University of Glasgow

Ali, Asif (2015) *Biomarkers for pancreatic cancer: identification, validation and clinical application.* PhD thesis.

http://theses.gla.ac.uk/6272/

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten:Theses
http://theses.gla.ac.uk/
theses@gla.ac.uk


# Biomarkers for Pancreatic Cancer: Identification, Validation and Clinical application

**Asif Ali**

**M.B.B.S, PGD EBM Ed**

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy (PhD)

Wolfson-Wohl Cancer Research Centre, Institute of Cancer Sciences, College of Medical, Veterinary & Life Sciences, University of Glasgow, Scotland, UK

March, 2015

# Abstract

Pancreatic cancer is common and aggressive: the main type is pancreatic ductal adenocarcinoma (PDAC). It is the fifth most common cause of cancer death in the UK with an overall five year survival of only 2-3%. Establishing the diagnosis of PDAC is important for optimal patient management but can be difficult and relies on imaging and cytology/pathology. Although imaging may be highly suggestive of PDAC, a pathological diagnosis is preferred prior to definitive treatment; therefore tissue samples are required. Cytology samples are obtained at endoscopy. Cytological analysis requires the identification of different cell types and in particular the distinction of malignant pancreatic epithelial cells from reactive pancreatic cells and other gastrointestinal contaminants. This requires experience and expertise and can be difficult. A tissue diagnosis is not achieved in a significant proportion of PDAC cases. Hence, an unmet clinical need exists for the diagnosis of PDAC from cytological samples. One potential way of improving the cytological diagnosis is to use immunohistochemistry (IHC) biomarkers as an adjunct to cytology in difficult to diagnose cases. Diagnostic IHC biomarkers have been investigated, but to date none has entered into routine clinical practice.

The aim of this study was to improve the diagnosis of PDAC from cytology samples. It is hoped that the identification and validation of IHC biomarkers in PDAC will help their clinical translation. For biomarker identification a meta-analysis of potential IHC diagnostic biomarkers investigated in PDAC was performed. Sixteen biomarkers were quantified in meta-analysis and the highest ranked biomarkers were KOC, maspin, S100P, mesothelin and MUC1. These biomarkers have not entered into clinical practice partly because they were investigated in separate studies with relatively small sample sizes and without a uniform and clinically appropriate cut-off.

Biomarkers identified in the meta-analysis were validated in a resection cohort from patients with pancreatico-biliary adenocarcinoma. The aim was to identify better biomarkers and cut-offs that could potentially be investigated in cytology samples. KOC, S100P, mesothelin and MUC1 were investigated in one set of tissue microarrays, while maspin was investigated in another set. Five cut-offs were carefully chosen for sensitivity/specificity analysis using receiver operating

characteristics curve analysis. Using 20% positive cells as a cut-off achieved higher sensitivity/specificity values: KOC 84%/100%; S100P 83%/100%; mesothelin 88%/92%; MUC1 89%/63%; and Maspin 96%/99%. Analysis of a panel of KOC, S100P and mesothelin achieved almost 100% sensitivity and specificity if at least two biomarkers were positive for both 10% and 20% cut-offs.

Clinical translation of biomarker requires a reliable and reproducible cut-off for interpretation of IHC. We identified three cut-offs for investigation to establish a consensus based cut-off(s) that could potentially be used by pathologists for PDAC and other cancers. A series of IHC images of microarray cores were used to investigate observer agreements for 10%, 20% and +2/+3 cut-offs. Seven pathologists participated in this study. The inter- and intra-observer agreements using the three cut-offs were reasonably good. For all three cut-offs a positive correlation was observed with perceived ease of interpretations ($p<0.01$ for all cut-offs). Finally, cytoplasmic only staining achieved higher agreement than cytoplasmic/nuclear and cytoplasmic/membranous staining. All three cut-offs investigated achieve reasonable strength of agreement modestly decreasing inter and intra-observer variability in IHC interpretation but 10% is slightly better than 20% and +2/+3 cut-offs.

Finally, KOC, maspin, mesothelin and S100P were investigated in archival cytology samples with the aim to generate a diagnostic panel which could potentially help as an adjunct to cytology. Using 10% cut-off achieved higher sensitivity/specificity values: KOC 92%/100%; maspin 54%/100% and mesothelin 72%/100%. But no staining was observed for S100P. In addition, analysis of a panel of KOC, maspin and mesothelin achieved 82% sensitivity and 100% specificity for 10% cut-off if at least two biomarkers in the panel were positive. The inter-observer agreement for 10% positive cells as IHC cut-off in cytology samples was very good for all three biomarkers.

In conclusion, a panel of KOC, maspin and mesothelin is a suitable diagnostic panel and 10% cut-off is a reasonable cut-off achieving high observer agreement. Their diagnostic accuracies approach those of optimal conventional cytology. These markers may be appropriate for further clinical validation and potentially routine use in difficult cases.

# Table of Contents

# List of Tables

# List of Figures

# Dedication

To my beloved mother Aseah Bibi whose prayers, encouragement and never ending love enabled me to achieve every dream of my life. To my father Abdul Wahid Khan who died when I was very young. To my cousin Salahuddin Khan, I owe you a huge amount that I cannot pay.

# Acknowledgement

I consider myself very lucky to have Dr Karin Oien as my supervisor. Karin is a brilliant academic, clinician and scientist and has great experience with extensive publications in cancer and pathology field. I believe that ALLAH has gifted Karin with extraordinary skills and a very caring and loving personality. She is an excellent mentor and has guided me through all the phases of PhD: in the first year to get my head round the relation of science to clinic; in the second year to work on my project independently under her kind guidance and encouragement; and in 3rd year to become an independent researcher and an academic writer. I must say that she has not only provided me the vision and direction in my PhD but has enormously helped me in developing my personality.

I am very thankful to my co-supervisor, Professor Owen Sansom for guidance and encouragement throughout my PhD especially during the time when Karin was on maternity leave.

I am also very thankful to Fraser Duthie, Jennifer Morton and Nigel Jamieson, who although not 'officially' recognised but helped and supported me like a co-supervisor. I am greatly indebted to Fraser who helped me with cytology project in particular but also provided me invaluable comments and feedback on various manuscripts and this thesis. Jennifer was always available to me without any appointment and willing to help with every aspect of my PhD from planning to practical experiments and to writing up my results. Throughout my PhD Nigel provided me constructive criticism on my work and we were always in touch via email communications. I am also grateful to the excellent support of Dr Daniel Mackay and Dr Zia Ul-Haq for making the meta-analysis possible.

I would also like to acknowledge the help and assistance of Fiona Conway, Joanne Thomson, Irene Todd and Margaret Jenkins (administration), my friends and colleagues Jennifer Rocissana, John Revie, Alan Bilsland, Sharon Burns, Clare Orange and Claire Cairney (PhD students and research scientist), Prof. Nicol Keith, Prof. Alan Foulis, Dr David Chang, all the faculty members of the Institute of Cancer Sciences and West of Scotland Pancreatic Cancer Unit, and many others for their support during my stay in Glasgow. Jennifer Rocissana in particular was always there when I was feeling tense or depressed and

counselled me with smiling face during those tense/depressed phases. Thank you very much to Jennifer Rocissana, Alan Bilsland and Sarah Bell for proof reading of my thesis. Sarah also enormously helped me with cytology project.

There are a number of colleagues within the department and outside whom I would like to acknowledge their input and excellent technical assistance: Rod Ferrier and Colin Nixon for the excellent assistance with immunohistochemistry (IHC). Coming from a clinical background, Rod was my first teacher and mentor in learning lab skills. Colin enormously helped me to learn IHC and different IHC protocols for optimisation; Fiona McGregor, Sarah Ghafoor for their assistance with IHC optimisation; Clare Orange for her extraordinary support with learning tissue microarray technology and digital pathology; John McCorriston and Lisa Irvine for their support with generating the needling cytology resource; and finally I would like to pay gratitude to the biobank staff at the Southern General Hospital especially Phil Duffy, Alexis Stewart, Alex Bell, Fiona Graham and Jane Hair.

Apart from my research, I enjoyed playing cricket for University of Glasgow for 3 years and remained unbeaten in this season (2013-2014). Thanks to all Cricket Mates (especially Frank, Imtiaz, Rafiq, Inayat and Gaurav) for providing me company and travelling throughout beautiful Scotland. I will miss Scotland for a lot of reasons and Cricket is one of them.

I am grateful to Khyber Medical University, Peshawar, Pakistan and Higher Education Commission of Pakistan for providing me with funding for my higher studies. I am particularly indebted to Prof. HafizUllah, Prof. Shad Muhammad, Prof. Jawad Ahmad, Col. Ihtesham, Mr. Arshad, Prof. Jawad and all KMU PhD Scholars.

I am also thankful to The Charles Wallace Pakistan Trust for their financial support during the last period of my project. Closer to my heart, I thank my friends in Pakistan, UK and other parts of the world for providing me with great company and always tolerating me.

The encouragement, prayers and support of my sister Farah, my brothers Khalid, Waqar, Sarfaraz, my cousins Sikander, Zulfiqar, Nasir and Johar throughout my

whole career have been invaluable. I would like to apologize to my wife, Sana and my two lovely daughters Anaya and Javeria for not giving them the proper attention and love which they deserved during my PhD. This really would not have been possible without their patience.

Lastly, I am extremely thankful to my Creator and Sustainer, ALMIGHTY ALLAH, who is all alone in HIS powers (Surely we belong to ALLAH & to HIM shall we return) for HIS countless bounties on me and for sending HIS Prophets from ADAM (PBUH) to the last Prophet, MUHAMMAD (PBUH) who showed us how to live a successful life.

# Author's Declaration

This thesis is submitted in fulfilment of the requirement for the degree of Doctor of Philosophy at the University of Glasgow. Unless stated otherwise, the work is that of the author. Parts of the research work included in this thesis has been published or submitted with co-authors.

# Definitions/Abbreviations (arranged alphabetically)

| | |
|---|---|
| **AJCC** | American Joint Committee on Cancer |
| **CI** | confidence interval |
| **CP** | chronic pancreatitis |
| **CT** | computerized tomography |
| **DAB** | diaminobenzidine |
| **$dH_20$** | de-ionised and distilled water |
| **DNA** | deoxyribonucleic acid |
| **ERCP** | endoscopic retrograde cholangiopancreaticography |
| **ESPAC** | European Study of Pancreatic Cancer |
| **EUS** | endoscopic ultrasound |
| **FEM** | fixed effect model |
| **FFPE** | Formalin fixed paraffin embedded |
| **FNA** | fine needle aspiration |
| **FOLFIRINOX** | fluorouracil, leucovorin, irinotecan and oxaliplatin |
| **FP** | false positive |
| **H&E** | haematoxylin and eosin |
| **HIER** | heat induced epitope retrieval |
| **IHC** | immunohistochemistry |
| **IMP3** | Insulin like growth factor messenger RNA binding protein 3 |
| **IPMN** | Intra-ductal papillary mucinous neoplasia |
| **KOC** | K homology domain containing protein overexpressed in cancer |
| **MDT** | multi-disciplinary team |
| **Maspin** | mammary serine protease inhibitor |
| **MRI** | magnetic resonance imaging |
| **mRNA** | messenger ribonucleic acid |
| **MUC1** | mucin 1 |
| **NP** | normal pancreatic tissue |
| **PEM** | polymorphic epithelial mucin |
| **PanIN** | pancreatic intraepithelial neoplasia |
| **PDAC** | pancreatic ductal adenocarcinoma |
| **PBA** | pancreatico-biliary adenocarcinoma |
| **PCR** | polymerase chain reaction |
| **QUADAS** | Quality Assessment of Diagnostic Accuracy Studies |
| **REM** | random effects model |
| **RevMan** | Review Manager |
| **RNA** | ribonucleic acid |
| **RT-PCR** | reverse transcription polymerase chain reaction |
| **S100P** | S100 calcium binding protein P |
| **TbT** | Tris buffered Tween |
| **TMA** | tissue microarray |

| | |
|---|---|
| **TNM** | Tumour Node Metastasis |
| **TN** | true negative |
| **TP** | true positive |
| **US** | ultrasound |
| **5-FU** | 5-fluorouracil |
| **5% cut-off** | 5% positive cells and any staining intensity as IHC cut-off |
| **10% cut-off** | 10% positive cells and any staining intensity as IHC cut-off |
| **20% cut-off** | 20% positive cells and any staining intensity as IHC cut-off |
| **+2/+3 cut-off** | Moderate-strong staining intensity and any proportion of positive cells as IHC cut-off |
| **HS20 cut-off** | Histoscore 20 as IHC cut-off |

| **Units** | |
|---|---|
| **°C** | degree Celsius |
| **L** | litre |
| **m** | milli |
| **m** | meter |
| **µ** | micro |
| **rpm** | revolutions per minute |

# Contributors

| | |
|---|---|
| **AA** | Asif Ali |
| **KAO** | Karin A Oien |
| **NBJ** | Nigel Balfour Jamieson |
| **OS** | Owen Sansom |
| **FD** | Fraser Duthie |
| **AB** | Alan Bilsland |
| **JM** | Jennifer Morton |
| **SB** | Sarah Bell |
| **CN** | Colin Nixon |
| **RF** | Roderick Ferrier |
| **VB** | Victoria Brown |
| **SD** | Simon Denley |
| **CO** | Clare Orange |

# 1 Introduction

The term diagnosis comes from the Greek *gnosis*, which means knowledge. Diagnosis is the process of defining the nature of a disease or disorder and then differentiating it from other likely conditions. The diagnostic process involves a combination of clinical features, imaging and pathology. As the pancreas is located deep within the abdominal cavity, multiple imaging modalities are used to make a correct diagnosis of tumours arising in the pancreas. To differentiate benign from malignant disease and to provide a tissue diagnosis, cell samples are obtained from suspected pancreatic lesions during endoscopy. Confirmation of the diagnosis is thus achieved from interpretation of these cell samples and is important for patient management. However, there are instances when the morphologic interpretation of cell samples is challenging.

Molecular diagnostics have emerged as a very important field in cancer diagnosis. Genomics and proteomics have enabled the identification of differentially expressed genes and proteins respectively in pancreatic cancer that could be used as diagnostic biomarkers. These biomarkers can be characterised at the protein level by immunohistochemistry (IHC). IHC is a routine technique in pathology laboratory. However, the amount of IHC biomarker research is not reflected in improving the clinical management of patients with pancreatic cancer.

This introduction will first discuss the normal pancreas and then pancreatic ductal adenocarcinoma (PDAC) (the most common pancreatic cancer) including the diagnostic process and management of PDAC. The role of cytology in PDAC diagnostic management will then be outlined, including factors influencing cytology diagnosis and the unmet clinical need to improve cytology diagnosis. Finally, the potential role of IHC biomarkers to help diagnosis will be discussed with special focus on biomarker development and factors delaying the clinical translation of biomarkers. These factors revolve around the identification and validation of biomarkers and cut-offs for diagnosis. The aims, objectives and implications are outlined at the end of the introduction.

The candidate biomarkers were identified by meta-analysis, and then validated on tissue arrays and cytology samples with investigation of appropriate cut-offs. A suitable diagnostic IHC biomarker panel and cut-offs were then suggested.

## 1.1 Pancreas anatomy and physiology

The human pancreas is pear shaped and weighs approximately 80-100 grams. It lies deep in the upper abdomen with the head of pancreas lying adjacent to the duodenum and the body and tail extending behind the stomach, crossing the midline and reaching as far as the spleen (Figure 1.1) (1, 2).

The pancreas is a compound gland containing both exocrine and endocrine tissue. The exocrine part is composed of two functional compartments: the acini (Latin, “grape” like) and ducts, and the endocrine part which is composed of islets of Langerhans (Figure 1.2 A&B).

The major part of the gland is arranged as pancreatic exocrine acini which produce digestive enzymes that help in the digestion of proteins, carbohydrates and fats. Proteins are digested with the help of proteolytic enzymes trypsin, chymotrypsin, elastase, and carboxypolypeptidase that cleaves proteins into peptides. Carbohydrates are digested by amylase that cleaves them into di- and tri-saccharides and fat digestion is facilitated by lipase, phospholipase and cholesterol esterase. Furthermore, sodium bicarbonate and water are secreted by ductal epithelial cells to neutralize gastric acid (3).

Acini secrete digestive enzymes into the ductal system which form the main pancreatic duct (duct of Wirsung). It is joined by the common bile duct to form the ampulla of Vater that opens into the second part of duodenum through the duodenal papilla guarded by the sphincter of Oddi.

The endocrine part is scattered throughout the pancreas, constituting approximately 2% of total pancreatic mass and secretes a variety of hormones including insulin and amylin by ß-cells, glucagon by α-cells, pancreatic polypeptide by γ-cells and somatostatin by δ-cells. Insulin and glucagon are major hormones that are responsible for glucose homeostasis (1-3)

Histology of the pancreas is very important as it provides the knowledge of the normal structural features of the pancreas and studies that focus on pancreatic cancer are dependent on this knowledge. Pancreatic ductal adenocarcinoma (PDAC) is a malignant tumour that arises in the ductal epithelium. As I will be discussing the expression of biomarkers in normal pancreas and PDAC knowledge of pancreatic histology will help us in better understanding the interpretation and utility of tissue based biomarkers.

## 1.2 Pancreatic Ductal Adenocarcinoma (PDAC)

The term "pancreatic cancer" is conventionally used as an alternative name for pancreatic ductal adenocarcinoma (PDAC) as it accounts for more than 85% of pancreatic neoplasms (4). PDAC arises in ductal epithelium and in 70% cases it arises in the head of pancreas with an aggressive clinical course.

### 1.2.1 Epidemiology

In 2010, pancreatic cancer was the 10$^{th}$ most common cancer in the UK and ranked 12$^{th}$ in men and 8$^{th}$ in women. Around 8,500 new cases were diagnosed in 2010 in the UK with a male to female ratio of almost 1:1. The crude incidence rate for pancreatic cancer is 14 new cases per 100,000 people. In 2010, an average of around 75% of cases were diagnosed in patients aged 65 years and over in the UK (5).

The incidence rates of pancreatic cancer have remained stable in Great Britain since the 1970s. Moreover, the survival rates for pancreatic cancer have slightly increased since the 1970s. The 5-year survival rate for people diagnosed with pancreatic cancer is around 4% but younger patients have a relatively better survival than older patients. It is the 5$^{th}$ most common cause of cancer related deaths in UK and in 2011, 8300 patients died from pancreatic cancer (5, 6). This poor prognosis is partly due to late clinical presentation with advanced disease, when the treatment options are limited and relatively ineffective (7).

### 1.2.2 Etiology

Smoking is by far the most common risk factor causing an estimated 25% of cases of pancreatic cancer in the UK. Other non-hereditary risk factors are old age,

obesity, chronic pancreatitis and eating processed meat (8, 9). Recent onset of diabetes may be an early sign of cancer development and patients with type I or II diabetes have approximately twice the risk of developing pancreatic cancer. Family history is also an important risk factor and people with one first-degree relative with pancreatic cancer have almost twice the risk of developing the tumour as compared to those without any family history (10, 11).

### 1.2.3 Tumorigenesis

At present three main PDAC precursor lesions have been identified based on clinical and histopathological research: pancreatic intraepithelial neoplasia (PanIN); intraductal papillary mucinous neoplasm (IPMN) and mucinous cystic neoplasm (MCN) (12, 13). The most common precursor lesion is pancreatic intraepithelial neoplasia (PanIN), which are a sequence of microscopic lesions generally arising in small pancreatic ducts (less than 5mm). PanIN-1 is a low grade lesion sub-classified as 1A (flat) and 1B (papillary) and characterized by nuclear atypia. PanIN-2 is an intermediate grade lesion with slightly more complex architecture compared to PanIN-1 and characterized by nuclear changes including nuclear crowding and pseudostratification, loss of polarity, pleomorphism and hyperchromasia. Mitoses in PanIN-2 are rare. PanIN-3 is a high grade lesion with widespread nuclear atypia, loss of nuclear polarity, frequent occurrence of mitoses and budding off of epithelial cells into the lumen. PanIN-3 is a pre-invasive lesion and is also known as “carcinoma in situ” as it has not breached the basement membrane (14, 15). The sequence of morphologic and genetic changes suggests that PanIn might be a precursor lesion for infiltrating ductal adenocarcinoma (Figure 1.3) (12, 16).

IPMN and MCN are macroscopic pre-invasive lesions. IPMNs are usually present in the head of pancreas and further classified as pancreatico-biliary, intestinal and gastric subtypes. MCN are usually solitary lesions found mainly in the body and tail of pancreas (17).

### 1.2.4 Pathology of PDAC

PDAC is a malignant neoplasm arising in the ductal epithelium of pancreas, showing glandular differentiation. It is a rapidly growing tumour and irrespective

of the size of the tumour, metastasizes to liver and lymph nodes. Macroscopically, they are typically solid tumours presenting as poorly circumscribed masses with an ill-defined and infiltrating margin.

Microscopically, the neoplastic epithelial cells are highly infiltrative demonstrating various patterns of glandular arrangement. The tumour glands are irregularly distributed without following the lobular architecture and spatial organisation of normal ducts. In addition, there is often a prominent desmoplastic stroma surrounding the tumour (Figure 1.4A). Morphologically, two types of PDACs have been recognised: pancreatico-biliary and intestinal type with the former associated with a worse prognosis than the latter (18, 19).

The diagnosis of PDAC can be challenging, both at the clinical and microscopic levels. The scirrhous (scar like) appearance of this tumour results from abundant host tissue stroma (desmoplastic stroma) (Figure 1.4A). This makes the differential diagnosis of PDAC from scarring inflammatory lesions like autoimmune pancreatitis and chronic pancreatitis (CP) difficult. Cytologic atypia induced by inflammation or degeneration of native ducts may mimic carcinoma. Conversely, well differentiated PDAC may resemble benign ducts (4, 20) (Figure 1.4B). Thus the discrimination of PDAC from CP is considered as a challenging differential diagnosis in diagnostic pathology (18). Despite their highly aggressive nature, most PDACs are well or moderately differentiated (Figure 1.4B&C) as compared to poorly differentiated tumours (Figure 1.4D). Both perineural and vascular invasion are common during microscopic examination.

About 60-70% of tumours develop in the head of pancreas; the rest arise in the body and tail of pancreas. PDAC arising in the head of pancreas presents earlier with jaundice and/or acute pancreatitis as compared to tumours arising in the body and tail of pancreas. Therefore, due to late clinical presentation, tumours arising in the body and tail are associated with poorer prognosis compared to pancreatic head tumours (7, 18, 19, 21).

The poor prognosis of PDAC is due to late clinical presentation with so-called “red flag” symptoms appearing once the tumour has extended beyond the confines of the pancreas. Local invasion is common and often involves duodenal

wall, common bile duct, ampulla of Vater, spleen, stomach, adrenal gland, peripancreatic tissue and adjacent vessels (7, 19, 22).

Knowledge of the pathology of the pancreas is important for understanding the current project. I will be validating the expression of IHC biomarkers in PDAC and this knowledge will help in observing and localising the differential expression of biomarkers in PDAC compared to normal tissue for potential diagnostic purposes.

## 1.3 Diagnosis of PDAC

Diagnosis of PDAC requires a combination of clinical findings, imaging and cytology/pathology. Here I will give an overview of the diagnostic process for PDAC but I will give a more detailed account of cytology in section 1.7, which will also provide the rationale for the current thesis.

### 1.3.1 Clinical presentation

Clinical presentation depends on the location of tumour and disease stage. Obstructive jaundice and weight loss are the commonest presenting features of tumours in the head of the pancreas. Abdominal pain and weight loss are the commonest presenting features of tumours arising in the body and tail. Pain is an important clinical symptom reported as a deep, dull and aching pain originating from the upper abdomen and radiating to the back.

Pancreatic cancer can cause other signs and symptoms such as asthenia, anorexia, hyperglycemia, abnormal liver function, panniculitis and superficial venous thrombosis. Locally advanced tumours can cause gastric outlet obstruction, duodenal obstruction and rarely gastrointestinal bleeding (7, 22, 23).

### 1.3.2 Imaging

Imaging contributes to the diagnosis of pancreatic cancer and is important for pre-treatment staging based on local invasion and metastases.

#### 1.3.2.1 Non-invasive imaging

Ultrasound, computed tomography scan, and magnetic resonance imaging are non-invasive imaging modalities for PDAC diagnosis with variable sensitivity and specificity.

Abdominal ultrasound (US) is safe, inexpensive and helps to formulate a differential diagnosis for possible causes of obstructive jaundice (22, 24). Computed tomography (CT) scan is a well validated imaging tool used for the diagnosis and staging of patients with suspected pancreatic cancer. CT scan allows visualisation and assessment of vascular and surrounding organs involvement by tumour (24, 25). Magnetic resonance imaging (MRI) can also be used for the diagnosis and staging of pancreatic cancer (24).

