeit with certain issues as highlighted above. Certainly, our study has shown the feasibility of correlating over-expressed genes in juice with that of tumour tissue and in carrying this work forward, consideration of microRNA markers in juice would seem the next appropriate step towards developing a biomarker panel in pancreatic juice.

**Metabolomic assessment of pancreatic cancer**

The future development of biomarkers is likely to involve a combination of genetic, proteomic and metabolomic variables. The human metabolome contains a small number of metabolites and alterations in the metabolite profiles represent the downstream effects of the genetic changes. Variations in the metabolite levels from the norm represent a combination of genetic and environmental changes in the metabolome. Earlier in this thesis, pancreatic juice was used as a substrate for a set of gene markers comparing those with the same over-expressed genes in pancreas tissue in pancreas cancer and pancreatic cystic tumours. In keeping with a biomarker panel incorporating both genetic and metabolomic changes, we aimed to perform a preliminary assessment of the metabolome in a complementary group of patients with pancreas cancer. An untargeted approach using GC-MS as a metabolic profiling technique in serum samples would provide us with a biological snapshot of the metabolome in pancreas cancer. We aimed to compare a cohort of patients with pancreas cancer, with matched controls using age, sex and BMI as the matching parameters. However, given the elderly age group, recent drop in BMI, associated features such as cachexia, anaemia and jaundice, we struggled to recruit appropriate controls from healthy volunteers. We eventually used age and sex matched controls stored for the HUSERMET study conducted by a group at Manchester University, as part of the investigation of the normal metabolome. This could undoubtedly confound any interpretation of the metabolomic profile from our patient group. Our results showed a significant disjunction
between the metabolite profiles between the patient population and the HUSERMET controls. The absence of accurate controls affects the reliable interpretation of alterations in the metabolite profiles within the study group. The HUSERMET controls, despite being matched, showed a variation in metabolite peaks and this could possibly have been related to metabolite degradation in the stored control samples. We also included patients with pancreas cancer and peri-ampullary malignancy and this could present a mixed metabolite snapshot and confound interpretation by the introduction of additional metabolite variables not be specific to pancreas cancer. So, despite an interesting collection of data, we could not export the study results to the development of provisional metabolic biomarkers. We did however show that there is an interesting distinction between the metabolite profiles from endocrine and ampullary cancer compared to pancreas cancer. A significant number of these metabolites were not clearly identified. The metabolites we found included a number of isotopes, salt complexes, fatty acid complex and metabolites from glucose and fatty acid metabolism. This opens avenues for further assessment using a fresh set of matched cases and collected controls, removing the discrepancy produced by the use of database controls and permit a closer examination of the preliminary findings from this study.

Earlier in this thesis, various efforts to study the pancreatic cancer metabolome were reviewed. A common feature to all the studies was the variation in the substrates, the methodology used for the examination of the profiles and a considerable variation in the metabolites identified. Metabolites have different physical and chemical properties that are detected using different methodologies and this fact underpins the difficulties in studying metabolite profiles. The lack of standardization of study protocols, the variety of methodologies and separation methods and statistical interpretation of the profiles generated makes the use of metabolomic profiles difficult to apply in clinical practice. This is complicated by the underlying physiological dysfunction that accompanies patients with cancer. We need to be able to differentiate between the physiological derangements secondary to the metabolic response to cancer compared to the downstream effects of the genetic changes in the cancer. A further difficulty in identifying metabolites of biological importance is the lack of a complete database of biologically significant metabolites. This being addressed by the human metabolome database, but is still developing.

Nonetheless, the use of metabolomic technologies expands our understanding of the downstream changes of the genetic processes in pancreas cancer. As we develop and
characterise the metabolites in pancreas cancer, combined with genetic data, a systems biology approach can help identify abnormalities in biological pathways in cancer. Theoretically, this could help us to delineate changes in pathways long before the effects are visible on scans or on clinical presentation. In the era of personalized medicine, the study of an individual’s metabolic fingerprint represents the closest representation of the biochemical processes occurring within.

Future work should focus on several areas. These include: 1) The impact of environmental factors such as smoking, jaundice and cachexia on the metabolite snapshot in individuals. 2) Studies looking at the diurnal variation in metabolite levels. 3) Studies assessing the reproducibility of identified metabolites using different techniques on the same sample type. 4) Based on the genetic abnormalities identified on various biomarker studies in a given individual with pancreas cancer, there is an opportunity to tailor substrate metabolic analysis towards those specific downstream pathways, potentially identifying features of the cancer sooner. This could, in the future, present a personalized expression of an individual’s biology and be used as a baseline to identify changes in the picture, particularly with relation to cancer.
References


232. Wasserman KH, JE; Sue, DY, Stringer, WW; Sietsema, KE; Sun, X; Whipp, BJ. Principles of exercise testing and Interpretation: Including Pathophysiology and Clinical applications. 5th ed: Lippincott Williams & Wilkins; 2011.


Appendix

1. Patient Information leaflet

Invitation to participate in a research study (Study group)

The genomic, metabolomic and protein marker profiling of pancreatic cancer, high risk chronic pancreatitis and pancreatic cystic tumours.

(Pancreas Pathway Project)

We would like to invite you to take part in the above research study. Before you decide you need to understand why the research is being done and what it would involve for you. Please take some time to read the following information carefully. Talk to others about the study if you wish.

This document is in two parts:

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

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

Part 1

What is the purpose of the study?