#### 1.3.2.2 Invasive imaging

Endoscopic retrograde cholangio-pancreato-graphy (ERCP) and endoscopic ultrasound (EUS) are minimally invasive imaging modalities to identify small mass lesions and abnormalities of pancreatico-biliary ductal systems not identified by non-invasive imaging (26-29). EUS and ERCP are also used to obtain cytology samples to confirm tissue diagnosis.

### 1.3.3 Cytology and histopathology

Imaging is suggestive of PDAC, but in the UK confirmatory tissue diagnosis is preferred for treatment planning purposes and is usually achieved from the interpretation of cytology samples obtained during ERCP and/or EUS. A detailed account of cytology is given in section 1.7/1.8.

Histopathologic reporting of the tissue specimens is achieved only in patients with resectable disease (15-20% of PDAC cases). The distinction of PDAC from reactive pancreatic changes (especially chronic pancreatitis) and other primary pancreatic or metastatic tumour is important as the treatment, prognosis and follow-up significantly differs. Meticulous microscopic examination of the pathology specimens is the backbone of this work up.

## 1.4 Staging system for pancreatic cancer

The staging system for pancreatic cancer shown in Table 1.1 is according to the most recent edition (7th Edition) of the American Joint Committee on Cancer (AJCC) (30).

The staging system especially T stage is very important for predicting the surgical resectability of primary tumour. Stages T1, T2 and T3 tumours are potentially resectable and stage T4 corresponds to unresectable disease (7). Staging is an integral part of the diagnostic work up for the management of patients with PDAC and pathology reports on resection specimens clearly states the pathologic findings of staging.

## 1.5 Management of pancreatic cancer

The treatment for pancreatic cancer depends on the stage of the cancer, the location of tumour and fitness of the patient and is best managed by a multidisciplinary team (MDT). The MDT is usually composed of surgeons, oncologists, radiologists, pathologists and other staff including nursing and palliative care staff and dieticians etc. Surgical resection with adjuvant chemotherapy is indicated for patients with Stage I/II disease. Neoadjuvant therapy is indicated in patients with Stage III borderline resectable cancers and if the tumour is successfully downstaged to stage II disease the patient will be offered surgical resection (31, 32). Stage III locally advanced disease should be treated with chemotherapy or chemoradiotherapy. Patients with good health and stage IV disease receive systemic therapy, whereas patients with poor performance status may be given the best supportive therapy (33).

### 1.5.1 Surgical resection

The most widely used operation for the treatment of PDAC involving the head of the pancreas is a Whipple pancreatoduodenectomy, which involves removal of the distal stomach. A variant of the Whipple procedure preserves pylorus and is known as pylorus preserving proximal pancreatoduodenectomy (PPPP) (34, 35). Unfortunately, tumours of the head of pancreas are unresectable in the majority of the cases and only about 20% of patients are suitable for surgery (7, 22, 36,

37). In high volume tertiary centres the operative mortality is less than 6% but the morbidity rate is 40% (38).

Tumours located in the neck and body of pancreas may require pancreatoduodenectomy, distal pancreatectomy or, rarely, a total pancreatectomy (39).

#### 1.5.1.1 Post-operative outcome

The median 5 year survival rate after resection is 10-20% and the median survival is 10-19 months (40, 41). Tumour size more than 3cm, positive resection margins, high histologic grade and lymph node invasion negatively affect survival, however, surgical resection still remains the only potentially curable option (40).

Generally, quality of life after PPPP is good apart from malabsorption which is a common complication and can require enzyme supplementation.

### 1.5.2 Adjuvant therapy

Adjuvant therapy may be given following curative resection of pancreatic cancer. The European Study Group for Pancreatic Cancer trial-1 (ESPAC-1) trial showed that adjuvant chemotherapy with 5- fluorouracil (5-FU) significantly improved the survival of patients with PDAC. The median survival was 20.1 months in patients who received chemotherapy compared to 15.5 months in patients who did not receive chemotherapy (42).

Another clinical trial, ESPAC-3 compared 5-FU with gemcitabine and found a median survival of 23 months and 23.6 months for patients treated respectively with 5-FU/folinic acid and gemcitabine (43). Gemcitabine has fewer side effects and is generally preferred over 5-FU. A further clinical trial, ESPAC-4 is in progress with the aim of comparing the survival after surgery for patients with PDAC treated with a combination of gemcitabine plus capecitabine with gemcitabine alone.

In patients with metastatic pancreatic cancer two regimens have recently emerged. In one regimen, FOLFIRINOX (5-fluorouracil, oxaliplatin, irinotecan,

leucovorin) as compared to gemcitabine showed survival advantage in patients with metastatic pancreatic cancer. The median overall survival was 11.1 months in the FOLFIRINOX group compared to 6.8 months in the gemcitabine group (44). In the other regimen nab-paclitaxel plus gemcitabine significantly improved survival as compared to gemcitabine alone (45). The median overall survival was 8.5 months in the nab-paclitaxel-gemcitabine group compared to 6.7 months in the gemcitabine group. There is a need to identify patients who will benefit from one of these treatment regimens.

### 1.5.3 Neoadjuvant therapy

The purpose of neoadjuvant therapy is to downstage locally advanced disease for potential resection, achieve R0 resections (clear resection margin) and reduce the recurrence of tumour (22, 46, 47). There is evidence that neoadjuvant therapy can enable resectability in up to 30%-40% of patients with locally advanced pancreatic cancer patients (47). Postoperative survival of these patients is comparable to patients with resectable disease. Patient selection is therefore an important aspect of neo-adjuvant therapy. Katz et al (48) reported that 78% (n=125) of patients (N=160) with borderline resectable pancreatic cancer received neoadjuvant therapy. These patients were then re-staged after neoadjuvant therapy and 41% (n=66) of these patients underwent pancreatectomy. The optimum regimen is not yet known, but combination chemotherapy and chemoradiotherapy have been used as neoadjuvant regimens.

### 1.5.4 Palliative management

Surgical, endoscopic, or radiological techniques are employed for palliative treatment of duodenal and biliary tract obstruction. A stent is put in the biliary duct and duodenum to relieve symptoms of obstructive jaundice and delayed gastric emptying. A palliative gastrojejunostomy is performed in patients found unresectable during laparotomy for resectable disease, and deemed to be at risk of developing gastric outlet obstruction (22, 49, 50).

Pancreatic insufficiency is managed by pancreatic enzyme supplements. Pancreatic fistulas can occur with an incidence of 0-25% and can either be treated conservatively or a re-operation may be needed (51).

## 1.6 Diagnostic terms

The terms ‘sensitivity’ and ‘specificity’ will be used throughout this thesis from the introduction, through to results and in the discussion. Therefore, it is necessary to define and discuss the importance of these terms for research diagnostics before going any further.

Sensitivity and specificity are measures of the diagnostic accuracy of a test used to diagnose a disease. We need a test that can categorise the population at risk into one of the two groups: people likely to have the disease; and people unlikely to have the disease.

### 1.6.1 Sensitivity

Sensitivity is “the ability of a test to correctly identify those patients with the disease” or more formally “the probability of a positive test result if disease is present” and is calculated by the formula:

$$\text{Sensitivity} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

OR

$$\text{Sensitivity} = \frac{\text{TP (Patients with the disease correctly identified by the positive test)}}{\text{TP+FN (Total number of patients with the disease)}}$$

True positive (TP): Patient with the disease is positive for test.

False Negative (FN): Patient with the disease is negative for test.

### 1.6.2 Specificity

Specificity is “the ability of a test to correctly identify those patients without the disease” or more formally “the probability of a negative test result if disease is not present” and is calculated by the formula:

$$\text{Specificity} = \frac{\text{True Negative}}{\text{True Negative} + \text{False Positive}}$$

OR

$$\text{Specificity} = \frac{\text{TN (Patients without the disease correctly identified by the negative test)}}{\text{TN+FP (Total number of patients without the disease)}}$$

True Negative (TN): Patient without the disease is negative for test.

False positive (FP): Patient without the disease is positive for test.

The specificity of a diagnostic test is considered relatively more important than sensitivity as low diagnostic specificity may categorise normal as diseased. Diagnostic tests are based on a continuous measurement or ordinal category but a cut-off is chosen for calculating sensitivity and specificity. Cut-points help in identifying an appropriate combination of sensitivity and specificity but generally as one increases, the other decreases (52-54). Table 1.2 can be used for calculating sensitivity and specificity.

## 1.7 Role of cytology in PDAC diagnosis

The role of cytology in the diagnosis and management of patients with PDAC is well established. Major developments in the sampling techniques have significantly improved the diagnostic management of these patients (55).

### 1.7.1 Indications for cytological sampling

Cytological sampling is most commonly indicated to confirm malignancy in patients with suspected pancreatic mass or stricture lesions on imaging. Confirmatory tissue diagnosis is necessary before chemotherapy or radiotherapy treatment, however a cytology specimen is not always required for resection when the suspicion of cancer is high; as generally, the resection will provide therapeutic benefit, and substantially delaying surgery to confirm a diagnosis may deny potentially curative treatment. Nonetheless, pre-operative tissue diagnosis provides more confidence to both the surgeon and patient and it allows optimal planning for surgery (28) (56-62).

## 1.7.2 Techniques of obtaining cytologic sampling

Endoscopic ultrasound fine needle aspiration (EUS-FNA) is usually used for sampling pancreatic mass lesions. ERCP is the conventional method of obtaining cytology samples, via brushing, of suspected pancreatic duct and bile duct strictures (63).

### 1.7.2.1 Endoscopic ultrasound fine needle aspiration (EUS-FNA)

Endoscopic ultrasound (EUS), like upper gastrointestinal endoscopy, is an outpatient procedure in which suspected lesions in the pancreas are imaged from stomach and duodenum. It is considered an accurate diagnostic imaging modality for pancreatic mass lesions especially for lesions less than 2cm (26, 64-66). However, EUS has low specificity in differentiating malignant from inflammatory pancreatic masses (29, 66). The addition of fine needle aspiration (FNA) provides samples for cytopathologic analysis. Indeed, this is a major advantage of EUS over other imaging modalities.

Endoscopic ultrasound-guided fine-needle aspiration (EUS-FNA) is used to obtain cytological samples from suspected pancreatic lesions, in particular mass lesions. It is a well-tolerated and minimally invasive procedure but the results are operator-dependent (67, 68). It has the ability to sample a wide area of a pancreatic mass lesion and allows multiple needle passes to be performed with low complication rate.

Generally, EUS-FNA is a safe procedure. Complication rates of EUS-FNA are 0-2% and include bleeding, pancreatitis, infections and bile peritonitis (69-72). The risk of tumour seeding with EUS-FNA is very low as compared to FNAs obtained by the percutaneous approach (73).

The sensitivity of EUS-FNA for diagnosing pancreatic mass lesions ranges from 73% to 100% and the specificity ranges from 75% to 100% (57, 60, 68, 70, 74-80). The specificity of EUS-FNA is high (100%) in most studies but the reported sensitivity in most studies is 80%-90%, therefore negative results do not absolutely exclude malignancy (59, 70, 81).

These statistical figures thus show that the sensitivity of cytological diagnosis is not perfect and the diagnosis from inconclusive EUS-FNA cytology cases could be improved, perhaps through additional biomarkers. Investigation of markers may also open the door for personalised treatment in modern oncology.

#### 1.7.2.2 Endoscopic retrograde cholangiopancreatography (ERCP) with brushing

Malignancies arising in the pancreatico-biliary (PB) ductal system may not be visible on radiographic imaging and thus may need endoscopic imaging (ERCP) (27, 63). Several benign PB conditions like chronic pancreatitis, primary sclerosing cholangitis and postoperative strictures can also present as stricture lesions. The distinction of malignant and benign lesions in the PB tree represents a diagnostic challenge (27).

ERCP with brushing is a minimally invasive technique for sampling stricture lesions in the PB ductal systems (82-84). The cytological material obtained helps to confirm the diagnosis of either benign or malignant disease (55).

The overall sensitivity of ERCP brushing cytology is less than EUS-FNA cytology with a range of 44-72%; however, the specificity approaches 100% with a range of 90%-100% (27, 60, 63, 85-89).

The sensitivity of brush cytology is low and could be improved through the addition of ancillary tests to the diagnostic work-up of patients with an indeterminate cytology diagnosis.

### 1.7.3 Evaluation of cytology samples

Cytological samples thus obtained are interpreted and reported by a cytologist. The cytology samples contain a mixed population of normal and abnormal cells. Initial assessment of cells involves the identification of normal pancreatic ductal and acinar epithelial cells, and recognition of contaminations from duodenal or gastric epithelium (Figure 1.5). The abnormal cells from neoplastic and non-neoplastic disease processes are assessed to make the cytology diagnosis.

The cytological features of PDAC include: irregular arrangement of cells in sheets, cluster, and three-dimensional cells and as single cells; overlapping pleomorphic nuclei, single and tall cells, high nuclear/cytoplasmic ratio, irregular nuclear membranes, macronucleoli, coarse and clumped chromatin, mitotic figures and necrosis (Figure 1.6) (90-92).

#### 1.7.3.1 Diagnostic categories of cytology

Cytopathologists often use the following broad categories for cytological reporting: non-diagnostic, negative for malignancy, atypical, suspicious for malignancy and positive for malignancy (60).

'Non-diagnostic' implies an overall low cellularity or a sample not appropriate for cytological evaluation due to blood or inflammatory cells. 'Negative for malignancy' means that the sample obtained contain normal or benign cells but this does not preclude malignancy and may possibly be a sampling error during the endoscopic procedure.

'Atypical' cases are reported when there is a low suspicion of malignancy based on mild cytologic and/or architectural atypia. This often occurs in the setting of the inflammatory process, reactive or degenerative changes due to instrumentation or stent placement and a hypocellular sample. The 'suspicious for malignancy' cases are reported when there is a high probability of malignancy but the degree of cytologic atypia is not exclusively confirmatory of malignancy and a benign etiology cannot be excluded with complete confidence (60). The 'positive for malignancy' cases are reported when the cytologic and architectural atypia suggests malignancy, in this case adenocarcinoma.

### 1.7.4 Factors affecting PDAC diagnosis from cytology samples

As shown in the previous sections, neither EUS-FNA nor ERCP with brushing is 100% sensitive in identifying the malignancy or 100% specific in excluding benign diseases. I will now discuss the factors leading to a decrease in the sensitivity and specificity of cytological interpretation. The factors influencing the diagnosis of PDAC form cytology samples are divided into quantitative and qualitative factors.

#### 1.7.4.1 Quantitative factors

Quantitatively, the cytological sample obtained may be of low cellularity with few, or even no malignant epithelial cells amongst a variety of cell types. This is related to sampling technique error. Adequate cellularity of samples can be achieved by training the endosonographer and the presence of an onsite cytopathologist for cytologic assessment (93). The British Society of Gastroenterology guidelines suggest an intensive year of training in EUS-FNA from a recognised skilled operator (94). Furthermore, the on-site evaluation of the specimen decreases the chances of inadequacy of sampling. This may also reduce the number of needle passes and hence the complications associated with FNA (95). Time and money are the most important factors precluding the on-site cytology evaluations in most institutions (93).

#### 1.7.4.2 Qualitative factors

Qualitatively, the malignant epithelial cells can be morphologically similar to reactive pancreatic cells, especially in well-differentiated adenocarcinomas. Chronic reactive changes arising from atrophy or inflammation in pancreatitis are common, and also make diagnosis of adenocarcinoma difficult.

**PDAC in the setting of chronic pancreatitis poses a diagnostic challenge**

Chronic pancreatitis (CP) is an important benign differential diagnosis of PDAC. The clinical and imaging similarities between PDAC and CP may make the diagnosis difficult (96, 97).

The sensitivity of EUS-FNA for diagnosing PDAC decreases to 62% in CP and to 50% in CP and obstructive jaundice with a biliary stent (98). This occurs because patients presenting with obstructive jaundice ordinarily undergo ERCP and a biliary stent is placed to relieve symptoms before they are referred for EUS±FNA. The inflammatory and degenerative changes induced by instrumentation (ERCP) and biliary stent may mimic well differentiated adenocarcinoma and make cytologic interpretation troublesome (78, 99, 100). CP leads to both false positivity and false negativity in the diagnosis of pancreatic cancer.

Cytomorphologic features such as a disorganised honeycomb pattern of epithelial groups, variable degrees of anisonucleosis and irregular nuclear membranes may provide a diagnostic indication of pancreatic cancer (91), but the reactive process in CP induces cellular atypia with overlapping cytologic features similar to those resulting from a well differentiated PDAC (101).

Therefore, chronic pancreatitis is by the far the most important confounding factor in the diagnosis of PDAC and caution should be exercised in a suspicious PDAC vs. CP scenario and again biomarkers may be helpful in such cases.

**Gastro-intestinal and pancreatic contaminants**

EUS-FNA can be performed via a transgastric or transduodenal route depending on the location of suspected lesion. A transduodenal approach is a common practice for lesions involving the head of pancreas, whereas the transgastric approach is generally used for lesions of the body and tail of the pancreas. The cytological samples thus obtained may contain gastric or duodenal contaminants and it is essential to recognise the morphology of these cells to avoid misinterpretation of cytology specimens (90). These epithelial cell contaminants may also originate from the adjacent normal pancreas and may contain blood which further renders the diagnosis challenging (68, 90).

One study demonstrated the presence of normal duodenal, gastric and pancreatic acini in 52%, 30% and 94% of positive/suspicious, atypical, and negative cytology cases respectively (90). A mixture of normal and abnormal elements when normal cells are in abundance may mask the few malignant cells or may mimic malignant cytology. Thus differentiation between normal and abnormal cells particularly becomes troublesome in well-differentiated adenocarcinoma. Examination of hematoxylin and eosin (H&E) stained sections from cell blocks may help in the final diagnosis in such cases. Apart from morphological features, biomarkers expressed in PDAC and not in these gastrointestinal contaminants might potentially help to confirm or exclude the diagnosis of PDAC.

**Cytopathologist expertise**

The interpretation of cytology specimens obtained via EUS-FNA or ERCP brushing is expertise dependent and can present many challenges for cytologists especially in institutions with a low number of cytology cases (102). One study evaluated the impact of experience of the cytopathologist on the sensitivity of EUS-FNA cytology. The sensitivity of difficult cytology cases improved after evaluations by an experienced cytopathologists (103). Well differentiated PDAC are especially difficult to diagnose from cytology specimens (104).

#### 1.7.4.3 False negative diagnosis leads to low sensitivity and is a relatively more significant problem than false positive diagnosis

False negative diagnosis from EUS-FNA samples for suspected pancreatic cancer occurs in a significant proportion of patients (78, 90, 105-107). The false negative rate of EUS-FNA increases further in patients with obstructive jaundice and biliary stent compared to patients without obstructive jaundice and without a biliary stent at presentation (81). In biliary stent cases, this occurs due to secondary reactive cellular atypia confounding the diagnosis of well differentiated cancer. Therefore a negative diagnosis of malignancy may not exclude malignancy due to false negativity (89). The causes of false negativity include sampling error, chronic pancreatitis, cytological misinterpretation, reactive changes due to bile stent and epithelial cell contamination (89, 107).

The false negativity of EUS-FNA in suspected pancreatic cancer could potentially limit its clinical utility as it may miss early resectable diseases (56, 108). These results suggest that the sensitivity of EUS-FNA cytology is not perfect and a false negative diagnosis can occur. Therefore, improving the sensitivity for patients with other ancillary tests, possibly biomarkers, will reduce the possibility of a false negative diagnosis and improve the sensitivity of the cytology diagnosis.

#### 1.7.4.4 False positive diagnosis leads to low specificity but is not a significant problem in PDAC cytology

The currently accepted standard is that a positive EUS-FNA result of a suspected pancreatic lesion provides evidence of malignancy and the cytological positivity can be regarded as a “true-positive” result (109, 110). In addition, a high

specificity (98%) reported in a recent meta-analysis by Hewitt et al suggests that false positivity of EUS-FNA is rare (68). However, false positive results can occur in EUS-FNA specimens from pancreatic lesions suspected of PDAC and recognising this risk is important due to its impact on the surgical and/or medical management of patients (93, 110-112).

A large retrospective cohort study (n=367) was conducted in a tertiary care referral hospital to assess the false positive rate of EUS-FNA cytology in patients presumed to have pancreatic cancer who underwent surgical resection. The false positive rate was 1.1% when "positive" cytology was interpreted as malignant but about 4% when both "positive" and "suspicious" cytology were interpreted as malignant. Among these false positive cases three were falsely diagnosed as positive for adenocarcinoma and six were falsely diagnosed as suspicious for adenocarcinoma (110). The causes of false positivity include chronic pancreatitis, cytological misinterpretation, reactive changes due to bile stent, tissue sampling technique error and epithelial cell contamination (111).

False positive diagnosis increases further when atypical and suspicious cytology are included in the positive diagnosis (68). Therefore, improving the specificity of patients with other means (for example biomarkers) apart from morphological interpretation will reduce the possibility of a false positive diagnosis and improve the specificity of cytology diagnosis.

### 1.7.5 Diagnosis is important for PDAC management

The clinical management of patients with both false negative and false positive diagnosis leads to significant outcomes. False positive diagnosis leads to unnecessary surgical or oncological treatment and false negative diagnosis delays treatment for patients with potentially resectable disease (107).

Pancreatoduodenectomy is a major surgical procedure and is associated with significant mortality and morbidity. The mortality rate ranges from 0% to 7% and the morbidity rate ranges from 21% to 59%, therefore unnecessary surgery of benign pancreatic lesions must be avoided (38, 61).

Preoperative EUS-FNA to confirm the diagnosis in every patient with resectable pancreatic cancer is debated (113, 114) but Kudo et al (61) showed that it does not adversely affect surgery or prognosis in these patients. Therefore, by providing a definitive diagnosis, EUS-FNA can potentially improve the outcomes of patients resected for pancreatic cancer.

Diagnostic confirmation of PDAC is especially important for patients with unresectable disease (locally advanced and/or metastatic disease) or not fit for surgery, which may require neoadjuvant therapy. Therefore, the significance of a false negative diagnosis increases even further when the oncologist is considering neoadjuvant therapy. In addition, false positive diagnosis might lead to offering neoadjuvant therapies to patients with benign disease (for example CP) which could potentially have deleterious effects. Patients with a high suspicion of PDAC but with false negative FNA cytology should not be precluded from surgical intervention (28). However, there are cases in the reported literature where pancreatic cancer could not be confidently ruled out and surgery was performed for misdiagnosed focal pancreatitis (76, 77). False negative diagnosis limits the clinical utility of EUS-FNA as it could potentially lead to missing the early resectable disease in patients suspected of pancreatic cancer (56, 112, 115). 'Atypical' cytology raise the clinical suspicion of pancreatic cancer but ordinarily this may not be sufficient evidence to subject patients to surgical resection (112). However, positive for malignancy and suspicious cytology usually results in surgical intervention.

The 'atypical' cytology diagnosis clearly decreases the sensitivity and leads to an indeterminate diagnosis and the sensitivity further decreases if 'suspicious' diagnosis is also included. Cytologic confirmation of a PDAC diagnosis is therefore very important for the management of patients with PDAC, and improving the diagnosis would mean improved PDAC management.

## 1.8 The unmet clinical need

The previous discussion indicates that the diagnosis of PDAC from EUS-FNA and ERCP brush cytology is not perfect and there is an unmet clinical need in improving the diagnosis (especially diagnostic sensitivity) of cytology. Improving

the diagnostic sensitivity will reduce the number of samples called ‘atypical’ or ‘suspicious’.

Management algorithms for PDAC are evolving and new strategies such as pre-operative chemoradiation to achieve R0 resections (clear resection margins) are being investigated. Thus the importance of confirming pre-treatment diagnosis from cytology samples is further increasing. Using ancillary tests (such as biomarkers) as an adjunct to cytology could be an invaluable tool in assisting in refining the diagnosis of PDAC and improving the diagnostic sensitivity.

## 1.9 Biomarkers could help improve the diagnosis of PDAC from cytology

One potential way of improving the diagnosis of PDAC from cytology samples is to use biomarkers as an adjunct to cytology. Biomarkers might potentially be indicated in the following clinical scenarios: EUS-FNA or ERCP brush cytology is unable to distinguish PDAC from a benign mass in the setting of chronic pancreatitis; hypocellular material from FNA; indeterminate cytological diagnosis but a high clinical suspicion of PDAC (110).

Therefore, in the subsequent sections I will discuss what are biomarkers? What are the phases of biomarker development from identification to validation and to clinical use? Why is IHC an attractive platform for biomarker investigation? What is the role of IHC biomarkers in improving the diagnosis of PDAC from cytology?

Then I will discuss: what are the major reasons or factors delaying or hindering the clinical translation of these biomarkers? What is the significance of meta-analysis in gathering, synthesising and ranking the published evidence on IHC biomarkers in PDAC? And why are TMAs an excellent high throughput platform for validation of biomarkers?

### 1.9.1 Tumour biomarkers

According to the Biomarkers Definitions Working Group (2001) (116), a biomarker is defined as “A characteristic that is objectively measured and evaluated as an

indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention."

Tumour markers are molecules produced by cancer cells and either endogenously present in the cellular compartments of cancer cells or are secreted from the cells. They are measured in blood, urine, stool, body fluids (for example pancreatic cyst fluid), or tissues of patient with cancer. Tumour markers are often proteins but genetic changes (gene mutations) and changes in gene expression patterns are also used as tumour biomarkers. The alterations in the tumour biomarkers help in the categorisation of patients into distinct groups (117).

Biomarkers are used in the clinical management of patients with tumours and can broadly be classified as: susceptibility biomarkers that help in the identification of individuals at risk of developing cancer; screening biomarkers that help in the early detection of cancer in the general or at risk populations of developing overt disease (i.e. detects subclinical disease); diagnostic biomarkers that help in the diagnosis of patients with the disease; prognostic biomarkers that help to predict the course of disease (for example survival); predictive biomarkers that help to predict response to therapy (for example a drug or surgical intervention) or monitor the efficacy of a therapy (116, 118-120); and pharmacogenomics biomarkers which are "measurable DNA and/or RNA characteristics that are indicators of normal biologic processes, pathogenic processes, and/or a response to therapeutic or other interventions" (121).

Some of these markers are already used in routine clinical practice (122-125). Perhaps the earliest markers to help in the clinical diagnosis were carcinoembryonic antigen (CEA) in colon carcinomas (126) and prostate specific antigen (PSA) (127) in prostate cancer.

### 1.9.2 Phases of biomarker development from identification to validation and to clinical use

Biomarker development passes through several phases of development from discovery to clinical practice (128). It involves a series of identification and validation steps before clinical application (Figure 1.7).

### 1.9.2.1 Identification of biomarkers

The first phase in biomarker development is the identification of suitable candidate biomarkers (129). Here I will discuss the identification of biomarkers through differential expression patterns of genes and proteins. Identification passes through the following stages.

***Discovery:*** Broadly two approaches can be used for biomarker identification. The first approach is to identify biomarkers based on current knowledge of pathophysiology of the disease through 'deductive reasoning'. The second approach is using molecular profiling techniques to identify candidates based on the differential expression between tumour and normal tissue (128). Differentially expressed genes between PDAC and normal or reactive pancreatic duct are identified through high throughput genomic and proteomic studies (130, 131).

***Demonstration:*** High throughput technologies generate a list of potential biomarkers but based on different statistical models the biomarker list is further refined and selected biomarkers are demonstrated by molecular techniques. The differential expression of genes is demonstrated by DNA microarray and polymerase chain reaction (132, 133), whereas, differentially expressed proteins are demonstrated by gel electrophoresis and mass spectrometry (134-136).

***Qualification:*** the purpose of the qualification phase is to confirm the differential expression of biomarkers using alternative techniques such as western blot and immunohistochemistry (134-137).

The purpose of the identification phase of biomarker development is to identify potential candidates with high sensitivity for detection. The emphasis therefore is to establish the association between biomarker expression and the tumour of interest. Biomarker identification thus uses a significant amount of resources, cost and utilizes modern technology.

### 1.9.2.2 Validation of biomarkers

After identification, the next phase is validation of biomarkers which is an important pre-requisite for clinical translation. 'Omics' technologies allow

identification of promising biomarkers but these biomarkers require verification, prioritisation and validation before they are used in clinical practice (138).

***Verification:*** Biomarker verification is carried out to test whether the candidate biomarker has sufficient potential for future validation studies. Pilot studies in a relatively small sample size are used to investigate biomarker expression in both tumour and normal samples from a variety of patients (139, 140). Verification begins to assess the specificity of biomarkers but still focuses on optimum sensitivity (128). This helps the researchers to select more specific candidates that are highly expressed in tumour for which they can invest their time, energy and money.

***Prioritisation of candidates:*** Prioritisation of candidate biomarkers from a list of potential biomarkers is very important for further clinical validation studies due to cost and limited clinical resources (129). The role of the candidate biomarker in tumour biology greatly facilitate this selection process as markers involved in the progression of tumour will prove potentially more useful in clinical practice (123, 141).

***Validation of selected candidates in large-scale studies:*** Biomarkers achieving a suitably good combination of sensitivity and specificity from pilot studies in the verification and prioritisation phases may be selected for further validation in large-scale studies. The further validation processes consist of three phases: analytical validation ensures the intra- and inter-laboratory reproducibility of the assay (for example IHC) achieving similar expression patterns; clinical validation ensures the diagnostic sensitivity and specificity of the biomarker is consistent for the outcome (for example differentiation between benign and malignant disease); and clinical utility of the biomarker assesses whether it improves the diagnostic management of patients (142).

Early validation studies are carried out on archival pathology specimens (tumour and normal). These samples are retrospectively identified and used to observe the expression and clinical utility of candidate biomarkers. These retrospective studies typically overestimate the actual sensitivity and specificity of biomarkers. Most of the reported literature on biomarker studies uses archival pathology samples (143, 144). However, the clinical utility of biomarkers can be

more clearly assessed in a prospectively designed study. But validation on archival samples is a pre-requisite before biomarker investigation in prospective clinical studies. Prospective clinical studies and biomarker trials lead to the qualification of biomarker for clinical use. Finally, biomarkers are used in clinical practice for the intended clinical use and they are monitored for their effectiveness. Figure 1.7 shows the path that a biomarker is likely to take from discovery to clinical use.

***Importance of validation studies:*** Biomarker validation is therefore an important but expensive and lengthy process and depends on the type of samples used for assessing the clinical utility. Successful implementation of biomarkers in clinical practice requires robust evidence from independent validation studies. A single study is unlikely to provide sufficient evidence for adoption of a biomarker in clinical practice. In a study (145) the magnitude of the effect size of proposed biomarkers in highly cited papers was examined. It was found that primary studies often report a larger effect size compared to the subsequent meta-analysis assessing the same associations (145). Therefore, clinical evidence from biomarker studies should be interpreted carefully and healthy scepticism is suggested (145). More large-scale and outcome-oriented validation studies expedite the clinical translation of biomarkers by providing a strong 'evidence base'.

In summary, careful identification of biomarkers and then validation in well-designed retrospective and prospective studies is a systematic strategy for developing clinically useful biomarkers.

### 1.9.3 Immunohistochemistry biomarkers in cancer pathology

Immunohistochemistry (IHC) is a tissue based method that allows the visualisation of specific antigens in tissues and cellular compartments based on antigen-antibody reaction using microscopy (122). IHC remains an important diagnostic tool even in the era of genomics and high throughput molecular diagnostics. IHC was first introduced in the 1960s and since then the amount of literature has increased exponentially. IHC has been used both in research and clinical settings. IHC characterises the expression of genes at the protein level and allows the observation and localisation of protein expression simultaneously

in tissue and cellular compartments (146, 147). It is a routine technique used in diagnostic pathology and is relatively inexpensive with widespread expertise in the technique. Therefore, biomarkers identified and validated by IHC have an enormous potential for clinical translation. In surgical pathology a range of biomarkers in clinical use is assessed by IHC (122, 123, 148).

Biomarkers identified by IHC have the advantage of defining the role of markers in the tissue context. They give insight into the expression of markers in specific cell types of tissue (malignant cells, stroma, and adjacent normal cells or other cell types) and the distribution of the marker in subcellular compartments (nuclear, membranous or cytoplasmic). Biomarker expression in a specific cell type (e.g. epithelial cells) or subcellular compartment (e.g. cytoplasm) might then be associated with tumour diagnosis. Biomarker expression in tumour can also be associated with the clinical follow-up (e.g. survival of patient) or the clinicopathologic characteristics (e.g. lymph node invasion) of the patient (149, 150).

Over the last few decades many IHC biomarkers have been investigated for improving the diagnosis and prognosis of tumours. The diagnostic IHC biomarkers help in the diagnosis and sub-classification of tumours. IHC biomarker c-kit helps in the diagnosis of gastrointestinal stromal tumours (GISTs) (124). p63 helps to detect the presence of basal cells which indicate normal prostate gland (125, 151). Furthermore, nuclear immunostaining of ki-67 as a proliferation marker (152), chromogranin A, CD56 and synaptophysin for the diagnosis of neuroendocrine tumours (153, 154) and the use of E-Cadherin in the differentiation of ductal and lobular carcinomas of the breast (155) are used in clinical practice.

An ideal diagnostic IHC biomarker should be 100% sensitive and specific which is almost never achieved as sensitivity increases at the expense of specificity and vice versa. The panel of biomarkers are thus becoming more relevant. These include CK20, P53, CK5/6, CD138, and Her2/Neu in the diagnosis of urothelial carcinoma in situ (156), a panel of napsin-A, Thyroid transcription factor 1, Cytokeratin 5, and P63 in differentiating adenocarcinoma from squamous cell carcinoma of the lung (157, 158)) and a panel of S100P and XIAP in the differentiation of pancreatic cancer from non-neoplastic pancreatic tissue (159).

In addition, IHC biomarkers are used for predicting the survival of patients, predicting the response to specific therapies and subsequent stratification of patients for different treatment options. Estrogen receptor (ER), progesterone receptor (PR) and HER-2/neu are used for the management of patients with breast cancer (123, 160-162). Panel of biomarkers are also used for prognostic and predictive purposes e.g. IHC4 (a panel of ER, PR, HER2 and Ki-67) is an assay which estimates recurrence risk for early stage breast cancer patients (163).

IHC biomarkers are therefore useful tools that could potentially be translated to clinical practice if suitable biomarkers are identified and validated in independent cohorts.

### 1.9.4 The role of IHC biomarkers in PDAC cytology diagnosis

The molecular profiles of PDAC identified through genomic and proteomic studies increase hopes that clinically useful biomarkers can be developed for PDAC. Genes showing differential expression between PDAC and normal pancreas could potentially be used as biomarkers (130). These genes are characterised at the protein level by IHC. IHC biomarkers have great translational potential in improving the diagnosis of PDAC from cytology samples.

For this project, the purpose of IHC diagnostic biomarkers in cytology samples is to differentiate PDAC from benign disorders and not the distinction of PDAC from other pancreatic or extra-pancreatic neoplasms. Importantly, these biomarkers are not intended to replace the cytomorphology but if used as an adjunct to cytology they could help improve sensitivity of PDAC diagnosis and subsequent surgical or oncological management.

It is evident from the PDAC literature that both resection and cytology specimens have been used for the investigation of potentially useful IHC biomarkers (164-167). A candidate diagnostic IHC biomarker may pass through different stages of identification and validation. Validation studies are performed first on resection specimens from patients with PDAC to optimise IHC methodology and to study the expression patterns of biomarkers in various tissue and cellular compartments. Biomarkers with high expression levels resulting in

higher sensitivity without compromising specificity are then validated in cytology specimens.

Cytological specimens processed into formalin fixed paraffin embedded cell blocks provide an ideal platform for IHC. Cell block sections may perhaps display better morphology of cells than direct smears and can also show a mixed population of benign and malignant cells (168).

A wide range of IHC biomarkers has been investigated in pancreatic resection specimens and EUS-FNA and biliary brush cytology specimens (141, 169-172). These candidate diagnostic IHC biomarkers have been investigated both as single biomarkers and as part of biomarker panels to improve the diagnosis of PDAC. (159, 170, 172-177).

However, despite the extensive amount of research on diagnostic IHC biomarkers for PDAC, none of these biomarkers has yet entered routine clinical practice. This increases the importance of identifying and addressing factors delaying the clinical translation of these biomarkers. Therefore, in the next section, I will discuss these issues hindering the clinical translation of IHC biomarkers.

## 1.10 Reasons delaying the clinical translation of biomarkers

Clinical translation of basic research is a priority for both academia and industry (178). Translational research has gained more importance from the fact that the level of investment in research is not reflected in clinical practice, and there is a concern that the benefit from the 'genetic revolution' is slow.

The amount of diagnostic biomarker research in every cancer type is enormous but unfortunately the translation of biomarkers to clinical practice is comparatively low (162). I have outlined below six major reasons or factors that contribute to the delay in the clinical translation of the candidate diagnostic biomarkers under investigation (Figure 1.8). All these reasons revolve around the identification and validation of biomarkers and their scoring systems and thresholds. These reasons are not limited to biomarkers in pancreatic cancer but can be applied to other cancers.

### 1.10.1 Lack of synthesis and quantification of existing evidence

The most important first step is identification of existing diagnostic biomarkers. Biomarkers have been investigated purposefully to address a clinical problem (157, 179) and in some cases they have been investigated in more than one study (174, 180). The results of some of these studies show promising results for differentiating one tumour type from another tumour type (179) or subtypes of the same tumour (157). These studies also show the utility of biomarkers in the differentiation of benign disease from malignant disease (132, 166). However, the synthesis of evidence and quantification of biomarker performance is described less in the literature especially for diagnostic biomarkers. Therefore, there may be a need to perform more focused systematic reviews and subsequent meta-analysis of biomarkers intended for specific clinical (diagnostic) problems. The clear advantage of this approach is the identification of suitable candidate biomarkers that have previously been investigated. The other important advantage of this selection process is that biomarkers investigated in different studies showing promising results are compiled for investigation in one potential validation study. This will surely provide more strength for a future validation study investigating these better biomarkers in a single setting.

### 1.10.2 Inadequate sample size

The sample size for biomarkers reported in literature is relatively small and this is especially true for novel biomarkers assessed in pilot studies (142, 181). This is realistic because a vast tissue resource will not be available for a new biomarker under investigation. Sufficient statistical power is thus not reached in most of these pilot studies which could potentially lead to promising biomarkers being overlooked in the enormous biomarker research field. Biomarkers identified in the meta-analysis might be carefully investigated in a sufficiently powered study using a large sample size. Obviously, investigation of biomarker expression in large sample size will further elucidate the diagnostic performance (sensitivity and specificity) of biomarkers. Biomarker expression is not homogenous in tumour tissue extracted from different patients (182, 183). This Inter-tumour heterogeneity of expression of biomarkers can be shown more clearly if the

sample size is large. This will help in better defining the diagnostic role of biomarkers in cancer. Biomarkers showing more inter-tumour heterogeneity are less sensitive in identifying the disease and are thus less accurate.

The other issue with sample size is the distribution of the number of samples between the two groups (disease and normal) investigated (142). For example, a diagnostic biomarker tested to differentiate between benign and malignant disease. The true diagnostic potential of a biomarker in this case can be measured if (ideally) the sample size is equally distributed between benign and malignant samples.

### 1.10.3 Lack of an optimal scoring system and threshold

Interpretation of IHC requires a robust and comprehensive scoring system that is able to quantify the extent of biomarker expression. From this scoring scheme, then, thresholds or cut-offs can be investigated for categorising patients into one of the two diagnostic categories (for example benign vs. malignant). There is no single uniform scoring system and researchers have used a wide array of scoring systems (157, 184-186). These scoring systems are based on: staining intensity; percentage of positively stained cells; combination of both staining intensity and positive cells; and semi-quantitative Histoscores; (160, 171, 187, 188). A semi-quantitative Histoscore (takes into account both intensity and proportion of staining) could possibly emerge as a standard scoring system. Histoscore quantify the expression level of biomarkers and allow for calculation of various potential cut-offs for diagnostic purposes.

The next issue after scoring is choosing an appropriate cut-off that is easy to use by practising pathologists. An optimal cut-off should be reliable and reproducible among pathologists. A systematic way of choosing a cut-off is to perform a receiver operating characteristic (ROC) curve analysis which provides diagnostic sensitivity and specificity of biomarkers on a range of cut-offs (189, 190). This helps the researcher to select an optimal cut-off which has both diagnostic potential and is easily scored by observers. This cut-off can then be used in future validation studies and studies involving observer variations between different scorers.

### 1.10.4 Limited use of panels of biomarkers

An ideal diagnostic biomarker should have homogenous expression within and between tumour tissue from the same cancer type (for example PDAC), but both intra-tumour and inter-tumour heterogeneity of biomarker expression exists (191). A single candidate biomarker is thus unlikely to work as a perfectly sensitive and specific biomarker (122, 192) in all patients. A panel of biomarkers is thus a plausible solution to address both inter- and intra-tumour heterogeneity. Most of the researchers in reported literature have investigated biomarkers singly with a limited panel approach. The panel approach has not been reported in instances when more than one biomarker was investigated in a single study (173).

Identification of suitable biomarkers and then exploring their diagnostic performance as a panel in a single experimental setting is a more powerful approach. The obvious strength of the panel approach is that it allows for the comparison of accuracy between biomarkers and panels of biomarkers. This comparison then determines an appropriate panel of biomarkers for future validation and clinical translational studies. Different biomarkers stain different cellular compartments and using a panel of biomarkers has this additional advantage of staining all major sub-cellular compartments. Clearly, a panel with more than one positive biomarker provides more confidence to the pathologist reporting the disease.

### 1.10.5 Technical differences in IHC between laboratories

Optimisation of IHC in biomarker research is very important to achieve appropriate staining in the tissue. The manufacturers normally provide information and suggested protocols for IHC but most research laboratories optimise antibodies. The aim of optimisation is to increase the strength and specificity of the signal while suppressing background signals and artefacts (122, 146).

Research laboratories usually employ different IHC experimental conditions including clone of primary antibodies, antigen retrieval methods (heat induced epitope retrieval vs. enzymatic retrieval), primary antibody dilutions and manual

or automated platforms (193). All these factors significantly contribute to the different sensitivity and specificity values reported for a candidate biomarker investigated in different studies. Studies have addressed this issue and have compared: different clones of antibodies; different PH of antigen retrieval buffers (for example pH 6 vs. pH 8) and dilutions of primary antibodies (for example 1/50 vs. 1/100) (194-197). These studies provide an insight for choosing better conditions for the optimisation of antibodies.

One approach to addressing the issue of technical heterogeneity is to systematically search the literature to identify IHC parameters for a biomarker that achieved an optimum combination of sensitivity and specificity. These parameters could then be used as a starting point for further optimisation.

Assay development is a critical component in the qualification of biomarker. Sometimes biomarkers fail to enter the list of potential biomarkers not because of the underpinning science, but because of issues around assay development and a lack of validation studies (178, 198).

### 1.10.6 The need for well-designed validation studies

Validation of potential IHC diagnostic biomarkers in independent tissue cohorts is probably the most important factor delaying clinical translation (181). Researchers investigate promising biomarkers, publish their work and sometimes leave excellent biomarkers without designing further validation studies. More focused and aim oriented validation studies could expedite the journey of biomarkers from bench to clinic (128, 142).

Validation studies in a step-wise fashion can be as follows. Validation of biomarkers in independent laboratories and patient cohorts, using the same IHC methodology, and the same scoring system and cut-offs. The expression level and subsequent diagnostic sensitivity and specificity should broadly be similar in validation studies. This will help to establish the reproducibility of the IHC methodology and cut-offs used for diagnostic purposes. Then establishing a multi-institutional validation study group and carrying out validation studies and addressing technical and other issues (199). For example a study group developed in pancreatic cancer research is ‘European Study Group for Pancreatic

Cancer' (ESPAC) (200). Finally, a prospective clinical study for investigating the optimum biomarker panel will provide more confidence to translational scientists and pathologists for further validation. In fact, academic-industry collaborations can further facilitate and expedite validation studies from bench to clinic.

In summary, biomarker development from identification to validation and clinical application would require: pooling of already existing data; synthesis and then analysis of evidence; identification of known promising biomarkers; addressing factors such as sample size, scoring systems and cut-offs that influence the validation of biomarkers; and finally using a panel approach and best IHC methodology in well-designed and aim oriented validation studies.

In the next section, I will discuss the importance of systematic review and meta-analysis in identifying suitable biomarkers that are already at the verification or validation stage of biomarker development.

## 1.11 Importance of systematic review and meta-analysis

Cancer biomarker discovery through 'omic' technologies has generated a long list of biomarkers. However, a labour intensive and costly filtering process is required to screen all biomarkers and identify potentially useful candidates. Various factors are involved in identifying suitable candidates including sensitivity i.e. the overexpression of biomarkers in cancer, specificity i.e. no or low expression in normal tissue and other tumour types and an appropriate assay (201).

One potential way of addressing the issues hindering the clinical translation is to identify biomarkers that have already passed through some phases and have been validated by IHC. The sensitivity and specificity values of these potential biomarkers will facilitate in further filtering the most suitable candidates. Thus, at the identification step of this project we performed a systematic review and meta-analysis of potential IHC diagnostic biomarkers.

Due to the scattered nature of literature on biomarker studies it was necessary to collect the existing evidence on diagnostic IHC biomarkers for PDAC. Every

time a novel diagnostic biomarker is reported or a validation study for an existing biomarker is reported, the translational scientist and pathologist compare it with the current evidence to assess its clinical significance. But simply summarising the scattered scientific information from the literature may not be enough. In addition, relying on the results of a single 'good' quality study especially if performed at one institution may not provide a rationale for the clinical use of a candidate diagnostic biomarker.

Systematic review and meta-analysis of randomised controlled trials is the highest level of evidence in 'evidence based medicine'. Systematic review attempts to identify the best research evidence from literature, appraise it critically, synthesise evidence for clinical practice and identify papers for meta-analysis (202, 203). Most of the reported studies in PDAC diagnostic biomarker research are either cohort (PDAC and non-neoplastic adjacent tissue from the same patient) or case control studies (PDAC and non-neoplastic tissue from separate groups of patients). Therefore, the best available evidence that could potentially be provided for diagnostic IHC biomarkers will be a systematic review and meta-analysis of cohort and case control studies.

Systematic review and meta-analysis not only quantifies and ranks the performance of potential IHC biomarkers but also assesses the quality of primary studies investigating biomarkers. It helps to evaluate primary studies by asking questions like: What is the design of the study? What is the sample size of primary studies and what is the sample distribution between malignant and benign samples? What is the IHC methodology and has the technique been described in sufficient detail to allow reproducibility in independent studies? Which scoring system and cut-off has been used and how much is their heterogeneity between studies? Were biomarkers investigated singly or in a panel? Indeed the answer to all these questions will help in evaluating the existing evidence, identifying issues with primary validation studies and making suggestions for future validation studies.

In summary, systematic review and meta-analysis is an invaluable tool for the identification of potential biomarkers from the literature, identifying translational issues with biomarker validation and providing a good base for future validation studies. After identification, these biomarkers were then

validated utilising tissue microarray technology. In the next section I will discuss the significance of tissue microarrays in biomarker research.

## 1.12 Tissue microarrays

Biomarkers identified through molecular profiling approaches or literature searches require independent validation in large scale tumour samples before clinical application. Previous tissue based validation platforms were whole tissue sections. But validation of IHC biomarkers on whole tissue sections is time consuming and is reagent and sample intensive. A high throughput platform exists for biomarker validation and it is called tissue microarrays (TMAs) (204, 205).

### 1.12.1 TMA technology

Utilising TMA technology allows for the examination of several hundred tissue cores from a number of patients on a single standard microscope slide (206). Briefly, TMAs are constructed by arraying small cores (0.6mm in diameter) of paraffin embedded tissue samples in a recipient wax block. Sections cut from the TMA block could then be stained by IHC using standard protocols.

The first step in the construction of TMAs is to identify formalin fixed paraffin embedded (FFPE) whole tissue blocks to act as donor blocks. The corresponding haematoxylin and eosin (H&E) section of tissue blocks is used to mark representative areas of tissue on the slide. Then FFPE blocks are sorted and arranged for array construction. The layout of the TMA is made with control samples from other tissue types to help orientate the subsequent sections after staining. The marked H&E sections are used to identify areas of interest in the donor block. A hollow needle is inserted into the donor block to ‘sample’ a tissue core. Another hollow needle is inserted into the recipient wax block and hole is created for insertion of the donor tissue core. This process is used following the TMA layout to create a full TMA block.

A very important consideration in the array design is taking into account the intra-tumour heterogeneity. The representative areas of the given tumour from multiple locations should be ‘sampled’ from a few donor blocks of the same

patient. The sampling of more cores from patient tissue blocks is associated with better representation of intra-tumour heterogeneity of biomarker expression. In a study it was shown that staining of one, two, three, four, six and ten cores represented the whole tissue section staining in 92%, 96%, 98%, 99%, 100% and 100% of cases (207).