Pancreas cancer is a relatively rare condition, but it is most often diagnosed when it is at a late stage. Therefore early diagnosis and treatment of pancreatic cancer offers a better chance of survival. With this project we aim to develop new markers of pancreatic cancer that would enable medical professionals to treat this condition when it is at an early stage, thus improving the chances of survival on the longer term. Our aim is to identify the basic structure of cells in pancreatic cancer and high risk conditions like chronic pancreatitis and cystic tumours of the pancreas, this would enable us to develop specific markers for each of the three conditions that are mentioned above.
The combination of genetic, metabolic and protein markers will help in the development of a “molecular signature” for each patient. The use of various methods will maximize the chance of a successful identification of biomarkers to help identify and facilitate early treatment of these diseases.

Why have I been selected?

The study involves patients with a suspected pancreas cancer or a high-risk lesion (pancreatic cystic lesions or high risk chronic pancreatitis) who are being assessed by the hepato-biliary surgical team to be considered for surgery or a non-surgical management.

Do I have to take part?

Taking part in the study is entirely voluntary. We will describe the study and go through this information sheet, which we will then give to you. We will then ask you to sign a consent form to show that you have agreed to take part. You are free to withdraw from the study at any time without giving a reason. This would not influence the treatment or the standard of care you receive.

What would happen to me if I take part?

If you agree to take part in this study, we will collect a blood sample from you when other routine blood tests are performed in the clinic and on the ward. In addition we will also collect a urine sample at your first clinic visit. At the time of surgery we will collect pancreatic juice, bile and tumour tissue. A pathologist will collect the tumour tissue so that it will not effect the reporting of your results. These samples may be stored for a period of up to 5 years.

Additionally we will have to take a sample of blood prior to discharge from hospital if you have had surgery and at your subsequent clinic visits along with radiological investigations at 3 and 6 months intervals. You would not need to make additional visits to the hospital because of the study.
Are there any drawbacks of taking part?

There is a small risk of bruising at the site where blood is taken. Having the small additional amount of blood taken for our study should not increase the discomfort or the risk of having the blood taken. The bile and pancreas juice samples taken at your operation will not effect the operation in any way. The samples are taken after dividing the bile and pancreas ducts as part of the normal steps in the operation. The cancer tissue will be analysed after all other tests are done and will not affect the report provided by the pathologist.

What are the benefits of this study?

There will be no direct benefit to you from this research study. However, once all the data has been analysed it is hoped that the genes or other metabolic biomarkers identified will enable the development of a method to monitor and diagnose pancreatic cancer patients in the future.

Who can I talk to about this study?

Our researchers, Mr Sudip Sanyal and Mr Asela Bandara, would be happy to sit down and answer any questions that you may have. They can be contacted via Prof Siriwardena’s secretary at 0161-2712534.

Will my taking part in the study be kept confidential?

Yes. We will follow ethical and legal practice and all information about you will be handled in confidence.

This completes part 1

If the information in Part 1 interests you and you are considering participation please read the additional information in Part 2 before making any decision.
Part 2

Are there any drugs or additional tests involved?

There are no drugs involved and the only tests are those that are routinely performed in the diagnosis and treatment of all patients undergoing evaluation No index entries found for a Whipple’s procedure.

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

In this instance your blood, bile and pancreas juice sample will be removed from storage and destroyed. All the data collected regarding you will be removed from the study records. This will not have any effect on your treatment.

Will my taking part in this study be kept confidential?

We will abide by the 1998 data protection act rules and regulations when handling all the data that have been collected for the study. The information will be collected from your hospital notes and coded, encrypted and stored in a password protected NHS computer. If you agree the data will be used for future studies for cancer detection tests. Only the principal investigator – Professor Ajith Siriwardena and co-researchers Mr Sudip Sanyal and Mr Asela Bandara will have access to the coded records. These records will be stored for 5 years and after which they will be securely removed. Within this period, your medical records may also be looked at, if required, by the researcher or someone appointed by him for that purpose.

Informing your General Practitioner

We will inform your General Practitioner about your participation of this study if you wish.

What will happen to the samples I give?

The blood, bile and pancreas juice samples will be stored for a period of up to 5 years. They will be analyzed in batches and discarded.
What will happen to the results of the research study?

The results of this study will be presented in scientific conferences and journals. Your personal details will not be identified in any of these presentations.

Who has reviewed the study?

An independent research ethics committee to protect your safety, rights and dignity looks at all research in the NHS. The North West Research Ethics Committee has approved this study.

In the event that something does go wrong and you are harmed during the research and this is due to someone’s negligence then you may have grounds for a legal action for compensation against Central Manchester University Hospitals NHS Foundation Trust but you may have to pay your legal costs. The normal National Health Services complaints mechanism will still be available to you (if appropriate).
Central Manchester University Hospitals
NHS Foundation Trust

Hepato-Pancreateco-Biliary Unit
Department of Surgery
Manchester Royal Infirmary
Oxford Road
Manchester M13 9WL

Study no:
Patient identification no:

CONSENT FORM

The genomic and metabolomic profiling of pancreatic cancer and high risk lesions. (chronic pancreatitis and pancreatic cystic tumours).
“The pancreas pathway project”

Researchers:
Mr S Sanyal/ Mr Asela Bandara
Dept. of Surgery, HPB Unit
Manchester Royal Infirmary
Manchester M13 9WL
TP: 01612764250

Please tick box

1. I confirm that I have read and understand the information sheet for the above study and have had the opportunity to ask questions. ☐

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 being affected. ☐

3. I agree to take part in the above study. ☐

4. I understand that relevant sections of my medical notes and data collected during the study, may be looked at by individuals from Central Manchester University Hospitals NHS Foundation trust, from regulatory authorities, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records. ☐

Name of Patient ______________________ Signature ______________________ Date __________

Name of Person taking consent
(If different from researcher) ______________________ Signature ______________________ Date __________

Researcher ______________________ Signature ______________________ Date __________

I for participant; 1 for researcher site file; 1 (original) to be kept in medical notes

200