Analysing multiple biomarkers on consecutive tissue sections can lead to rapid exhaustion of precious clinical material. Tissue cores 'sampled' for the construction of TMAs do not compromise the tissue block that could be retained in case it may be required for future diagnostic purposes such as immunostaining (204). In addition, multiple tissue cores can be sampled from one tissue block, generating multiple replicates of TMA blocks. For example, if we have 10 whole tissue blocks and each tissue block can be used for staining of 100 tissue sections then the total number of sections from these tissue blocks can be 10×100=1000. Now if we take multiple tissue cores from each of these 10 tissue blocks and constructs 5 replicates of TMAs. The resulting number of section from 5 TMA blocks will be 5×10×100=5000. Hence the number of replicates of TMA slides can significantly increase from a set of clinical tissue specimens utilising TMA technology.

TMAs are thus a speedy and cost-effective approach for biomarker research in molecular pathology that ensures the abundant supply of tissue for multiple techniques. Figure 1.9A shows an H&E image of a TMA section and Figure 1.9B&C show magnified view of TMA cores.

### 1.12.2 Uses of TMAs

TMAs are now routinely used at most research institutions for IHC based biomarkers in a high throughput fashion across large cohorts. In the past decade a wealth of research emerged utilising TMAs for validating diagnostic, prognostic and predictive biomarkers (208, 209).

The design of arrays differs for addressing different clinical questions and researchers have used tailor-made arrays to investigate specific research questions. TMAs designed specifically for diagnostic biomarkers have an array of tumour and normal tissue cores (210). They may also consist of different

histologic tumours from the same organ (157). Multi-tumour TMAs consist of tumour types from different organs and can be used for screening of biomarker expression patterns. TMAs have been used for assessing biomarker expression in a tumour progression model for example PanIn model of pancreatic cancer (211). TMAs can provide efficient platforms for investigation of prognostic and predictive biomarkers. The survival and clinicopathologic variables of patients including chemotherapy or other interventions are linked to tissue cores on TMAs. This enables the assessment of biomarker expression on TMAs in relation to patient survival, prediction of therapy and clinicopathologic variables (212, 213). TMAs also allow the standardisation of IHC parameters and assay development (214). TMAs provide an excellent and efficient platform for investigation of a panel of candidate biomarkers, thereby helping to demonstrate the most reasonable panel discriminating a disease process from normal or an optimum panel predicting survival and response to specific therapies.

In summary, TMAs can substantially accelerate the validation of novel molecular discoveries coming from genomic and proteomic research. The clinical, pathological, follow-up and therapy response information from cancer patients can be quickly and efficiently correlated with multiple biomarkers. They can also be used as an invaluable tool for sub-classification of tumours which could subsequently lead to more personalised medicine approaches. TMAs thus enable assessment of biomarker in a high throughput manner saving clinical material, time and cost.

## 1.13 Project aim, objectives and implications

As described in the above sections, cytology samples obtained via EUS-FNA and ERCP with brushing are very important for confirming the diagnosis of PDAC. However, cytology diagnosis is not perfect and false positive and false negative diagnoses have significant influence on the management of patients with suspected pancreatic cancer.

Clearly, there is an unmet clinical need in PDAC diagnosis potentially leading to increased risk of ‘incorrect’ or ‘missed’ diagnosis and need additional diagnostic modalities such as tissue based biomarkers (172, 215-217). Therefore, the

identification and validation of diagnostic biomarkers in PDAC is vital to improving the management of this deadly disease. IHC biomarkers have long been established in clinical laboratories and provide an excellent and efficient platform for potential diagnostic biomarkers. The workflow of the current PhD project is illustrated in Figure 1.10.

### 1.13.1 Aim

The overall aim of the project is the identification of potentially diagnostic IHC biomarkers for PDAC and then validation of their diagnostic utility for improving the diagnostic management of patients with suspected PDAC from cytology samples.

### 1.13.2 Objectives

The following is a list of specific objectives:

**Identification of biomarkers**

1. To review the current evidence on diagnostic IHC biomarkers investigated in PDAC literature.

2. To identify candidate diagnostic IHC biomarkers through quantification and ranking in a meta-analysis

3. To assess the quality of primary papers reporting diagnostic IHC biomarkers and to assess the heterogeneity of IHC methodology, microscopic interpretation of immunostaining and diagnostic cut-offs.

**Validation of biomarkers**

4. To evaluate the expression level and validate the sensitivity and specificity (singly and in panel) of candidate diagnostic IHC biomarkers identified in meta-analysis in local PDAC and cholangiocarcinoma (CCC) resection tissue specimens.

5. To investigate and determine appropriate diagnostic cut-offs resulting in suitable sensitivity and specificity through receiver operating characteristics curve.

6. To investigate various IHC cut-offs for observer agreement amongst practising pathologists and to identify factors influencing the interpretation and scoring of different cut-offs.

7. To develop a resource of 'needling' cytology samples to enable optimisation of candidate IHC biomarkers for use in pancreatico-biliary cytology specimens.

8. To evaluate the expression level and validate the sensitivity and specificity (singly and in panel) of selected biomarkers in archival cytology samples.

Through this systematic approach it is envisaged that this project will identify and validate biomarkers with a potential to support cytological diagnosis and improve patient management.

### 1.13.3 Implications

The pre-treatment diagnosis of patients in suspected pancreatic cancer remains a challenge. This unmet clinical need could potentially be improved with IHC biomarkers. IHC is a routine and well established tool in pathology laboratories. Therefore, the identification and validation of potential IHC biomarkers with a higher diagnostic sensitivity and specificity will improve the current diagnostic management of patients with pancreatic cancer. Importantly, these biomarkers are not intended to replace cytomorphology but their proposed use as an adjunct to cytology could help diagnose the indeterminate or difficult to diagnose cases.

**Table 1.1: Pancreatic cancer staging**

| Stage | Tumour Grade | Nodal Involvement | Distant Metastasis | |
|---|---|---|---|---|
| Stage 0 | Tis | N0 | M0 | Carcinoma in situ (also includes PanIN-3) |
| Stage IA | T1 | N0 | M0 | Tumour limited to the pancreas, ≤2 cm in longest dimension |
| Stage IB | T2 | N0 | M0 | Tumour limited to the pancreas, >2 cm in longest dimension |
| Stage IIA | T3 | N0 | M0 | Tumour extends beyond the pancreas but does not involve the celiac axis or superior mesenteric artery |
| Stage IIB | T1, T2 or T3 | N1 | M0 | Regional lymph node metastasis |
| Stage III | T4 | N0 or N1 | M0 | Tumour involves the celiac axis or the superior mesenteric artery (unresectable disease) |
| Stage IV | T1, T2, T3 or T4 | N0 or N1 | M1 | Distant metastasis |

**Note:** T= Primary tumour, N= Lymph nodes and M=distant metastases. (American Joint Committee on Cancer 7th Edition (30))

**Table 1.2: The 2×2 table for calculating the sensitivity and specificity**

| | | True diagnosis of the disease (Histopathology or Clinical Follow up) | |
|---|---|---|---|
| | | D+ | D- |
| **Biomarker** | B+ | a (TP) | b (FP) |
| | B- | c (FN) | d (TN) |
| | | Sensitivity a/a+c | Specificity d/b+d |

**Figure 1.1: Anatomy of pancreas**

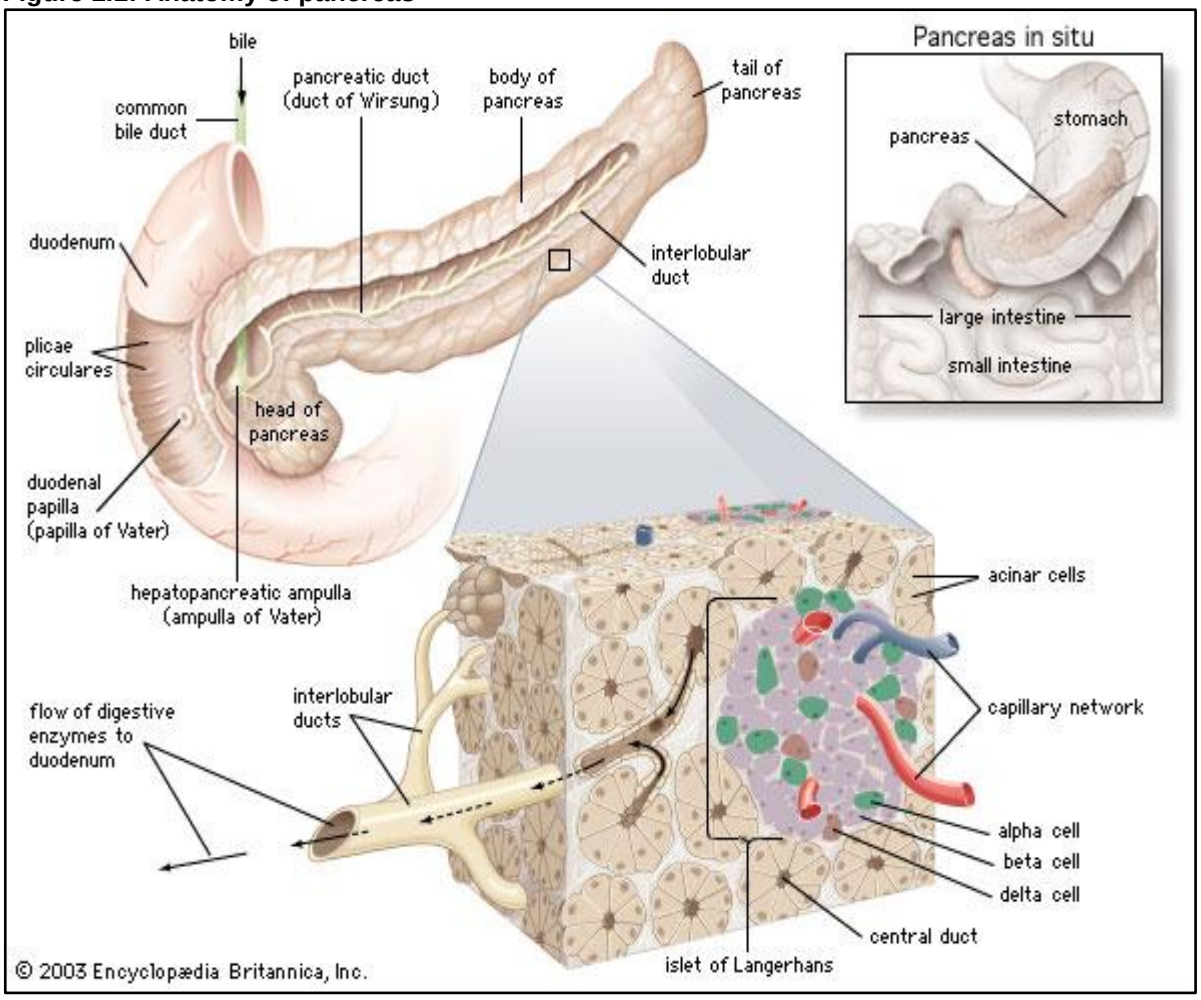



**Figure Legend:** Pancreas Anatomy, Art. By courtesy of Encyclopaedia Britannica, Inc., copyright 2003; used with permission.

**Figure 1.2: Histology of normal pancreas**

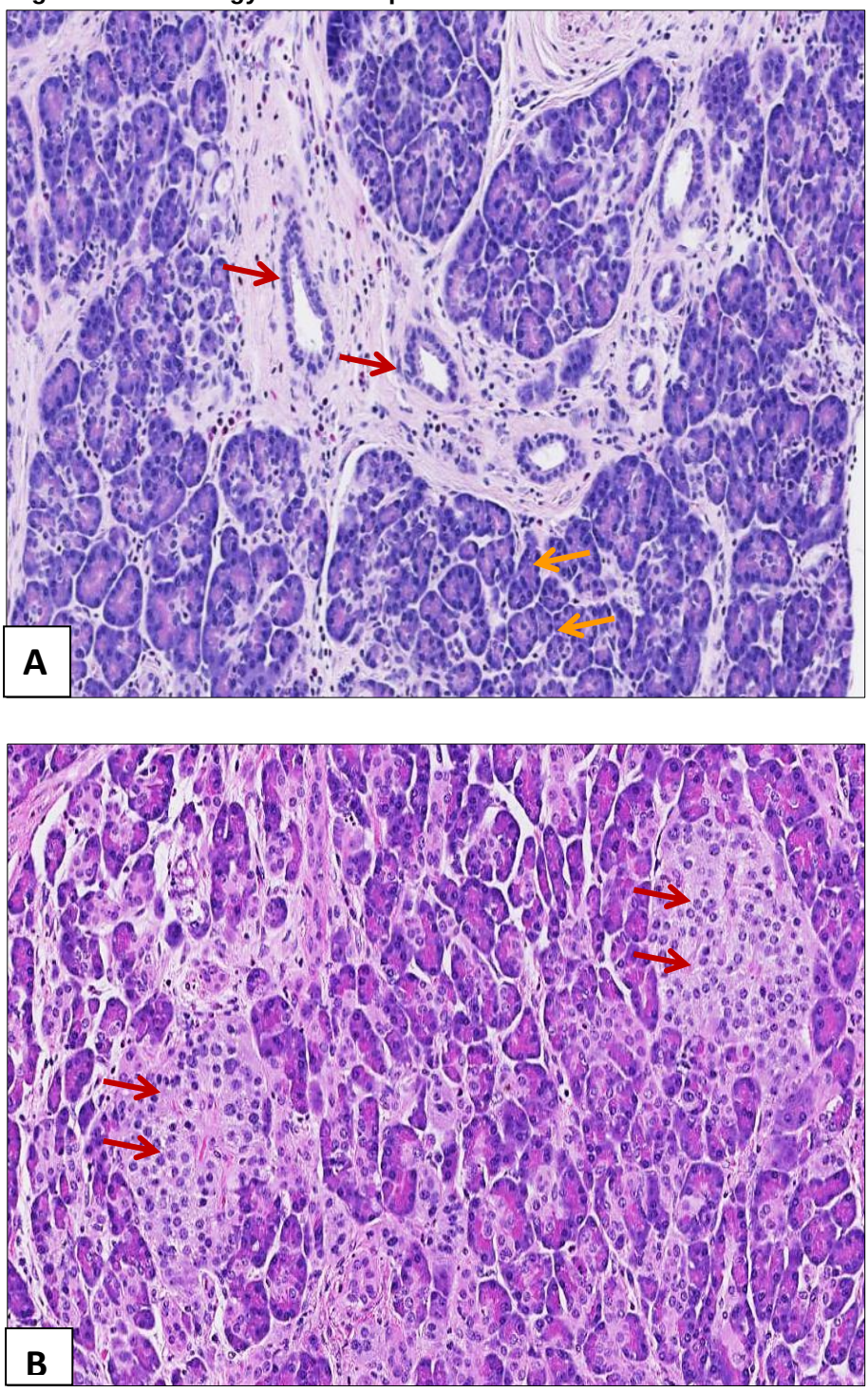



**Figure Legend: A)** Normal pancreatic tissue. Red arrows= normal pancreatic ducts are lined with cuboidal-low columnar epithelium, Orange arrows=normal pancreatic acini arranged in lobules constitute majority of the parenchyma. The apical portion of the cells is lightly eosinophilic (pink) due to the presence of zymogen granules and the basal part of cytoplasm has basophilia (purple). The nucleus is polarized to the periphery and the cells are arranged in round units to form acinus. **B)** Red arrows= Islet's of Langerhans consist of round collection of endocrine cells which have amphophilic cytoplasm and nuclei have fine dotted chromatin organization.

**Figure 1.3: PanINgram**

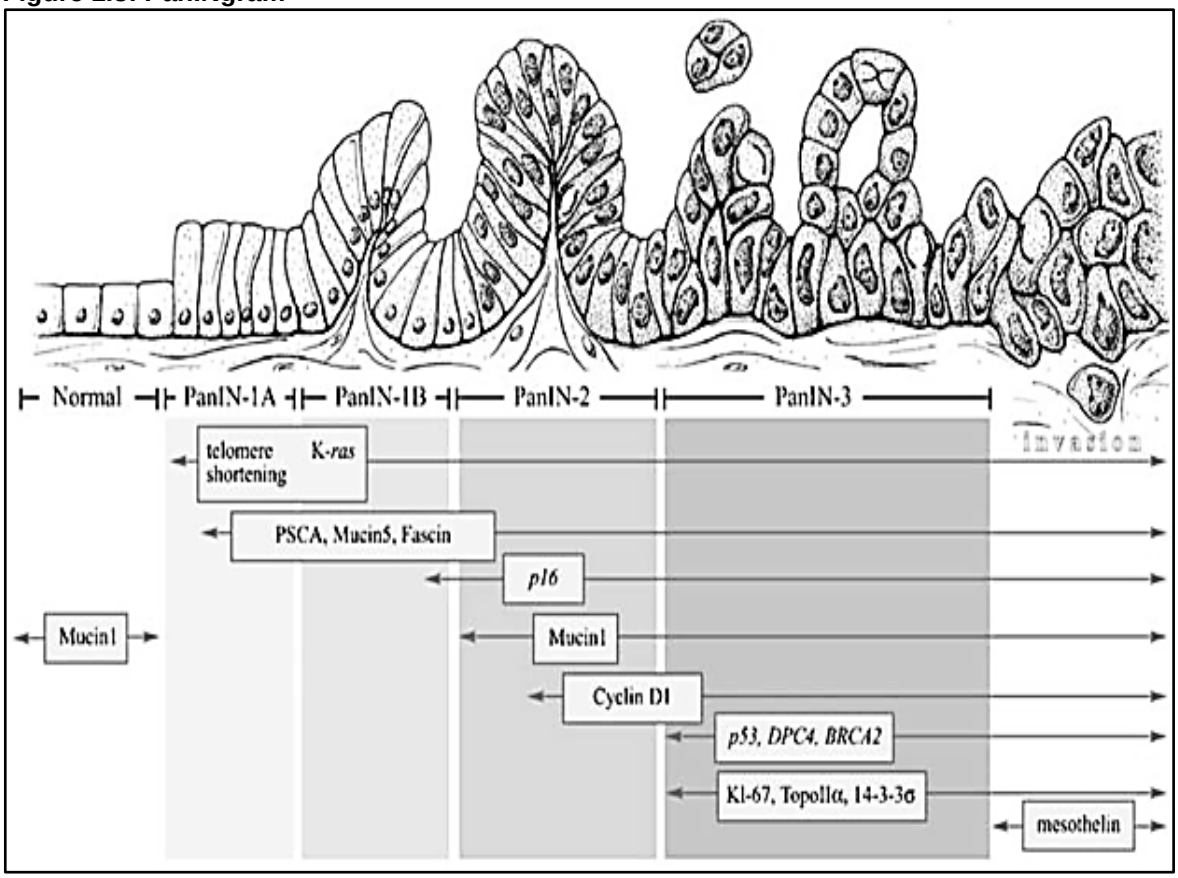



**Figure Legend: "**A PanINgram illustrating our current understanding of the molecular changes in the multistep progression model of pancreas adenocarcinomas" Reprinted by permission from Macmillan Publishers Ltd: Modern Pathology, Maitra et al. © 2003 (211).

**Figure 1.4: Haematoxylin and eosin images of PDAC**

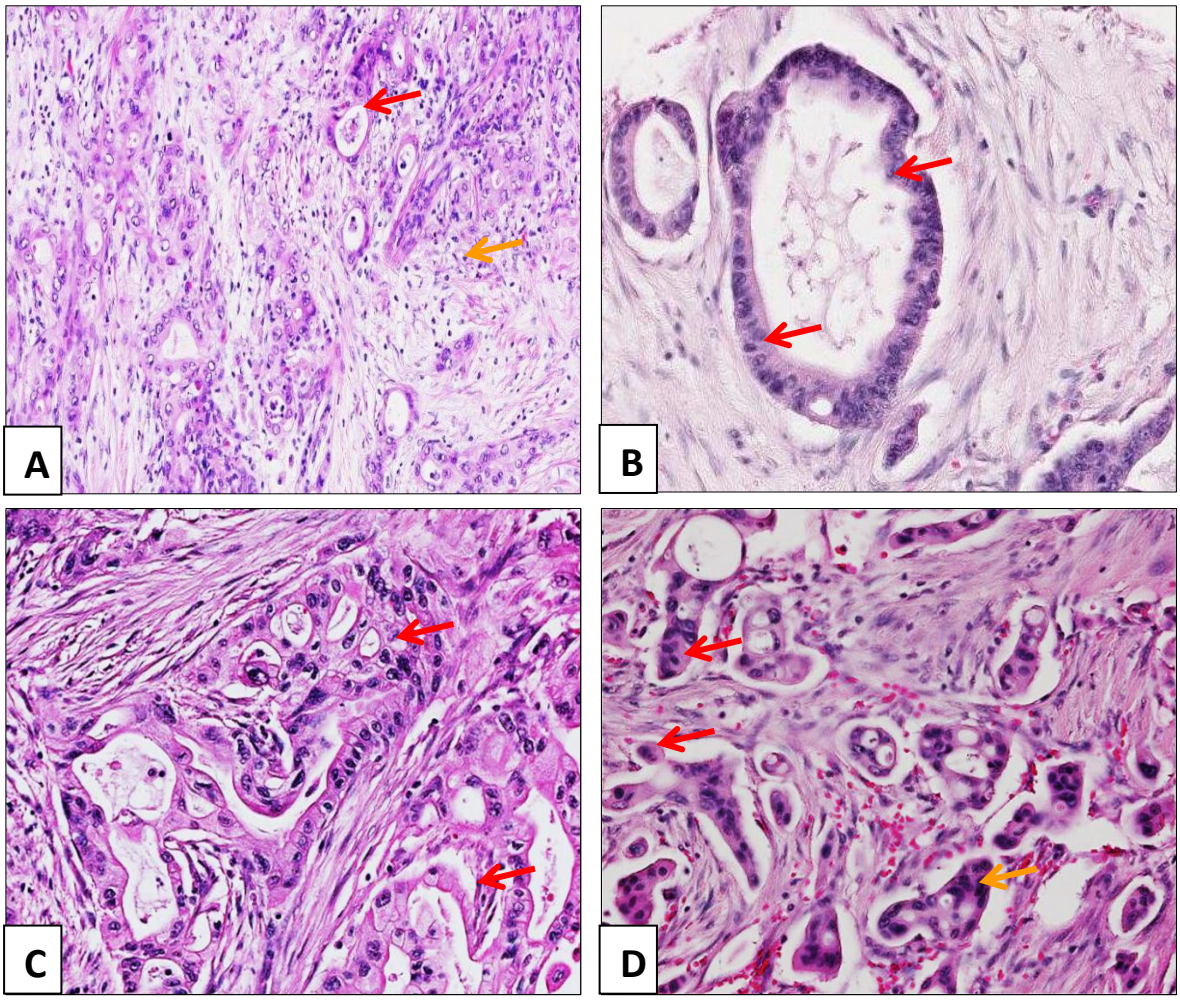



**Figure legend: A)** Pancreatic ductal adenocarcinoma is characterised by infiltrating tubular glands (red arrow) which are embedded in desmoplastic stroma (orange arrow), **B)** Well differentiated pancreatic ductal adenocarcinoma. Well-formed glandular structures lined by cuboidal cells similar to non-neoplastic ducts (red arrows). **C)** Moderately differentiated pancreatic ductal adenocarcinoma. The degree of cytologic and nuclear atypia is more and there is loss of polarity with irregular glandular arrangement (red arrows), **D)** Poorly differentiated pancreatic ductal adenocarcinoma. The degree of cytologic and nuclear atypia (orange arrow) is more with less glandular arrangement (red arrow).

**Figure 1.5: Benign pancreatic cytology**

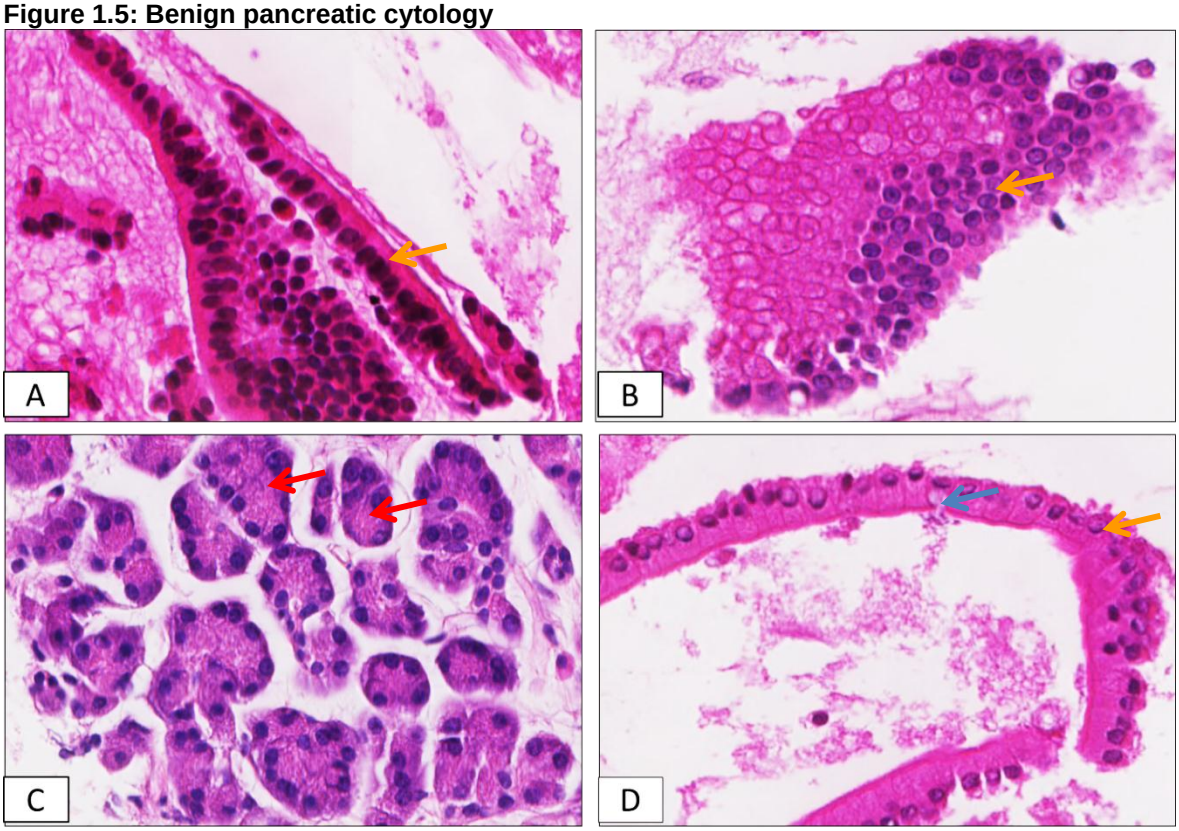



**Figure Legend: A&B)** Normal pancreatic ductal epithelial cells: flat sheet of cells with regular and evenly spaced nuclei (orange arrows), **C)** Normal pancreatic acinar epithelial cells: Group of cells forming acini (red arrows). Cells have epical eosinophilic cytoplasm and nuclei are oriented peripherally, **D)** Normal duodenal epithelial cells contamination of cytology samples: Columnar epithelium (orange arrow) with goblet cells (blue arrow) and brush border microvilli.

**Figure 1.6: Fine needle aspiration cytology samples of PDAC.**

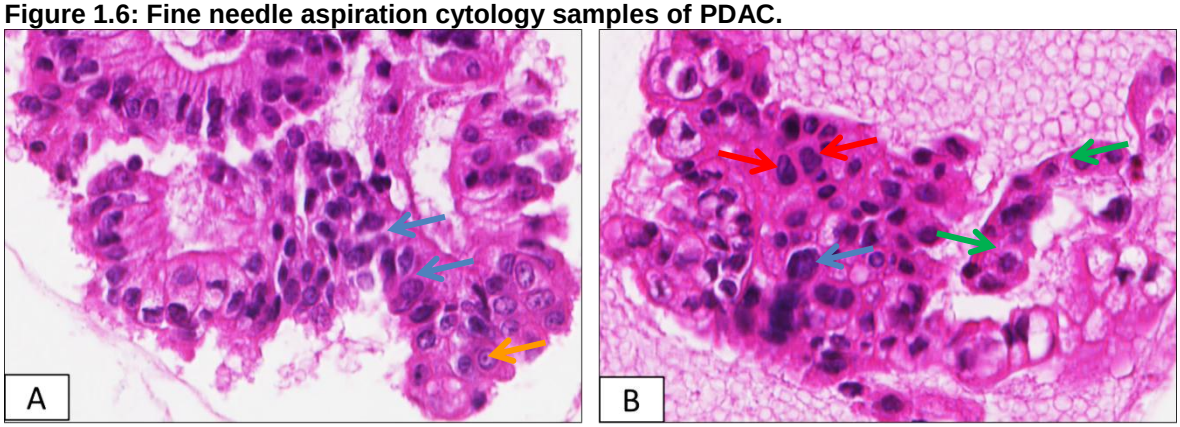



**Figure Legend: A&B)** Cluster of malignant cells (blue arrows), nuclear pleomorphism (red arrows), prominent nucleoli (orange arrow), loss of polarity (green arrow) and nuclear overlapping (heamatoxylin and eosin of cell blocks**), B)** drunken honeycomb appearance (irregular, enlarged, crowded ductal epithelial sheets).

**Figure 1.7: The pathway of biomarker development from discovery to clinic.**

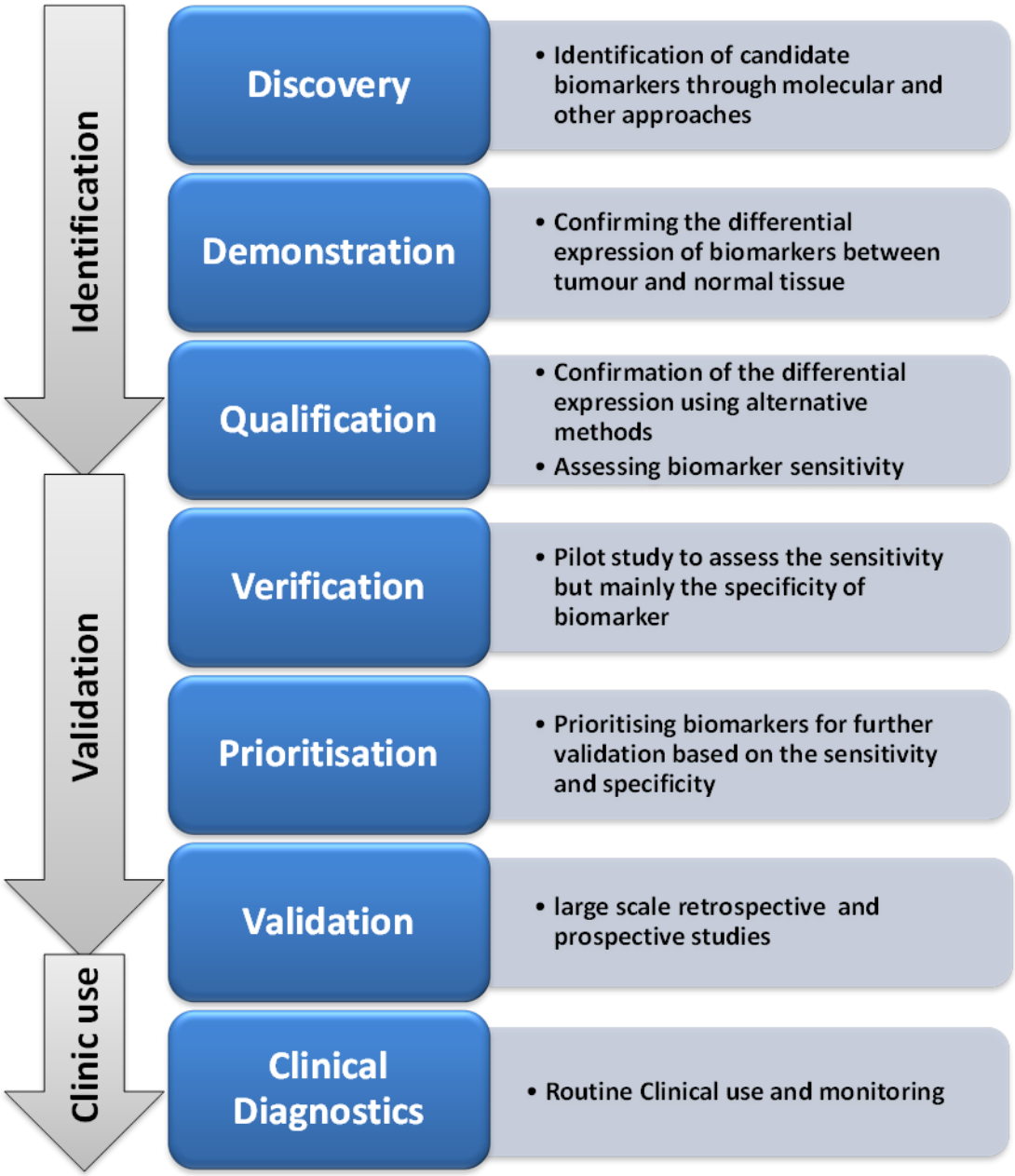



**Figure Legend:** The sequential stages from discovery to clinical diagnostics are shown in the middle vertical block. Furthermore, each stage is further elaborated in the right side vertical block. Adapted from a model based on Lee et al 2007 (119).

**Figure 1.8: Diagrammatic representation of reasons delaying clinical translation of biomarkers**

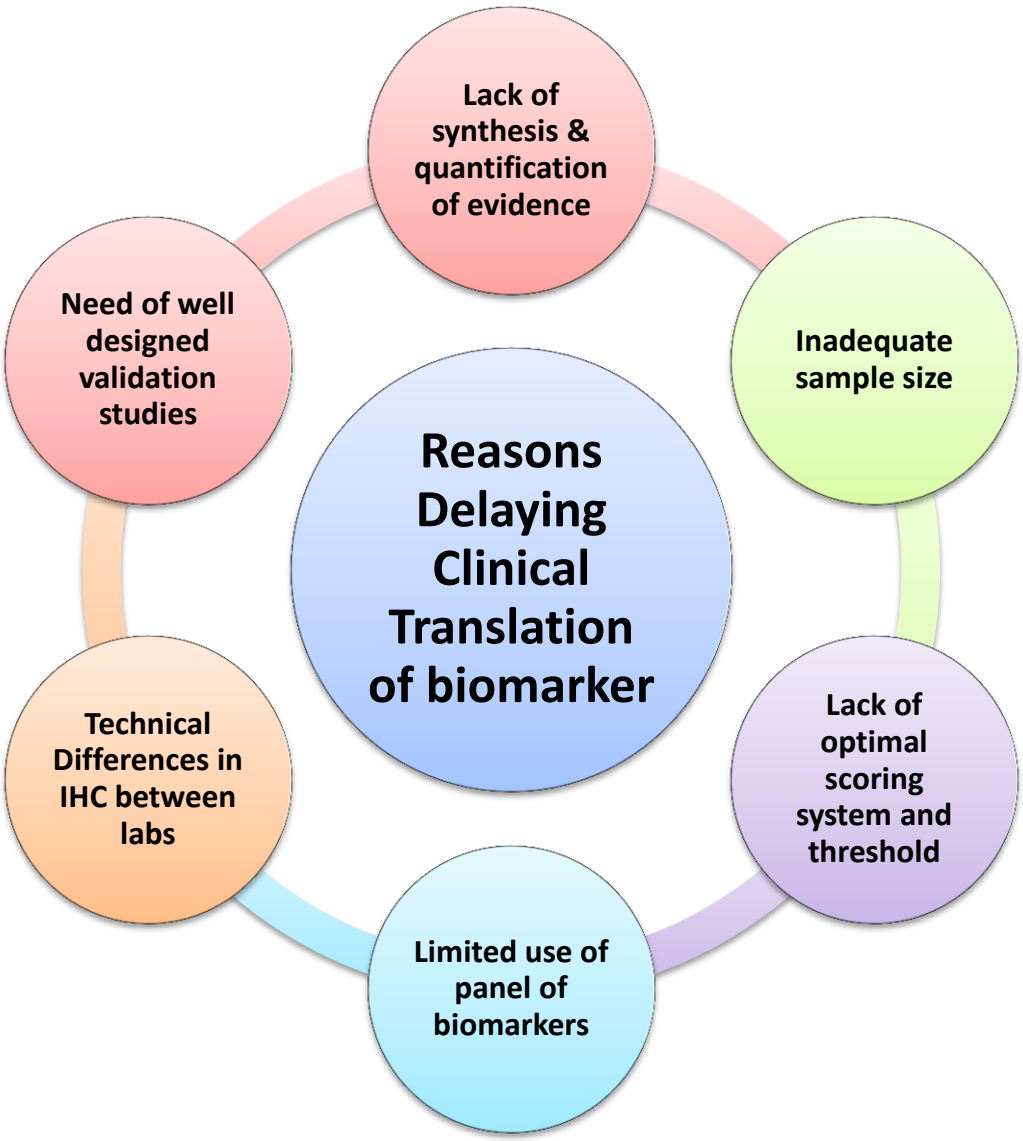


**Figure 1.9: Tissue microarrays (haematoxylin and eosin)**

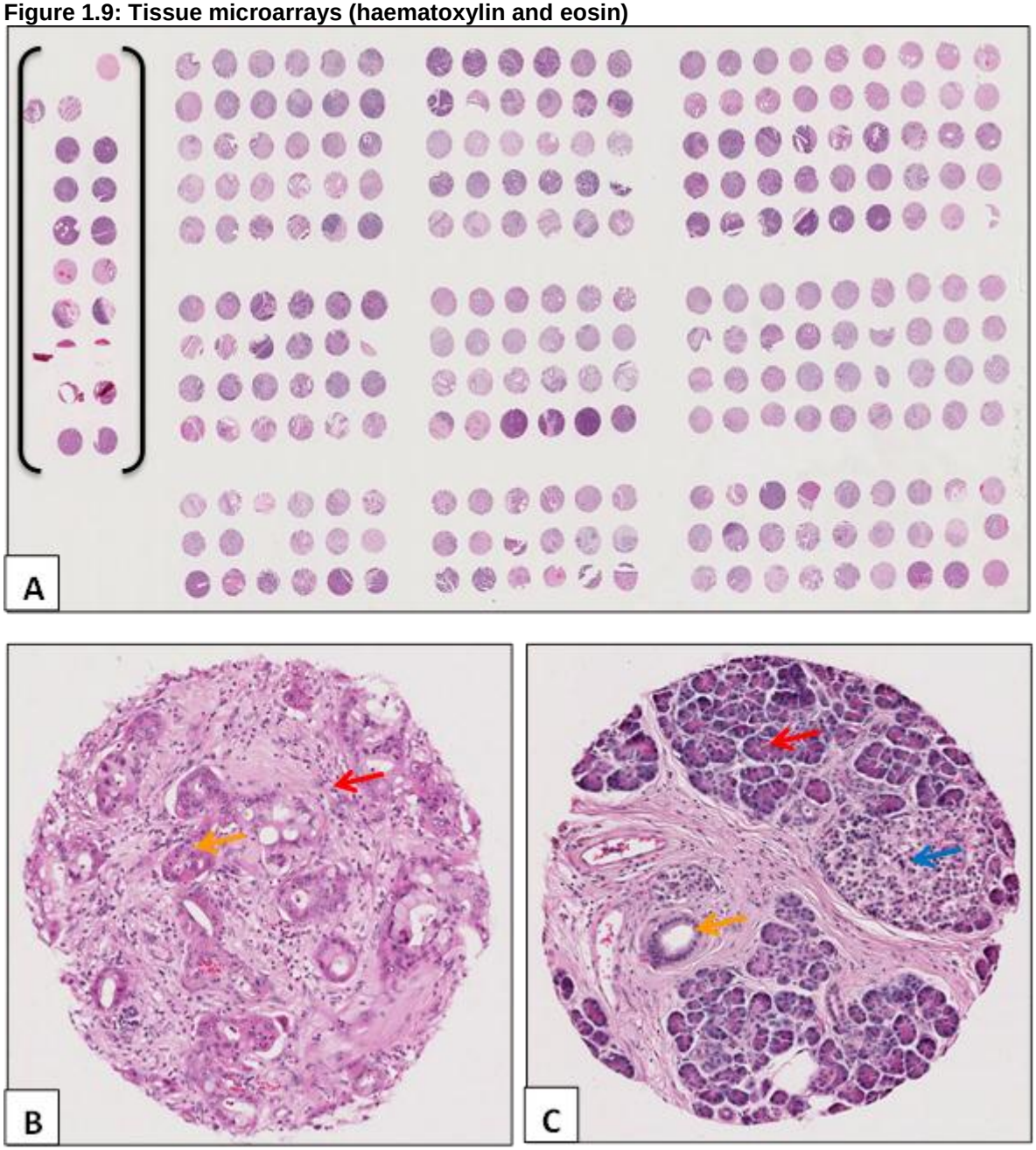
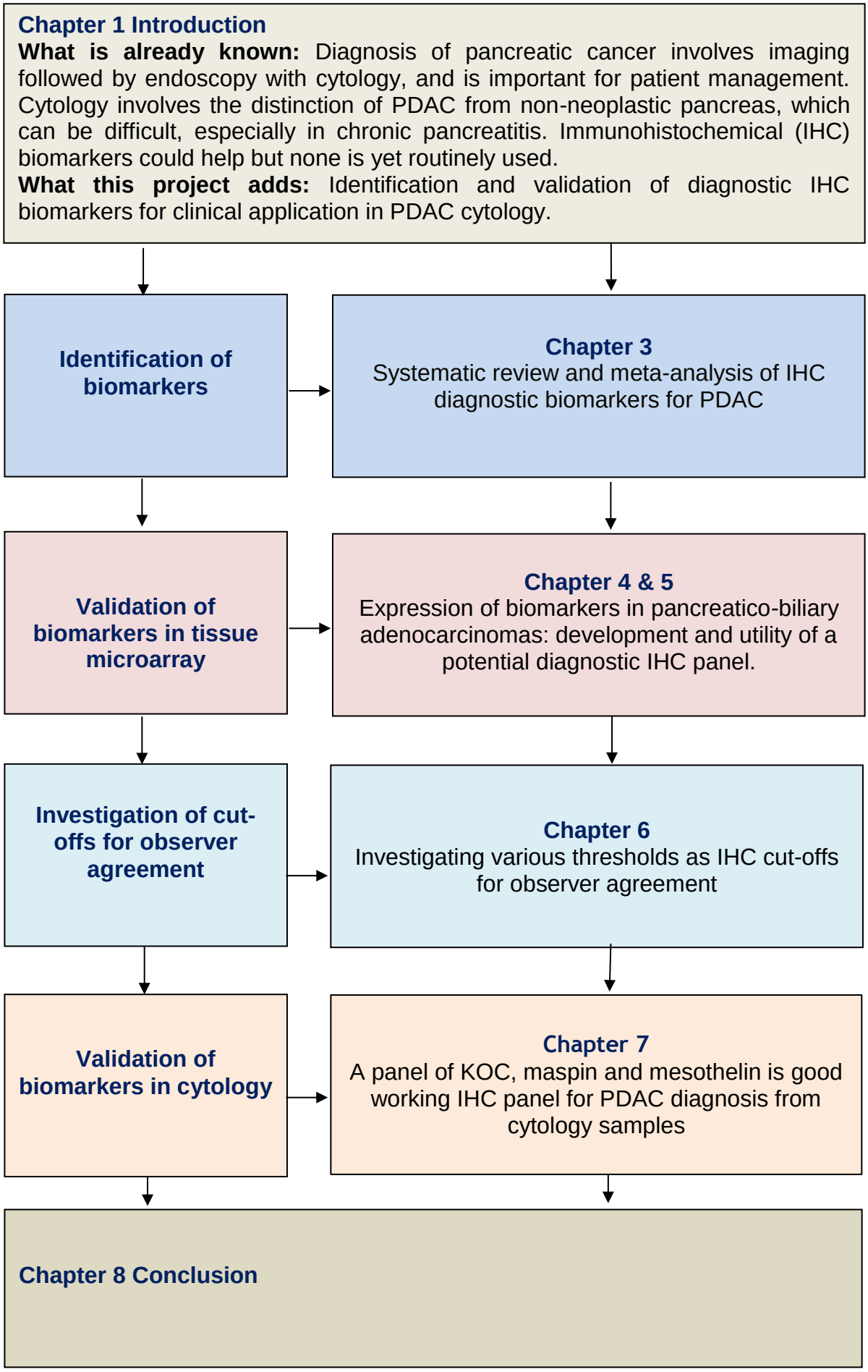



**Figure Legend: A)** an image of a tissue microarray (TMA) section containing both tumour (PDAC) and normal cores (normal pancreas). Tissue cores in the bracket are from other tumour types to help orientate the slide. **B)** A tumour core from patient with PDAC showing invading glands (orange arrow) and dense stroma (red arrow). **C)** A normal core from the adjacent normal pancreas from patient with PDAC showing normal duct (orange arrow), normal acini (red arrow) and Islet of Langerhans (blue arrow).

**Figure 1.10: Experimental flow chart of current PhD project.**



# 2 Materials and methods

# 2.1 Materials

Materials used for the current PhD project that helped to carry out four studies discussed in chapter 3, 4, 5 and 6 of this thesis are described below.

## 2.1.1 Systematic review and meta-analysis of immunohistochemistry (IHC) diagnostic biomarkers

### 2.1.1.1 Softwares and Databases

**The Cochrane Collaboration (Oxford, UK)**

The software package RevMan 5.1 for systematic review and meta-analysis

www.cochrane.org

**The R project for statistical computing**

R 2.15 with MADA package (Meta-Analysis of Diagnostic Accuracy)

www.r-project.org/

**Meta-DiSc: software for meta-analysis of test accuracy data**

MetaDisc version 1.4

www.hrc.es/investigacion/metadisc_en.htm

**Ovid database (New York, USA)**

Ovid MEDLINE®, Coverage: 1948 - Present

Embase, Publisher: Elsevier, Coverage: 1974 - Present

www.ovid.com

**National Center for Biotechnology Information (NCBI)**

www.ncbi.nlm.nih.gov/pubmed/

**Microsoft Office (Reading, UK)**

Excel spreadsheet programs

## 2.1.2 Statistical analysis of diagnostic biomarkers

### 2.1.2.1 Softwares

**IBM Corp, SPSS. (New York, USA)**

SPSS version 21.0

**GraphPad Software, Inc. (California, USA)**

Version 5

**Minitab (Coventry, UK)**

Version 16

**Microsoft Office (Reading, UK)**

Excel spreadsheet programs

### 2.1.3 Equipment

The equipment that is standard in most research laboratories were used for this project and is not separately listed by supplier, but included: refrigerators, −20$^{O}$C freezers, wet and dry ice, vortex mixers, microcentrifuge, microwave oven, pressure cooker, glass pipettes, bottles, staining troughs, universal container, measuring cylinders and flasks.

Other instruments used are listed below.

**Anachem Ltd (Luton, UK)**

Pipettes (2µl, 20µl, 200µl, 1000µl)

**DAKO (Ely, UK)**

DAKO Autostainer Link 48

DAKO Pre-Treatment (PT) module

**Leica Biosystems (Milton Keynes, UK)**

Multistainer Leica ST5020

**Hamamatsu (Japan)**

NanoZoomer Digital Pathology Scanner

**Millipore (Watford, UK)**

Milli-Q plus PF water purification system

## 2.1.4 General plasticware and needles

**Elkay laboratory Products (UK) Ltd (Basingstoke, UK)**

Microcentrifuge tubes (0.5, 1.5 and 2.0ml, both flip-cap and screw cap)

Standard pipette tips (10µl, 20µl, 200µl and 1000µl)

**Bectin Dickinson (Oxford, UK)**

Disposable Needles (22 gauge)

Disposable Syringes (5 mls)

## 2.1.5 Chemicals, buffers and solutions

Stock solutions were made up with distilled water.

**Hayman Ltd (James Borrough) (Witham, UK)**

Absolute alcohol (ethanol, analytical reagent grade, 100%, 70% solutions)

**Cell Path**

Haematoxylin Z

Gill's Haematoxylin

Putts Eosin

**Sigma-Aldrich (Dorset, UK)**

Scott's Tap Water substitute

Leica Biosystems (Milton Keynes, UK)

Acid Alcohol

DPX mounting medium

Sodium Citrate Buffer (0.1M Sodium citrate)

Ethylenediaminetetraacetic acid disodium salt dihydrate (EDTA) buffer (1mM EDTA (pH 8.0))

Tris EDTA buffer (1mM EDTA (pH 9.0))

TBS/Tween

2.5ml Tween

Xylene

Thrombin

**Hologic (West Sussex, UK)**

PreservCyt solutions

## 2.1.6 Tissue microarrays

Two sets of tissue micorarrays (TMAs) were used for investigating the expression of biomarkers. One set contain 3 TMAs and the other set contain 7 TMAs. In addition, a test array called 'PDAC practice TMA' was used for optimisation of antibodies for IHC. A detailed account of these TMAs is outlined in the Methods section 2.2.2.

## 2.1.7 Cytology

Two sets of cytology samples were used for investigation of biomarkers. One set was called 'needling' samples and was used for optimisation of antibodies and as control samples. The other set was archival cytology samples and was used to evaluate the diagnostic utility of biomarkers. A detailed account of these cytology samples is outlined in the Methods section 2.2.3.

**Sectioning of sections-Leica Microsystems (UK) Ltd (Milton Keynes, UK)**

1512 Rotary microtome

## 2.1.8 Immunohistochemistry

### 2.1.8.1 Primary antibodies

**DAKO Ltd (Ely, UK)**

KOC/IMP3 (Monoclonal Mouse Anti-Human; L523S, Clone 69.1)

**Leica microsystems (Novocastra antibodies) (Milton Keynes, UK)**

Mesothelin (Monoclonal Mouse Anti-Human; Clone 5B2)

MUC1 (Monoclonal Mouse Anti-Human; Ma695)

**BD Bioscience (Oxford, UK)**

S100P (Mouse monoclonal Anti-Human; clone 16)

Maspin (Mouse monoclonal Anti-Human; clone G167-70)

**Other reagents from DAKO (Ely, UK) for immunohistochemistry**

EnVision™ Detection Systems

Peroxidase block

Diaminobenzidine (DAB+) chromogenic substrate system
Secondary anti-mouse antibody

### 2.1.9 Tissue Microarray Scoring

#### 2.1.9.1 Tissue microarray image acquisition and archiving

**Slidepath (Dublin, Ireland)**

Digital Image Hub and Distiller version 2.2 are software that were used for archiving tissue microarray images for scoring purposes, capturing images and labelling of samples.

## 2.2 Methods

The methods are divided into six sections to facilitate the understanding of methodology of subsequent results chapters. The subsections are: Systematic review and meta-analysis of IHC diagnostic biomarkers; tissue microarrays resource; cytology resource; IHC optimisation and staining of sections; microscopic assessment and scoring of IHC; and observer variations in the assessment of cut-offs for IHC biomarkers.

### 2.2.1 Systematic review and meta-analysis of IHC diagnostic biomarkers

#### 2.2.1.1 Search strategy and identification of papers

The aim of the search was to retrieve papers describing tissue based IHC biomarkers having diagnostic potential and showing differential expression between PDAC and non-neoplastic pancreas.

The literature was searched using EMBASE and MEDLINE databases from inception to March 2012. Table 2.1 shows the search terms and MeSH headings used for retrieving relevant literature. The reference lists of papers judged relevant were searched to ensure the identification of additional papers missed by our primary search. The papers identified were then screened for systematic review and

meta-analysis. The papers selected for meta-analysis were also checked for methodological quality. A detailed account of this selection process and quality of papers is outlined in chapter 3.

#### 2.2.1.2 Meta-analysis of biomarkers

The meta-analysis aimed to generate a list of diagnostic biomarkers for PDAC, assessed in either surgical or cytology specimens, ranked by sensitivity and specificity.

To achieve this, first, coupled Forest plots were generated for biomarkers separately in resection specimens and in cytology samples. These Forest plots display within study estimates and confidence intervals for sensitivity and specificity for each biomarker.

Second, bivariate summary estimates for sensitivity and specificity were generated for each biomarker. For a biomarker described in at least 3 studies, we used a bivariate random effects model (REM) to get combined summary estimates for sensitivity and specificity (218); for a biomarker described in only two studies, we used a bivariate fixed effect model (FEM) (219). (By definition, no biomarkers described in only one study progressed to meta-analysis.)

Third, combined summary receiver operating characteristic (ROC) curves were generated, as follows. For each biomarker, the sensitivity of each study was plotted against (1 minus specificity) to yield an individual point in a space plot, which is the basis of a ROC curve. For each biomarker, the individual points for each of the multiple studies were placed in one plot, and then a summary ROC curve (SROC) was drawn. This provides a meta-analytical summary for a given biomarker and displays any variation between studies. Then, for all biomarkers assessed in one specimen type, all SROC curves were drawn in a so-called combined SROC plot and compared. This combined SROC plot, together with the bivariate summary estimates, provides a ranked list of biomarkers ordered according to pooled sensitivity and specificity.

The software package RevMan 5.1 (from The Cochrane Collaboration) (220) was used to generate coupled Forest plots and combined summary ROC curves. R

2.15 with MADA package (Meta-Analysis of Diagnostic Accuracy) (221, 222) was used for bivariate REM; and MetaDisc version 1.4 (223) was used for bivariate FEMs.

## 2.2.2 Tissue microarray resource

Two sets of tissue microarray (TMA) sections were used for analysing the expression of diagnostic biomarkers investigated in this thesis. One set of TMAs from surgical specimens was constructed by the researcher Simon Denley (SD) (a full list of contributors identified by initials appears at the beginning of the thesis) and the use of these TMAs for biomarker studies has been previously described (224). The second set of TMAs from surgical specimens was constructed by researcher Nigel Balfour Jamieson (NBJ) and the use of these TMAs for biomarker studies has been previously described (225). For the purpose of this thesis the former set was called 'SD TMAs' and the latter as 'NBJ TMAs'. These two sets of TMAs were used because the TMAs from 'SD TMAs' had a loss of TMA cores and hence the staining for one biomarker was performed on TMAs constructed from 'NBJ TMAs'. Both sets of TMAs are from the same surgical and pathology units in Glasgow and the details are presented below.

### 2.2.2.1 SD TMAs

Histological sections from three tissue microarrays (TMAs) containing samples from 99 surgically resected pancreatico-biliary adenocarcinoma (PBA) patients (Pancreatic ductal adenocarcinoma (PDAC) =85, Cholangiocarcinoma (CCC) = 14) were used for IHC. All resectional surgery was performed in the West of Scotland Pancreatic Unit (WSPU), Glasgow Royal Infirmary (GRI), UK, during a 10-year period (1995 to 2004). All patients were consecutive but tumours from other pancreatic pathologies for example intra-ductal papillary mucinous neoplasms and neuroendocrine tumours were excluded. Formalin fixed paraffin embedded (FFPE) tumour specimens were archived in the Department of Pathology, GRI and were used for the construction of TMAs. These TMAs contain five tissue cores (3 tumours and 2 normal) for each patient. Tumour cores are adenocarcinoma cores from patients with PBA, whereas normal cores are from adjacent normal ducts and acini. The male to female ratio was almost 1:1 and the median age of

patients was 65 (range, 38-77) years. These TMAs were used for IHC staining of KOC, S100P, mesothelin and MUC1.

#### 2.2.2.2 NBJ TMAs

Histological sections from seven TMAs containing samples from 137 surgically resected PBA patients (PDAC=114, CCC= 23) were used for IHC. All resectional surgery was performed in the WSPU, GRI, UK, during a 15-year period (1992 to 2007). All patients were consecutive but tumours from other pancreatic pathologies for example intra-ductal papillary mucinous neoplasms and neuroendocrine tumours were excluded. FFPE tumour specimens were archived in the Department of Pathology, GRI and were used for the construction of TMAs. The use of these TMAs has been previously described (225). These TMAs contain eight tissue cores (6 tumours and 2 normal) for each patient. Tumour cores are adenocarcinoma cores from patients with PBAs, whereas normal cores are from adjacent normal pancreatic ducts and acini. The male to female ratio was almost 1:1 and the median age of patients was 65 (range, 41-77) years. These TMAs were used for the staining of maspin.

#### 2.2.2.3 PDAC practice TMA

The PDAC practice TMA was used for optimisation of antibodies. This microarray contained 20 tissue cores from patients with PDAC.

***Sectioning of TMAs:*** The tissue microarray blocks were sectioned using Leitz 1512 microtome. A 2.5µm thick section was cut from each TMA block and put onto a "charged slide". These slides were then baked at 65˚C for one hour and stored at 4˚C. These TMA blank sections were prepared by CO (senior lab scientist) and were used for IHC of biomarkers.

#### 2.2.2.4 Ethical approval

Ethical approval has been granted by the North Glasgow University Hospitals NHS Trust Ethics Committee and by the National Health Service Greater Glasgow and Clyde Ethics Committee. This ethics approval includes the use of archival pathology specimens, where the patients were not given the opportunity to

donate their tissue. The ethical approval for the use of TMAs for research purposes for SD TMAs and NBJ TMAs was obtained by SD and NBJ respectively.

## 2.2.3 Cytology resource

Two sets of cytology samples were used for optimising and analysing the expression of diagnostic biomarkers investigated in this thesis. One set of cytology samples was called 'needling' samples generated from obtaining cells from fresh Whipple resections. The second set of cytology samples was retrieved from the cytology archives in the form of FFPE cell blocks. The former set was used for optimising the diagnostic antibodies and the latter set was used for evaluating the diagnostic sensitivity and specificity of biomarkers. The details of these cytology samples are presented below.

### 2.2.3.1 'Needling' cytology samples

The cytology samples were obtained by 'needling' the fresh Whipple specimens (tumour, adjacent normal pancreas and normal duodenum) using a sterile and disposable needle with syringe attached to it. These samples were therefore called 'needling' cytology samples. Sample collection was under the guidance of the on-duty pathologist (Fraser Duthie (FD), Alan K Foulis (AKF), Sarah Bell (SB) and Karin A Oien (KAO)) for the gastrointestinal team, with help as required from the author Asif Ai (AA).

The primary tumours were surgically resected at the WSPU, GRI, UK during a 5-month period (22 October 2013 to 17 March 2014). Immediately following surgical resection the tissue specimens were transported on ice by the bio-bank staff to the Department of Pathology, Southern General Hospital (SGH), Glasgow, UK. 'Needling' cytology samples were prospectively collected from ten Whipple resections. Samples obtained were processed into FFPE 'cell blocks' and archived in the Department of Pathology, SGH. The male to female ratio was 1.5:1 and the median age of patients was 70 (range, 50-80) years. The method used to generate this resource is outlined below.

### 2.2.3.2 Generating 'needling' cytology resource

The technique of obtaining the cell samples and processing into cell block was performed as close as possible to the fine needle aspiration in a routine clinical practice. A needle of similar gauge (22 gauge) to that used in fine needle aspiration of the pancreas was used and on average 8-10 needle passes with and without aspiration via a syringe was performed for each 'needling' sample. Such samples were obtained from three different sites i.e. pancreatic tumour, normal adjacent pancreas and normal duodenum. The 'needling' specimens obtained were then put in three different and labelled universal containers containing the "PreservCyt" solution. These samples obtained were then quickly transferred to the Department of Cytology, SGH, Glasgow, UK.

These samples were then processed into FFPE cell blocks by a team of cytology colleague (Lisa Irvine (LI), John McCorriston (JM)) who are also involved with processing the pancreatic and other cytology samples in routine clinical practice. The processing of 'needling' samples into cell blocks was performed in a way exactly similar to the routine pancreatic cytology samples. Briefly, saline was added to the top of universal container with fragments of cell/tissue, centrifuged, decanted and three such saline washes were performed. Afterwards, two drops of human plasma and then two drops of thrombin solution were added to the cell pellet with the help of a disposable pipette. The cell pellet was left until a cell clot was formed. The cell clot was then gently tipped and put in universal container filled with formalin. The fixation in formalin was carried out for 12 hours and the specimen was then embedded in paraffin to form a "cell block". This procedure is similar to that used for routine diagnostic cytology samples, thus the resulting samples should be comparable to standard cytology and enable their use for optimisation and as control samples for IHC or other research techniques. The methodology of obtaining these 'needling' samples is shown in Figure 2.1.

The 'needling' samples were obtained from ten Whipple resections i.e. six PDAC, one ampullary carcinoma (AC), one neuroendocrine tumour (NET), one intraductal papillary mucinous neoplasm (IPMN) and one serous cystadenoma. The 'needling' samples were procured from tumour, adjacent normal pancreas and normal duodenum from PDAC, AC and IPMN cases. This makes a total of 27

'needling' samples processed into cell blocks and archived in the Department of Cytology, SGH, Glasgow, UK for this project and future research.

Furthermore, the corresponding resections specimens of PDAC 'needling' samples were also obtained from the Department of Pathology. The H&E sections of these cases were retrieved from the Department of Pathology, SGH, Glasgow, UK and were reviewed on a multi-header microscope by a pathologist (SB) and the author (AA). After careful inspection, from the H&E sections, tissue blocks with malignant tissue and adjacent normal pancreas were identified. These tissue sections will allow the comparison of biomarker expression in cytology and corresponding tissue sections.

The purpose of the 'needling' resource was to optimise diagnostic IHC antibodies before staining the archival cytology specimens and to use as control specimens. However, this resource is not limited to this project and could potentially be used for future biomarker research.

#### 2.2.3.3 Archival cytology samples

The archival cytology samples for the immunostaining were carefully selected. After discussion and with the help of gastro-intestinal pathologists (KAO, FD and SB), the cytology files of 57 patients were reviewed. The heamatoxylin and eosin (H&E) stained sections of cell blocks of 57 cytology samples obtained by EUS-FNA and ERCP brushing for suspected PDAC between May 2012 and October 2012 were identified from the cytology files of Department of Cytology, Southern General Hospital, Glasgow, UK.

The H&E sections of cytology samples were then checked for adequacy of cellularity on a multi-header microscope with the help of pathologists FD and SB. FD in particular is an expert on pancreatic cytology. The criteria for adequacy of cellularity from H&E stained sections was the presence of cytologically distinct sheets/groups of malignant and benign cells respectively in malignant and benign cytology samples. By applying this criterion to H&E sections, 36 cases with insufficient/scant cellularity were excluded from this study. The remaining 21 cases included in the study were: ten patients with a cytology diagnosis of adenocarcinoma, one patient suspicious for malignancy and ten patients with

benign cytology diagnosis. The male to female ratio was almost 1:1 and the median age of patients was 69 (range: 46-86) years.

***Sectioning of cell blocks:*** FFPE cell blocks from both 'needling' and archival resource were sectioned using Leitz 1512 microtome. A 4µm thick section was cut from each cell block and put onto a "charged slide". These slides were then baked at 65˚C for one hour and stored at 4˚C. These cytology sections were prepared by RF (senior lab scientist) and were used for IHC of biomarkers.

#### 2.2.3.4 Ethical approval

The author (AA) gained ethical approval for both 'needling' cytology and archival cytology samples from National Health Services, Greater Glasgow and Cylde (NHS GG&C), Bio-repository Ethics Committee. Three ethical applications were sequentially written by the author (AA): one application was for the 'needling' cytology project; the second application was an amendment of the 'needling' cytology project to allow for the procurement of corresponding surgical pathology specimens for comparing biomarker expression in tissue vs. cell samples; and the third application was for staining the archival cytology samples for the final and proposed biomarker panel. The approvals of all three ethical applications are present in the Appendix.

### 2.2.4 IHC optimisation and staining of sections

This section is divided into three subsections as follows: the IHC methodology; the optimisation of IHC on tissue sections; and the optimisation of IHC on 'needling' samples.

#### 2.2.4.1 Immunohistochemistry methodology

The technique of IHC was the same for both TMA sections and cytology sections and is presented below. IHC was performed using an automated platform (DAKO Autostainer Link48) for KOC, maspin, mesothelin, MUC1 and S100P antibodies. The steps are detailed below.

**De-waxing and rehydration:** The purpose is to remove the wax and hydrate the tissue to allow the aqueous solutions to penetrate the sections. FFPE sections

were de-waxed in xylene for 5 minutes and then rehydrated for 1 minute each in graded alcohols, 100%, 100% and 70%. Sections were then washed in deionised water for 5 minutes.

**Antigen Retrieval:** Formalin fixation leads to cross-linking within tissue which might mask the antigenic site of the protein. Antigen retrieval is therefore necessary to unmask the antigen for the antigen-antibody reaction. In this thesis two methods were used for antigen retrieval i.e. heat induced epitope retrieval (HIER) and enzyme digestion. HIER was performed on sections requiring KOC, maspin, mesothelin and MUC1 antibody staining using DAKO's Pre-Treatment (PT) module. Sections were incubated in different retrieval buffer solutions depending on the antibody and tissue or cell section used at 98ºC for 25 minutes. Enzymatic digestion using proteinase-k for 10 minutes at room temperature was used for the antigen retrieval of S100P antibody. Sections were then washed with Tris Buffered Tween (TbT).

**Blocking the endogenous peroxidase activity:** Endogenous Peroxidase was blocked using DAKO EnVision Peroxidase block for 5 minutes. Sections were then washed with TbT.

**Antibody incubation:** Primary antibody was applied for 60 minutes with final antibody dilutions after optimisation. After incubation, the sections were washed with TbT. Secondary antibody (DAKO EnVision, anti-mouse) was then applied to the sections for 40 minutes. Sections were then washed with TbT.

**Visualisation method:** Visualisation of the antigen-antibody reaction is required to identify the location and intensity of staining. For all antibodies 3,3'-Diaminobenzidine tetrahydrochloride (DAB) was applied for 10 minutes to achieve brown staining in subcellular compartments (cytoplasm, cell membrane and/or nucleus). The reaction was then terminated by placing the sections into deionised water.

**Counterstaining:** This is required to stain areas in the sections without antibody to allow visualisation of other cellular and stromal elements. For this purpose Haematoxylin Z was applied for 7 minutes to stain nuclei. Sections were then washed in tap water for 1 minute, dipped once in 1% acid alcohol for 2 seconds,

washed again in tap water for 30 seconds, placed in Scotts tap water substitute solution for 2 minutes and finally washed in tap water for 1 minute.

**Dehydration and mounting:** After staining the slides are dehydrated and put in xylene before mounting to preserve the tissue for microscopic examination. The sections were dehydrated for 1 minute each in graded alcohols 70%, 100% and 100% and then put in 3 changes of xylene for 5 minutes each. DPX mountant was used to attach a glass coverslip to the slide.

The counterstaining and cover-slipping was performed on a Lecia Autostainer ST5020). Slides were allowed to dry for 30 minutes and examined under the microscope.

#### 2.2.4.2 Optimisation of IHC on tissue sections

All antibodies were optimised before staining the final microarrays. The plan of optimisation was to review the literature (Table 2.2) and identify preferred IHC parameters reported for each antibody. This provided a starting point for optimisation with some modifications. The IHC conditions with variable parameters as shown in Table 2.3 were then used to stain the FFPE tissue sections from patients with PDAC. The IHC conditions achieving optimal staining were then repeated on PDAC practice TMA containing 20 tissue cores. The optimal staining was defined as diffuse and strong intensity staining in malignant ducts, very low or no staining in normal tissue and lowest possible background staining.

Four biomarkers KOC, mesothelin, MUC1 and S100P were used to stain the SD TMAs. One biomarker maspin was used to stain the NBJ TMAs due to the loss of cores in SD TMAs in the later sections used for the current and other research projects. The Optimisation of IHC and the final staining of SD TMAs for KOC, mesothelin and MUC1 were performed by a laboratory scientist VB under the supervision of KAO. The Optimisation of IHC and the final staining of SD TMAs for S100P and NBJ TMAs for maspin were performed by the author (AA) and CN. Nevertheless, scoring of all immunostaining was independently performed by the author (AA). The optimisation is outlined as follows.

**SD TMAs were used for staining of KOC, mesothelin, MUC1 and S100P:** For **KOC**, the HIER (pH6, citrate buffer) was evaluated against three primary antibody dilutions 1/50, 1/100 and 1/200. Optimal staining on tissue sections was achieved using HIER (pH6, citrate buffer) and a 1/50 antibody dilution. This optimised protocol was then repeated on a practice TMAs again achieving optimal staining. Finally the three SD TMAs were stained with appropriate positive and negative control sections. The staining was expected in the cytoplasm based on the literature and product datasheet and this pattern was achieved in this study.

For **mesothelin**, the HIER (pH6, citrate buffer) was evaluated against three primary antibody dilutions 1/20, 1/40 and 1/80. Optimal staining on tissue sections was achieved using HIER (pH6, citrate buffer) and a 1/20 antibody dilution. This optimised protocol was then repeated on a practice TMA again achieving optimal staining. Finally the three SD TMAs were stained with appropriate positive and negative control sections. The staining was expected to be both cytoplasmic and membranous based on the literature and product datasheet and this pattern was achieved in this study.

For **MUC1**, the HIER (pH6, citrate buffer) was evaluated against three primary antibody dilutions 1/50, 1/100 and 1/200. Optimal staining on tissue sections was achieved using HIER (pH6, citrate buffer) and a 1/200 antibody dilution. This optimised protocol was then repeated on a practice TMA again achieving optimal staining. Finally the three SD TMAs were stained with appropriate positive and negative control sections. The staining was expected to be both cytoplasmic and membranous based on the literature and product datasheet and this pattern was achieved in this study.

For **S100P**, the enzymatic antigen retrieval (Proteinase k for 10 minutes) was evaluated against three primary antibody dilutions 1/50, 1/100 and 1/200. Optimal staining on tissue sections was achieved for proteinase k and a 1/100 antibody dilution (Figure 2.2). This optimised protocol was then repeated on a practice TMA again achieving optimal staining. Finally the three SD TMAs were stained with appropriate positive and negative control sections. The staining was expected to be both cytoplasmic and nuclear based on the literature and product datasheet and this pattern was achieved in this study.

**NBJ TMAs were used for staining of maspin:** The HIER (PH6, citrate buffer) was evaluated against two primary antibody dilutions 1/75 and 1/150. Optimal staining on tissue sections was achieved using HIER (PH6, citrate buffer) and a 1/75 antibody dilution (Figure 2.3). This optimised protocol was then repeated on a practice TMA again achieving optimal staining. Finally the seven NBJ TMAs were stained with appropriate positive and negative control sections. The staining was expected to be both cytoplasmic and nuclear based on the literature and product datasheet and the same was achieved in this study.

The final IHC conditions to stain the TMAs are outlined in Table 2.4 and they were called 'IHC parameters optimised for TMAs'.

#### 2.2.4.3 Optimisation of IHC on 'needling' samples

All antibodies were optimised before staining the archival cytology samples. The 'needling' samples (tumour, normal pancreas and normal duodenum) from five patients with PDAC were used for optimisation purposes. The optimisation was carried out in two phases as follows.

**First Phase:** In the first phase, the 'IHC parameters optimised for TMA' (Table 2.4) were used to stain the 'needling' cytology samples. The staining achieved for KOC, mesothelin and maspin on PDAC 'needling' samples was considered sub-optimal (detailed immunostaining and scoring results will follow in Chapter 7). For S100P no staining was achieved in all samples stained.

**Second Phase:** In the second phase, the IHC conditions were changed to achieve optimal staining. These optimised IHC parameters were called 'IHC parameters optimised for cytology'. The IHC parameters including antigen retrieval and antibody dilutions were adjusted to achieve optimal staining and the optimisation parameters are shown in Table 2.5. The following is an outline of the optimisation of KOC, mesothelin, maspin and S100P for cytology.

For **KOC** optimisation for cytology, the antigen retrieval buffers were compared i.e. HIER (pH6, citrate buffer) with HIER (pH9, Tris EDTA buffer). The rest of the IHC conditions exactly matched the 'IHC parameters optimised for TMAs'. Optimal staining was achieved with HIER (pH9, Tris EDTA buffer). This optimised

protocol was then repeated on one 'needling' sample achieving similar staining. Finally the 21 archival cytology samples were stained with appropriate positive and negative 'needling' control sections. The staining was expected in cytoplasm based on the TMA work in this thesis, literature and product datasheet and the same was achieved in this study in archival cytology samples.

For **mesothelin** optimisation for cytology, again the antigen retrieval buffers were compared i.e. HIER (pH6, citrate buffer) with HIER (pH8, EDTA buffer). The rest of the IHC conditions exactly matched the 'IHC parameters optimised for TMAs'. Optimal staining was achieved with HIER (pH8, EDTA buffer). This optimised protocol was then repeated on one 'needling sample' achieving similar staining. Finally the 21 archival cytology samples were stained with appropriate positive and negative 'needling' control sections. The staining was expected in both the cytoplasm and in the membrane based on the TMA work in this thesis, literature and product datasheet and the same was achieved in this study in archival cytology samples.

For **S100P** optimisation, three different antigen retrieval conditions i.e. proteinase k for 10 minutes, HIER (pH 6.0, citrate buffer), and HIER (pH 9.0, Tris EDTA buffer) were compared. In addition, two antibody dilutions 1/50 and 1/100 were used for each of the three antigen retrieval conditions. The rest of the IHC conditions exactly matched the 'IHC parameters optimised for TMAs'. However, no staining was observed in any samples stained with S100P antibody.

For **maspin** optimisation, the antibody dilution was changed i.e. 1/75 dilution was compared with 1/150 dilution. The rest of the IHC conditions exactly matched the 'IHC parameters optimised for TMAs'. Optimal staining was achieved with 1/150 antibody dilution. This optimised protocol was then repeated on one 'needling' sample achieving similar staining. Finally the 21 archival cytology samples were stained with appropriate positive and negative 'needling' control sections. The staining was expected in both the cytoplasm and in the nucleus based on the TMA work in this thesis, literature and product datasheet and the same was achieved in this study in archival cytology samples.

**Staining of the corresponding tissue sections:** The corresponding resection specimens of the 'needling' cytology samples for all antibodies including S100P

showed optimal and expected staining in both the first and second phase of optimisation.

The detailed immunostaining and scoring results of optimisation will follow in Chapter 7. The final IHC conditions ('IHC parameters optimised for cytology') to stain the archival cytology samples are outlined in Table 2.6.

## 2.2.5 Microscopic assessment and scoring of IHC

The H&E and IHC staining of TMAs and cytology samples were microscopically assessed. This section is divided into four subsections: TMA scoring; cytology scoring; evaluation of immunostaining; and statistics and data analysis.

### 2.2.5.1 TMA scoring

Microscopic analysis was undertaken blinded to the diagnosis or other parameters. The expression of KOC, mesothelin, MUC1, S100P and maspin across all TMAs was independently scored by the author (A.A). The scoring of tumour and normal tissue cores was performed based on the morphology of tissue and, with help, as required from the corresponding H&E stained core. A consensus was reached on cores found difficult to score with a consultant histopathologist (KAO).

Stained TMA sections were scanned (Hamamatsu Slide Scanner) and images uploaded in Distiller 2.2 (Leica Biosystems). This allowed the Distiller 2.2 to function as a digital microscope to visualize and score the TMAs. The OpTMA module in the Distiller 2.2 allows the user to view, navigate, annotate, zoom and score digital images. Digital images were opened in OpTMA and each tissue core in the TMA was presented with a scoring sheet allowing manual scoring using set criteria (Table 2.7- Scoring sheet). The scores for tissue cores were automatically saved after navigating to the next image in the folder. In addition, the scores from each tissue core were linked to a unique patient ID to allow matching scoring from different cores to a patient. All scores were exported in an Excel spreadsheet from Distiller 2.2 for analysis.

**Digital pathology role in TMA scoring**

The sampling strategy of TMA technology provides a high degree of sophistication by arraying small and regularly sized and shaped tissue cores (225). The staining of TMAs can thus be easily scored using both traditional microscope and digital pathology software that act like a digital microscope.

Traditional microscope based scoring of TMAs is sometimes challenging due to large number of densely packed cores on one slide (Figure 1.9). This can also result in a tiresome user interaction while developing consensus on difficult cores (unless it is a multi-header microscope). In comparison, digital pathology provides an efficient platform for scoring of TMAs which facilitates easy orientation of TMA slide for scoring and excellent user interaction on a personal computer (236).

Using a digital microscope, tissue cores on TMAs can be scored in a blinded fashion without knowledge of the nature and clinical information of the core. The scoring is automatically saved as one move along a TMA and if interrupted scoring can be resumed at anytime from anywhere with internet access and a computer. The scoring of TMA tissue cores is linked to patients' clinical data which can be exported in an excel spread-sheet. The TMA scoring and clinical data are arranged on a per core and per patient basis for subsequent analysis. These digital microscopes also allow zooming in and out of images and provide high quality annotated images of publication quality.

An additional advantage of some of these digital microscopes (for example Distiller and Digital Image Hub, Leica Microsystem) is that they provide dynamic images (with flexibility of zooming in and out) that can be emailed to a colleague for opinion. The scoring of multiple biomarkers on the same TMAs scanned into digital microscope facilitates assessing different panels of biomarkers.

Therefore, in a research setting digital pathology provide attractive platforms for scoring of TMAs and analysis of biomarker expression. This greatly facilitates and speeds up the analysis of the clinical utility of biomarkers.

### 2.2.5.2 Cytology scoring

The assessment and interpretation of cytology samples and the resulting IHC staining was undertaken in conjunction with pathologists (F.D and SB). One pathologist (FD) in particular has significant expertise in pancreatico-biliary cytology. All scoring was performed using a light microscope.

The purpose of scoring of 'needling' cytology samples was to compare the level of staining for different IHC parameters. The scoring performed by the pathologist was such that in a 'needling' sample with a mixed population of cells only atypical and malignant cells were scored for IHC staining. This was carried out with a corresponding H&E stained section available for each IHC stained section.

The scoring of archival cytology samples was performed by the pathologist (SB) and the author (AA). Scoring was performed independently and the scorers were blinded to the final diagnosis. Moreover, the scorers were asked to score the IHC staining without paying attention to the cellular morphology. However, if there were very distinct populations of normal and abnormal cells, it was scored twice for both normal and malignant cells.

### 2.2.5.3 Evaluation of immunostaining

The biomarker expression in both tumour and normal tissue was recorded utilising a semi-quantitative Histoscore. The Histoscore takes into account both the staining intensity (0=absent, 1=weak, 2=moderate and 3=strong) and proportion of positive cells with each of these staining intensities. A Histoscore is calculated by the formula [0 × % negative cells + 1 × % weakly stained cells + 2 × % moderately stained cells + 3 × % strongly stained cells]. The Histoscore thus has a range of possible scores between 0 and 300. As an example, a tissue core having 30% cells (of interest) with weak staining intensity; 40% cells with moderate staining intensity and 30% cells with strong staining intensity will have a total Histoscore of: (30x1)+(40x2)+(30x3) = 200. Histoscore provides a better estimate of biomarker expression than simply reporting the staining intensity or proportion of positive cells. A Histoscore was generated for each tissue core on TMAs and each cytology sample.

**TMAs:** A median Histoscore was taken as a final score for a patient with multiple tissue cores. As an example the final Histoscore of 180 was taken for diagnostic purposes for a patient with Histoscores of 150, 180 and 200 respectively for three tumour cores. Following IHC semiquantitative analysis of protein expression, three different scoring systems were used to assess expression level of biomarkers and sensitivity/specificity analyses. All three scoring systems were basically derived from the Histoscores as follows: scoring based on Histoscore value (range, 0-300); scoring based on percentage of positive cells (range, 0%-100%); scoring based on staining intensity (1=weak, 2=moderate, 3= strong).

**Cytology:** Again for scoring of the cytology samples the same three scoring systems were used as follows: scoring based on Histoscore value (range, 0-300); scoring based on percentage of positive cells (range, 0%-100%); scoring based on staining intensity (1=weak, 2=moderate, 3= strong). However, an additional scoring system based on categorical cut-off of 10% positive cells was also used. In this case, the scorers were asked to score a cytology case as 'positive' if more than or equal to 10% cells in the cytology sample were positive for a biomarker. While scoring a cytology case as 'negative' if less than 10% cells in the cytology sample are positive for a biomarker. Finally, the inter-observer agreement between two scorers (SB, a pathologist and AA, the author) for 10% cut-off was assessed using kappa statistics.

#### 2.2.5.4 Statistics and data analysis

The mean expression of each biomarker in the PBA tumour cores was compared with the mean expression in normal tissue cores. Statistical significance was calculated using an independent sample t-test to generate the p value. The independent sample t test was used rather than the paired sample t test because a full set of matching tumour and normal tissue cores was not available for approximately 5% of patients. This was due to loss of tissue cores during processing, which is expected in a proportion of samples. Sensitivity/specificity analyses were carried out for biomarkers, both individually and in panels of 2-4 biomarkers, and compared. We used two different panel approaches for sensitivity/specificity analyses. One approach assigns the case into the positive category if the tumour expresses only one biomarker in the panel. The other

approach assigns the case into the positive category if the tumour shows staining for at least 2 biomarkers in the panel.

A combined summary receiver operating characteristic (SROC) curve was generated to compare different panels of biomarkers. P value <0.05 was considered statistically significant. SPSS-21 and RevMan-5.1 were used for statistical analysis. Observer variations in the assessment of cut-offs for IHC biomarkers were also considered.

### 2.2.6 Investigating various thresholds as immunohistochemistry cut-offs for observer variations

The study was conducted to examine the inter-observer and intra-observer agreements between pathologists as an indicator of reliability and reproducibility of cut-offs (thresholds). The intra-observer part of the study examined individual pathologists' responses in a test-retest setting with retesting conducted three weeks after the inter-observer part and with all images re-arranged in different order to remove recall bias.

#### 2.2.6.1 PDAC images

A series of 36 images of pancreatic cancer tissue microarray cores for four diagnostic IHC biomarkers (nine images each from KOC, maspin, mesothelin and S100P) were used for this study. These cores have previously been studied and were retrospectively identified for the current study. These cores were carefully selected for each biomarker based on a variable range of staining intensity and proportion of positive cells. Some cores with no immunostaining were also included. The expression of KOC was cytoplasmic, maspin has both cytoplasmic and nuclear expression but the scorers were asked to score only cytoplasmic staining for maspin and disregard nuclear staining, mesothelin expression was cytoplasmic and/or membranous and, S100P expression was cytoplasmic and/or nuclear.

#### 2.2.6.2 Participants

Seven pathologists (3 experienced pathologists and 4 junior pathologists) participated in the study. Experienced pathologists have clinical pathology

experience of more than 15 years, while junior pathologists have 3-7 years of experience. Pathologists were coded as A, B, C, D, E, F and G.

#### 2.2.6.3 Scoring cut-offs

A PowerPoint presentation containing these 36 images was prepared by the author (A.A). The images were arranged based on biomarkers with reference staining intensities (weak, moderate, severe) provided at the start for each biomarker. A scoring sheet with instructions on scoring was prepared with the help of pathology colleagues (Table 2.8). All the participating pathologists participated in one session for the inter-observer part of the study.

After a short presentation (5-10 min) on the purpose of this study, the scoring sheets were distributed between all seven pathologists. Each image was shown for only one minute. The pathologists were asked to interpret the immunostaining of each image for the three cut-offs as a binary categorical variable, “present” or “absent”. The three cut-offs were: 10% cut-off (10% positive epithelial cells of any staining intensity), 20% cut-off (20% positive epithelial cells of any staining intensity) and +2/+3 cut-off (moderate or strong staining of any cells). For example, a 10% cut-off is “present” when more than equal to 10% cells are positive in the desired subcellular compartment and is “absent” when less than 10% cells are positive. Each core was also recorded as being easy (1) or challenging (2) to score. The scoring sheet was also provided with a column for any comments by the scorers.

All seven pathologists participated in the intra-observer part of the study three weeks after inter-observer session. The tissue core images were the same but rearranged in different order to remove recall bias.

#### 2.2.6.4 Statistics and data analysis

**Inter-observer agreement**

To determine inter-observer agreement for each of the three cut-offs, each pathologist’s interpretation of immunostaining was compared with that of the other pathologists in a pairwise manner. This generated three sets of inter-observer kappa (k) scores: one each for the 10%, 20% and +2/+3 cut-offs.

k scores generated served as a measure of strength of agreement between pathologists for all three cut-offs. k scores reflect strength of agreement between two observers adjusted for chance agreement. k scores range from 0 to 1 and we used the standards suggested by Landis and Koch (226) for the interpretation of strength of agreement. k scores are shown in six categories from 0-1 and each category is colour coded. White: 0, no agreement; greyscale indicates increasing agreement in the ranges 0.01-0.20= slight agreement (lightest grey), 0.21-0.40 fair agreement; 0.41-0.6, moderate agreement; 0.61-0.8, substantial agreement; 0.81-1.0, almost perfect (black). To determine whether these three cut-offs are statistically different from each other, the paired sample t test (for large sample size) and Wilcoxon signed ranked test (for small sample size) were used to compare the pairwise k scores.

**Perceived ease of scoring**

To determine which cut-off is most easily scored, these three cut-offs as predictor variables were put in a linear regression model against perceived ease of scoring as dependant variable.

**Intra-observer agreement**

To determine reproducibility of these three cut-offs, kappa scores were generated comparing scoring and re-scoring of the same image arranged in different order three weeks apart for each pathologist. The intra-observer part was carried out in two sessions with three pathologists participating in one session and four participating in another session.

A p value $<0.05$ was considered statistically significant. SPSS version 21 was used for statistical analyses.

**Table 2.1: Search Terms and Mesh Headings.**

| | |
|---|---|
| **Study Population:** | |
| **1.** | exp Pancreatic Neoplasms/ |
| **2.** | (pancrea* adj5 neoplas*).tw. |
| **3.** | (pancrea* adj5 cancer*).tw. |
| **4.** | (pancrea* adj5 carcin*).tw. |
| **5.** | (pancrea* adj5 adenocarcin*).tw. |
| **6.** | 1 or 2 or 3 or 4 or 5 |
| **Study Type:** | |
| **7.** | exp Diagnosis/ |
| **8.** | diagnos*.tw. |
| Other diagnostic terms important for search. | |
| **9.** | exp "Sensitivity and Specificity"/ |
| **10.** | sensitivity.tw. |
| **11.** | specificity.tw. |
| **12.** | predictive value*.tw. |
| **13.** | likelihood ratio*.tw. |
| **14.** | diagnostic accuracy*.tw. |
| **15.** | 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 |
| **Study Intervention:** | |
| **16.** | exp Immunohistochemistry/ |
| **17.** | (immunohistoch* or immunostaining or immunocytoch* or IHC).tw. |
| **18.** | (biomarker* or marker* or protein).tw. |
| **19.** | 16 and 17 and 18 |
| **Combining terms:** | |
| **20.** | 19 and 15 |
| **21.** | 20 and 6 |

**Table 2.2: IHC conditions used for antibodies in PDAC in literature.**

| Protein Assayed | Primary antibody | Antigen retrieval | Antibody dilutions | Incubation time | Incubation temperature |
|---|---|---|---|---|---|
| **KOC** | Mouse Monoclonal, Clone 69.1 | HIER<br>Citrate buffer pH 6.0 and Tris EDTA pH 9.0 | 1/50, 1/500 | 45, 60 minutes | RT |
| **Mesothelin** | Mouse monoclonal, Clone 5B2 | HIER<br>Citrate buffer pH 6.0 and EDTA pH 8.0 | 1/15, 1/20, 1/30 | Not available for most studies, Overnight | RT, 4°C |
| **MUC1** | Mouse monoclonal Clone DF3, Clone Ma695, Clone Mh1 | HIER<br>Citrate buffer pH 6.0 | 1/75, 1/100, 1/200, | 60 minutes, overnight | RT, 4°C |
| **S100P** | Mouse monoclonal, Clone 16 | Proteinase K (10 min) | 1/100 | 30minutes, | RT |
| **Maspin** | Mouse monoclonal, Clone G167-70 and Clone, EAW24 | HIER<br>Citrate buffer pH 6.0 | 1/20, 1/50, 1/75, 1/800, 1/3000 | 60 minutes, 12 hours | RT |

**Note:** the IHC conditions used for resection specimens and cytology specimens appears to be the same in the studies.

**Abbreviations:** HIER= Heat induced epitope retrieval, RT= room temperature.

**Table 2.3: The IHC conditions from literature review used as a starting point for IHC optimisation on tissue sections in our laboratory**

| Protein Assayed | Primary antibody | Antigen retrieval | Antibody dilutions | Incubation time | Incubation temperature |
|---|---|---|---|---|---|
| **KOC** | Mouse Monoclonal, Clone 69.1 | HIER Citrate buffer pH 6.0 | 1/50, 1/100, 1/200 | 60 minutes | RT |
| **Mesothelin** | Mouse monoclonal, Clone 5B2 | HIER Citrate buffer pH 6.0 | 1/20, 1/40, 1/80 | 60 minutes | RT |
| **MUC1** | Mouse monoclonal Clone Ma695, | HIER Citrate buffer pH 6.0 | 1/50, 1/100, 1/200, | 60 minutes | RT |
| **S100P** | Mouse monoclonal, Clone 16 | Proteinase K (10 min) | 1/50, 1/100, 1/200 | 60 minutes | RT |
| **Maspin** | Mouse monoclonal, Clone G167-70 | HIER Citrate buffer pH 6.0 | 1/75, 1/150 | 60 minutes | RT |

**Abbreviations:** HIER= Heat induced epitope retrieval, RT= room temperature.

**Table 2.4: Details of the final IHC parameters used on TMA sections for five antibodies**

| Antibody | Company | Clone of antibody | Antigen Retrieval | Antibody Dilution | Incubation temperature | Duration of incubation |
|---|---|---|---|---|---|---|
| **SD TMAs staining** | | | | | | |
| **KOC/IMP3** | DAKO | L523S, 69.1 | HIER (Citrate buffer pH 6) | 1/50 | 25˚C | 60 min |
| **S100P** | BD Biosciences | 16 | Proteinase K (10 minutes) | 1/100 | 25˚C | 60 min |
| **Mesothelin** | Novocastra | 5B2 | HIER (Citrate buffer pH 6) | 1/20 | 25˚C | 60 min |
| **MUC1** | Novocastra | Ma695 | HIER (Citrate buffer pH 6) | 1/200 | 25˚C | 60 min |
| **NBJ TMAs staining** | | | | | | |
| **Maspin** | BD Biosciences | G167-70 | HIER (Citrate buffer pH 6) | 1/75 | 25˚C | 60 min |

**Abbreviations:** HIER= Heat induced epitope retrieval, RT= room temperature.

**Table 2.5: The IHC conditions used for optimisation on 'needling' cytology sections**

| Protein Assayed | Primary antibody | Antigen retrieval | Antibody dilutions | Incubation time | Incubation temperature |
|---|---|---|---|---|---|
| **KOC** | Mouse Monoclonal, Clone 69.1 | HIER<br>Citrate buffer pH 6.0<br>Tris EDTA buffer pH 9.0 | 1/50 | 60 minutes | RT |
| **Mesothelin** | Mouse monoclonal, Clone 5B2 | HIER<br>Citrate buffer pH 6.0<br>EDTA buffer pH 8.0 | 1/20 | 60 minutes | RT |
| **S100P** | Mouse monoclonal, Clone 16 | Proteinase K (10 min) , Citrate Buffer pH 6.0, Tris EDTA pH 9.0 | 1/50, 1/100 | 60 minutes | RT |
| **Maspin** | Mouse monoclonal, Clone G167-70 | HIER<br>Citrate buffer pH 6.0 | 1/75, 1/150 | 60 minutes | RT |

**Abbreviations:** HIER= Heat induced epitope retrieval, RT= room temperature.

**Table 2.6: Details of the final IHC parameters used on archival cytology sections for three antibodies**

| Antibody | Company | Clone of antibody | Antigen Retrieval | Antibody Dilution | Incubation temperature | Duration of incubation |
|---|---|---|---|---|---|---|
| **KOC/IMP3** | DAKO | L523S, 69.1 | HIER (Citrate buffer pH 9) | 1/50 | 25°C | 60 min |
| **Maspin** | BD Biosciences | G167-70 | HIER (Citrate buffer pH 6) | 1/150 | 25°C | 60 min |
| **Mesothelin** | Novocastra | 5B2 | HIER (Citrate buffer pH 8) | 1/20 | 25°C | 60 min |

**Abbreviations:** HIER= Heat induced epitope retrieval

**Table 2.7: Scoring sheet for TMA cores with an example highlighted in red**

| | | | | | |
|---|---|---|---|---|---|
| **Core Type** | | | | | |
| | Tumour | **Tumour** | | | |
| | Normal | | | | |
| **Scorable Tissue** | | | | | |
| | Yes | **Yes** | | | |
| | No | | | | |
| **Biomarker expression** | | | | | |
| | Categories | Absent | Weak | Moderate | Strong |
| | 1-5% | | | | |
| | **6-10%** | **10%** | | | **10%** |
| | **11-25%** | | **20%** | | |
| | 26-40% | | | | |
| | **41-60%** | | | **60%** | |
| | 61-75% | | | | |
| | 76-90% | | | | |
| | 91-100% | | | | |
| **Final IHC Positvity** | **90%** | | | | |
| **Final Histoscore** | **170** | | | | |
| **Overall Staining Intensity** | Absent | Weak | Moderate | Strong | |
| | | | **Moderate** | | |
| **Scorer Comment** | | | | | |

**Note:** This algorithm of scoring has previously been used by our group (Denley et al 2013) (224).

**Table 2.8: Scoring sheet for observer agreement (Biomarker with cytoplasmic staining)**

| Core# | 10% + cells | 20% + cells | Moderate-Strong intensity (+2/+3) | Interpretation: Easy (1), Challenging (2) | Comments |
|---|---|---|---|---|---|
| **1.** | | | | | |
| **2.** | | | | | |
| **3.** | | | | | |
| **4.** | | | | | |
| **5.** | | | | | |
| **6.** | | | | | |
| **7.** | | | | | |
| **8.** | | | | | |
| **9.** | | | | | |

**Instructions for scoring:** Kindly mark each core as either + OR ⁻ for the 3 thresholds i.e. 10% positive cells, 20% positive cells and moderate-strong staining intensity. Score the ease of interpretation as 1 (easy) OR 2 (Challenging). If you wish to make comments, please use the last column.

**Figure 2.1: The methodology of 'needling' cytology samples collection.**

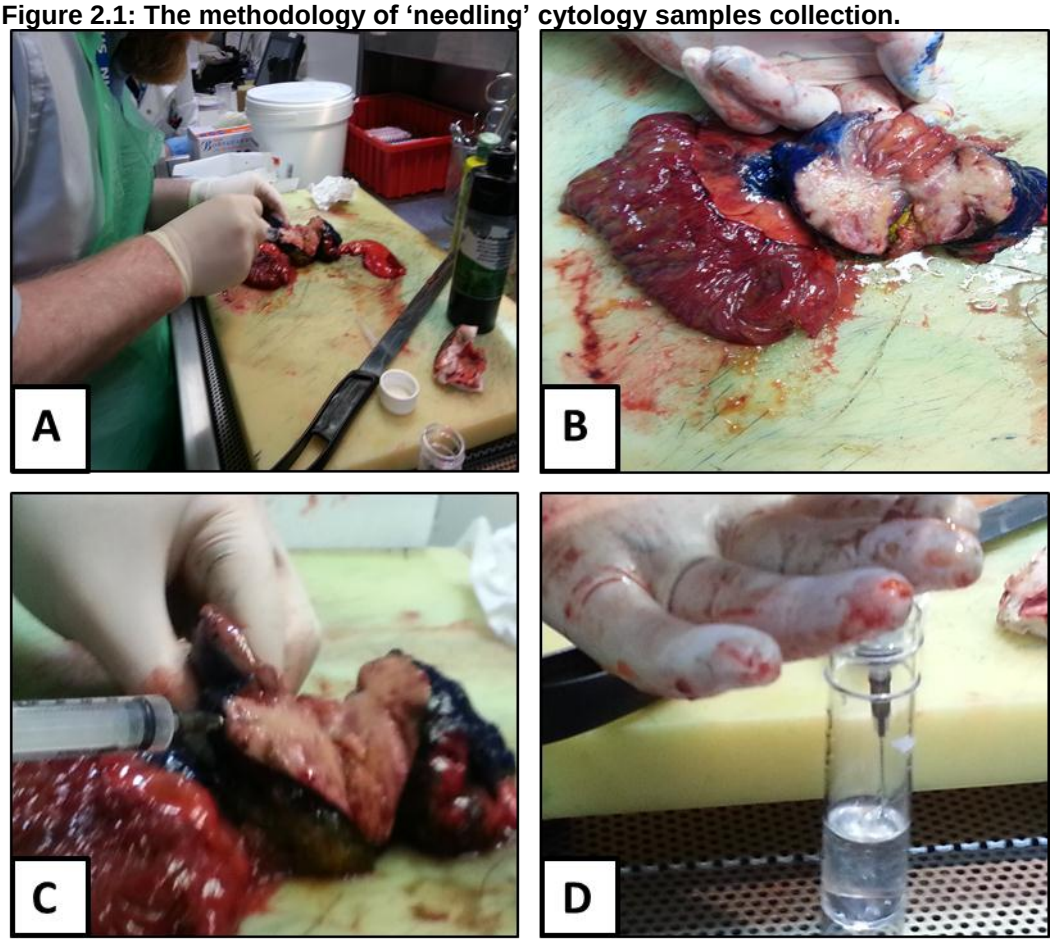



**Figure legend: A)** Opening pancreatic specimen **B)** Tumour visible on sectioning **C)** Collecting "needling cytology" sample **D)** Putting cells into PreservCyt solution.

**Figure 2.2: Optimisation of S100P on PDAC resection specimens**

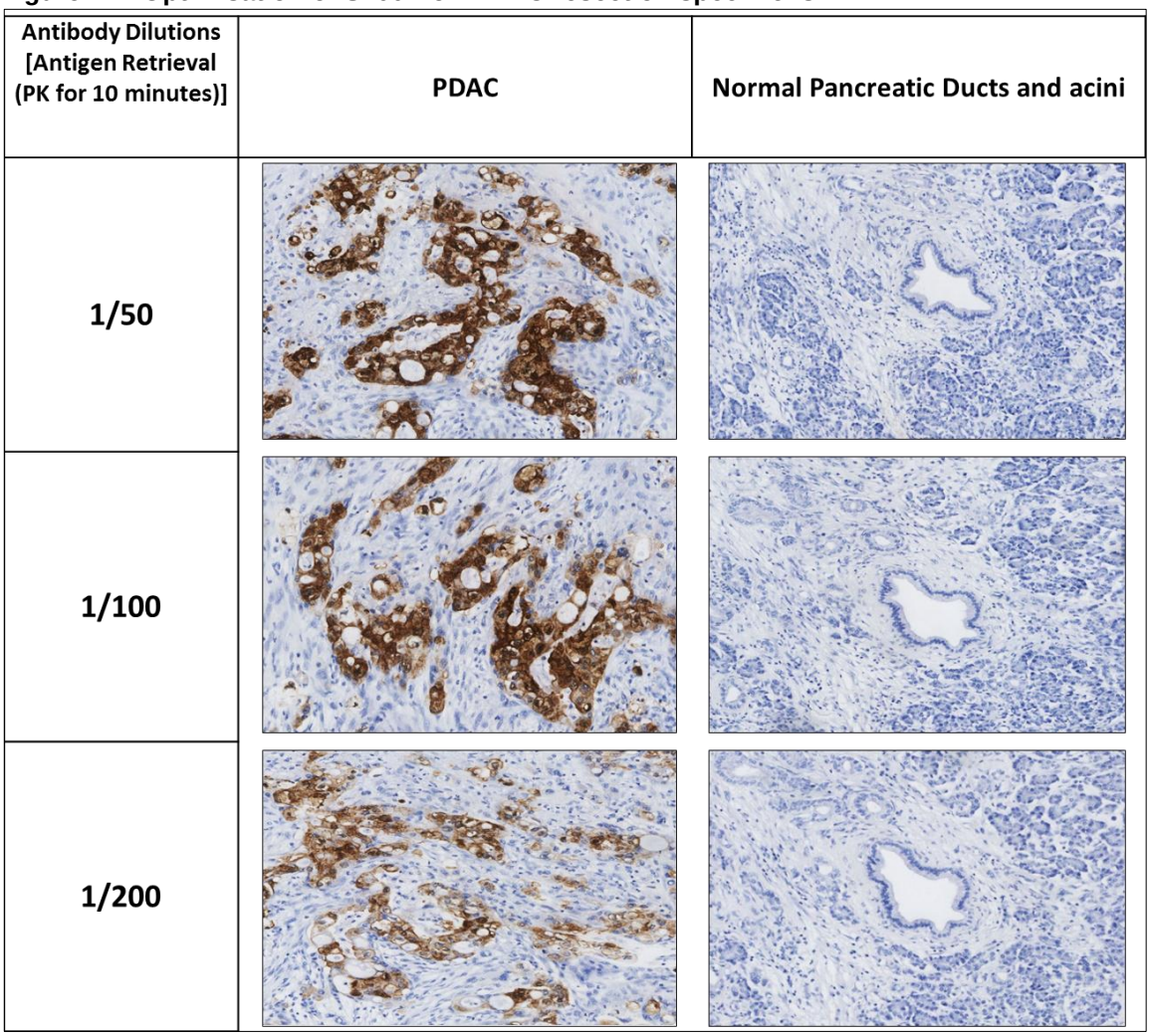



**Figure Legend:** Optimisation of S100P antibody for IHC on malignant ducts and adjacent normal ducts and acini from PDAC tissue sections. Note the strong and diffuse cytoplasmic and nuclear staining of S100P in malignant ducts for 1/50 and 1/100 dilutions. The staining intensity decreases with 1/200 dilution. No staining in normal ducts and normal acini was observed with any antibody dilution.

**Figure 2.3: Optimisation of maspin on PDAC resection specimens**

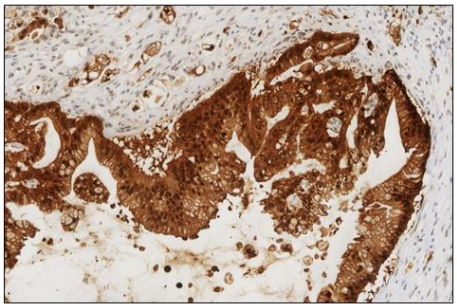
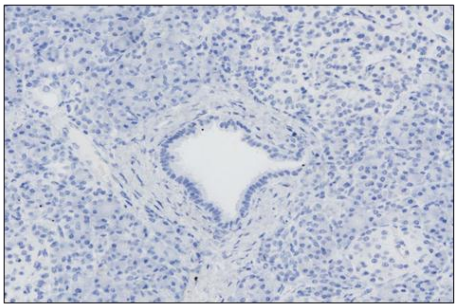
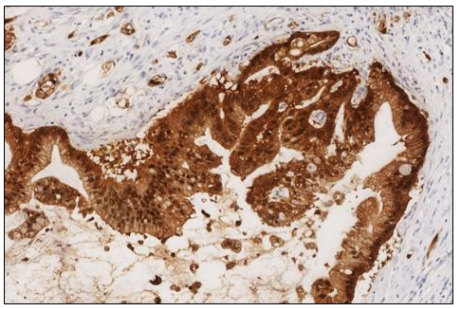
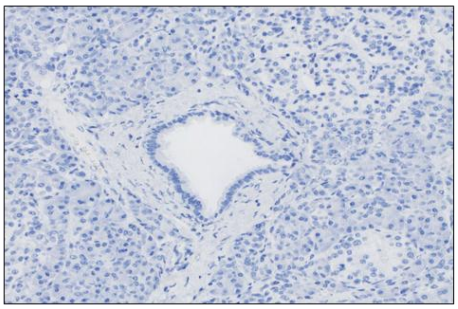
| Antibody Dilutions [Antigen Retrieval (PH6 Citrate Buffer)] | PDAC | Normal Pancreatic Ducts and acini |
|---|---|---|
| 1/75 |  |  |
| 1/150 |  |  |

**Figure Legend:** Optimisation of maspin antibody for IHC on malignant ducts and adjacent normal ducts and acini from PDAC tissue sections. Note the strong and diffuse cytoplasmic and nuclear staining of maspin in malignant ducts for 1/75 and 1/150 dilutions. No staining in normal ducts and normal acini was observed with any antibody dilution.

# 3 Systematic review and meta-analysis of immunohistochemical diagnostic biomarkers for PDAC

## 3.1 Chapter Summary

Pancreatic ductal adenocarcinoma (PDAC) has an overall five year survival of only 2%. Diagnosis involves imaging followed by endoscopy with cytology, and is important for patient management. Cytology involves the distinction of PDAC from non-neoplastic pancreas, which can be difficult, especially in chronic pancreatitis. Immunohistochemical (IHC) biomarkers could help but none is yet routinely used. This study provides a systematic review and meta-analysis on diagnostic IHC biomarkers for PDAC.

The literature was searched using EMBASE and MEDLINE databases from inception to March 2012. The publications on IHC markers differentially expressed between human PDAC and non-neoplastic pancreas were sought. The study characteristics including specimen type, biomarkers assessed and staining results were catalogued. In the meta-analysis, for each biomarker, coupled Forest plots, bivariate summary estimates and combined summary receiver operating characteristic curves were generated, in turn, then compared and ranked according to pooled sensitivity/specificity.

2089 papers were initially identified. 64 studies reporting 49 biomarkers were selected for systematic review. From these, 45 studies reporting 16 biomarkers progressed to meta-analysis. Meta-analysis of IHC biomarkers assessed in resection specimens showed 12 differentiating PDAC from non-neoplastic pancreas. The highest ranked biomarkers according to pooled sensitivity/specificity values were: S100P (100% sensitivity/100% specificity); maspin (92%/97%); KOC (IMP3) (85%/98%); and MUC4 (82%/93%). Meta-analysis of cytology specimens showed seven biomarkers. The highest ranked were: KOC (85%/100%); SMAD4 (80%/100%); S100P (91%/91%); mesothelin (64%/92%); and MUC1 (83%/77%).

The highest ranking IHC markers for PDAC were KOC, S100P, maspin, mesothelin and MUC1. Their diagnostic accuracies approach those of optimal conventional cytology. These markers may be appropriate for further clinical validation and potentially routine use in difficult cases.

## 3.2 Introduction

Establishing the diagnosis of PDAC is important for optimal patient management but can be difficult and relies on imaging and pathology. Pre-treatment confirmation of diagnosis is achieved from cytology samples but the diagnosis from morphological interpretation of cytology is not perfect. This unmet clinical need in PDAC diagnosis from cytology samples has been discussed in detail in chapter 1. In this chapter, I will discuss the identification of potential diagnostic biomarkers that could help us improve the diagnosis of PDAC. The identification will involve a systematic review and meta-analysis of potential biomarkers already investigated in PDAC. Better biomarkers identified will then be validated in our surgical and cytology cohorts.

We know that many genes are differentially expressed between benign pancreas and PDAC and could potentially be used as an adjunct to cytology to improve diagnostic accuracy (132, 133, 227). These genes are characterised at the protein level by IHC (147). IHC is a standard technique in pathology and a wide range of IHC biomarkers has been investigated in pancreatic resection specimens and EUS-FNA and biliary brush cytology specimens (169-172, 228). No biomarker has yet entered routine clinical practice but, used as an adjunct to cytology, they could help to increase diagnostic accuracy in indeterminate or otherwise difficult-to-diagnose cases (172).

As mentioned earlier, a wide range of diagnostic IHC biomarkers have been published and I believe that a review of these biomarkers is the first step in the identification and selection for clinical translation. Although, outstanding reviews (229-231) on IHC biomarkers have been published for PDAC, none has specifically evaluated and quantified diagnostic IHC biomarkers.

### 3.2.1 Aims

The aim was to explore the current evidence on diagnostic IHC biomarkers in PDAC literature and to review, quantify and rank the performance of these biomarkers for PDAC. A further aim was to assess the quality of primary papers reporting diagnostic IHC biomarkers and to observe the range of IHC methodology, microscopic interpretation of immunostaining and diagnostic cut-

offs. This should enable us to identify suitable candidate biomarkers to be investigated for their clinical utility for PDAC diagnosis.

## 3.3 Results

The results are reported in four main sections. The first section describes the screening process and selection of papers for systematic review and meta-analysis. The second section presents the meta-analysis of 16 potentially diagnostic biomarkers for PDAC. The third section describes the remaining 33 biomarkers from the systematic review not included in the meta-analytical model. The fourth section provides an overview of the scoring systems used to evaluate immunostaining and the cut-offs used for diagnostic purposes.

### 3.3.1 Screening and selection of papers based on inclusion and exclusion criteria

The prisma (http://www.prisma-statement.org/) flow chart was used as a template for the screening and selection of papers. At initial screening, duplicate publications were removed. The titles and abstracts of papers were inspected for potential eligibility for inclusion by the author (A.A). For the papers which passed this screening process, full text articles were then retrieved and checked for suitability.

The inclusion criteria were: studies on tissue biomarkers which showed differential expression between PDAC and non-neoplastic pancreas (normal pancreas, chronic pancreatitis or both); assessment of tissue biomarkers using IHC; studies on cytology samples, surgical resection specimens or both; either prospective or retrospective study designs; reference diagnosis confirmed through either tissue histology or death from PDAC; and availability of either absolute numbers for biomarker sensitivity and specificity or of the raw data (for example a 2×2 table) to enable their calculation.

The exclusion criteria were: studies reporting tumours which were not PDAC for example cholangiocarcinoma, ampullary or duodenal adenocarcinoma, and pancreatic neuroendocrine or acinar tumours or tumours from other organs; studies reporting only benign pancreatic pathology; studies reporting biomarkers which were not potentially diagnostic for example prognostic, predictive or

related to tumorigenesis only; studies on fluids for example pancreatic cyst fluid or juice, bile or serum; and molecular markers tested in tissue but not validated by IHC.

#### 3.3.1.1 64 studies reporting 49 biomarkers were included in the systematic review; and from these, 45 studies reporting 16 biomarkers progressed to meta-analysis.

Figure 3.1 presents a prisma flow chart describing in detail the process of selection of studies for systematic review and meta-analysis based on the inclusion and exclusion criteria. 2089 studies were initially identified and 1618 studies remained after the removal of duplicates. Further studies were screened out based on abstract and title, to leave 227 studies which were checked at the full text level for eligibility. Eventually 64 studies reporting 49 biomarkers were included in the systematic review; and from these, 45 studies reporting 16 biomarkers progressed to meta-analysis.

The general characteristics of the 64 selected studies including reference, year and country of publication, biomarker investigated either in resection specimens or cytology specimens or both, biomarker investigated in either a PDAC vs. normal pancreas or PDAC vs. chronic pancreatitis comparison or both and the design of study are summarized in Table 3.1.

Next I will discuss the meta-analysis of 16 biomarkers, the outcome variables for each biomarker from the primary papers included in the meta-analysis and the methodological qualities of primary papers included in the meta-analysis.

### 3.3.2 Sixteen biomarkers studied by IHC were reported in more than one study and were therefore included in the meta-analysis

The meta-analytical results are reported separately for biomarkers studied in resection specimens and for those studied in cytology samples. This grouping of biomarkers into different specimen types is based on the way that these biomarkers were reported in primary papers. The analysis is further subdivided into biomarkers differentiating PDAC from normal pancreas and biomarkers differentiating PDAC from chronic pancreatitis. Again this categorisation is based on the manner that biomarkers were reported in primary papers. It is important

to note here that there is overlap of biomarkers between resection specimens and cytology specimens i.e. a biomarker will be assessed both in resection specimen and cytology specimens.

#### 3.3.2.1 Twelve out of sixteen biomarkers which progressed to meta-analysis were tested in resection specimens

These 12 biomarkers are: maspin (171, 172, 232-235), mesothelin (172, 173, 236-239), KOC (IMP3) (186, 240, 241), fascin (172, 242, 243), COX2 (185, 244, 245), S100P (159, 166), claudin18 (227, 246), S100A6 (166, 247), S100A4 (166, 248), PSCA (249, 250), CEACAM6 (172, 251) and MUC4 (171, 173, 252). The sensitivities and specificities of these biomarkers are shown in the coupled Forest plots in Figure 3.2. Out of these 12 biomarkers for PDAC, 11 were compared with normal pancreas, four were compared with chronic pancreatitis and three were compared with both (Table 3.2).

**Eleven out of twelve biomarkers tested in resection specimens in the literature differentiated PDAC from normal pancreas**

11 biomarkers differentiated PDAC from normal pancreas (Table 3.2). For five biomarkers, maspin, mesothelin, KOC, fascin and COX2, three or more studies were available and so a random effects model (REM) could be applied. Overall sensitivities were 83-92% and overall specificities were 89-98%. Pooled sensitivity/specificity (Table 3.2) and combined SROC curve (Figure 3.3A) were used to rank the biomarkers according to combined sensitivity/specificity. Maspin was highest ranked with a pooled sensitivity of 92% and pooled specificity of 97.5%, followed by mesothelin, KOC, fascin and COX2.

For six biomarkers, S100P, claudin18, S100A6, S100A4, PSCA and CEACAM6, only two studies were available and so a fixed effect model (FEM) could be applied. Overall sensitivities were 75-100%, and overall specificities were 72-100%. Pooled sensitivity/specificity (Table 3.2) and combined SROC curve (Figure 3.3B) were used to rank the biomarkers according to combined sensitivity/specificity. S100P was highest ranked with pooled sensitivity and specificity of 100%, followed by claudin 18, S100A6, S100A4, PSCA and CEACAM6.

The biomarker expression in PDAC compared to normal pancreas provides an insight into the diagnostic role of biomarkers. A biomarker highly expressed in PDAC is likely to be highly sensitive in identifying the malignancy and a biomarker with less expression in normal pancreas is likely to be highly specific. In summary, maspin, KOC, mesothelin and S100P are good candidates as they achieve a higher diagnostic accuracy in categorising PDAC and normal pancreas in resection specimens.

**Four out of twelve biomarkers tested in resection specimens in the literature differentiated PDAC from chronic pancreatitis**

Furthermore, four biomarkers differentiated PDAC from chronic pancreatitis (Table 3.2). For three biomarkers, KOC, mesothelin and maspin, three or more studies were available and so REM could be applied. Overall sensitivities were 85-87% and overall specificities were 68-98%. Pooled sensitivity/specificity (Table 3.2) and combined SROC curve (Figure 3.3C) were used to rank the biomarkers according to combined sensitivity/specificity. KOC was highest ranked with a pooled sensitivity of 87% and pooled specificity of 98%, followed by mesothelin, and maspin.

One biomarker MUC4 was assessed in 2 studies was suitable for FEM differentiating PDAC from chronic pancreatitis (Table 3.2 and Figure 3.3D). Pooled sensitivity and specificity for MUC4 were 82% and 93% respectively.

The biomarker expression in PDAC compared to chronic pancreatitis provides a further insight into the diagnostic role of biomarkers. A biomarker with less expression in the reactive process makes it a desirable candidate for differentiating PDAC from a reactive but benign disease process. In summary, both KOC and mesothelin are primarily good candidates as they achieve a higher diagnostic accuracy in categorising PDAC and chronic pancreatitis in resection specimens.

Overall, 12 potentially diagnostic biomarkers with reasonable diagnostic accuracy were investigated in resection specimens from patients with PDAC and non-neoplastic pancreas. The high ranking biomarkers were KOC, maspin, mesothelin and S100P. The ranking was based on best combination of sensitivity

and specificity values which was facilitated by SROC curves. Next I assessed the potential diagnostic utility of these and other biomarkers in cytology specimens reported in the literature.

#### 3.3.2.2 Seven out of sixteen biomarkers which progressed to meta-analysis were tested in cytology specimens

These seven biomarkers are: KOC (IMP3) (186, 217, 240, 253), mesothelin (172, 173, 254, 255), MUC1(170, 176, 188), p53 (165, 187, 256-259), SMAD4 (139, 228), S100P (159, 177), and MUC5AC (176, 188). The sensitivities and specificities of these biomarkers are shown in the coupled Forest plots in Figure 3.4. Out of these seven biomarkers for PDAC, two were compared with normal pancreas, and six were compared with chronic pancreatitis and one was compared with both (Table 3.3).

**Two out of seven biomarkers tested in cytology specimens in the literature differentiated PDAC from normal pancreas**

For two biomarkers, KOC and SMAD4, only two studies were available and so FEM could be applied. The pooled sensitivity and specificity of KOC was 85% and 100%. The pooled sensitivity and specificity of SMAD4 was 80% and 100% (Table 3.3). SMAD4 was higher ranked than KOC based on the combined SROC curve (Figure 3.5A).

Although the KOC offers better sensitivity indices, SMAD4 is higher ranked due to comparatively large sample size. In summary, both KOC and SMAD4 are good candidates as they achieve a higher diagnostic accuracy in categorising PDAC and normal pancreas in cytology samples. The 2 studies reporting SMAD4 are from the same research group.

**Six out of seven biomarkers tested in cytology specimens in the literature differentiated PDAC from chronic pancreatitis**

Six biomarkers differentiated PDAC from chronic pancreatitis (Table 3.3). For four biomarkers, KOC, mesothelin, MUC1 and p53, three or more studies were available and so REM could be applied. Overall sensitivities were 64-84% and overall specificities were 77-92%. Pooled sensitivity/specificity (Table 3.3) and

combined SROC curve (Figure 3.5B) were used to rank the biomarkers according to combined sensitivity/specificity. KOC was highest ranked with a pooled sensitivity of 84% and pooled specificity of 92%, followed by mesothelin, MUC1 and P53.

For two biomarkers, S100P and MUC5AC, only two studies were available and so FEM could be applied. The pooled sensitivity and specificity of S100P was 90.9% and 90.9%. The pooled sensitivity and specificity of MUC5AC was 87.3% and 58.8% (Table 3.3). S100P was higher ranked than MUC5AC based on the combined SROC curve (Figure 3.5C).

In summary, KOC, SMAD4, S100P, mesothelin and MUC1 are good candidates as they achieve a higher diagnostic accuracy in categorising PDAC and chronic pancreatitis in cytology samples.

Overall, seven potentially diagnostic biomarkers with reasonable diagnostic accuracy were investigated in cytology specimens from patients with PDAC and non-neoplastic pancreas. The high ranking biomarkers were KOC, SMAD4, mesothelin, S100P and MUC1. The ranking was based on best combination of sensitivity and specificity values which was facilitated by SROC curves.

#### 3.3.2.3 Linking meta-analytic results: The highest ranking candidate diagnostic IHC biomarkers are KOC, S100P, maspin, mesothelin and MUC1

In total, 16 biomarkers were quantified and ranked to identify high ranking candidates for potential clinical translation. As reported in the literature, biomarkers have been investigated in resection and cytology specimens. Therefore, I performed analysis separately for resection and cytology specimens. In addition, sub-classification was done based on whether biomarker expression in PDAC was compared with normal pancreas and/or chronic pancreatitis as described in primary research studies.

Twelve biomarkers investigated in resection specimens were suitable for meta-analysis. Five biomarkers: KOC, maspin, mesothelin, S100P and claudin 18 appear promising in differentiating PDAC from non-neoplastic pancreas (normal pancreas and chronic pancreatitis). They achieve reasonable diagnostic

sensitivity and specificity. Similarly, seven biomarkers investigated in cytology specimens were found suitable for meta-analysis. The five biomarkers KOC, SMAD4, mesothelin, S100P and MUC1 appear to be superior in differentiating PDAC cytology from benign cytology.

KOC, mesothelin and S100P have been extensively investigated in both resection and cytology specimens. KOC was found to have similar diagnostic accuracy; however, the specificity of mesothelin remained the same but the sensitivity differed between resection and cytology specimens. The sensitivity and specificity of S100P dropped in cytology specimens as compared to resection specimens.

These biomarkers are intended to improve the diagnosis of cytology samples and in particular the distinction of PDAC from chronic pancreatitis. Thus I used the following criteria to select the high ranking candidates: pooled sensitivity and specificity values from meta-analysis; ranking by SROC curves; investigation in a PDAC vs. chronic pancreatitis category; and validation in cytology samples by independent laboratories. Based on the above criteria KOC, S100P, mesothelin, maspin and MUC1 are high ranking potential candidates. It might be useful to further validate these biomarkers ideally in one experimental setting.

#### 3.3.2.4 The outcome variables for each biomarker from the primary papers included in the meta-analysis

The outcome variables of all 16 biomarkers have been detailed in Table 3.4. Each biomarker is reported with study details, including sample size, microscopic assessment and scoring system, threshold of immunostaining for positive or negative diagnosis, staining compartment of cell (cytoplasmic, nuclear and membranous) and the diagnostic sensitivity and specificity values in the type of specimen (cytology and/or resection specimen).

These factors are important in evaluating the biomarkers for potential clinical translation. Sample size is undoubtedly the most important factor giving the effect size of the study. Scoring system and microscopic assessment of immunostaining is the first step in evaluating a potential IHC biomarker and as evident from Table 3.4 a variety of scoring systems have been used. The next

step is the cut-off point whereby categorisation of two conditions (for example malignant vs. benign) based on immunostaining is performed. In addition, Table 3.4 outlines the sensitivity and specificity values of biomarkers reported or calculated in various studies. These values were used to pool evidence from various articles to quantify and rank the performance of biomarkers which were used in different meta-analytical models.

Finally, in Table 3.5 I have provided the details of the IHC methodology reported in primary papers for each biomarker included in the meta-analysis. These details include the primary antibody used with supplier details, the method of antigen retrieval used and the dilution of primary antibody used for IHC.

#### 3.3.2.5 Methodological quality of included studies evaluated by QUADAS checklist showed varying pattern of quality for 10 items assessed

The aim of meta-analysis is to review and quantify the available research evidence on a particular objective from primary studies. If these primary studies are biased then synthesising evidence from them without any consideration to their quality will introduce bias to the meta-analysis. Quality assessment should thus be an integral part of a meta-analysis that allows assessing the effect of different biases and sources of variations (260, 261).

The aim of diagnostic accuracy studies is to assess a particular test in identifying a diseases process and these studies have the following basic structure. An “index test” is the test of interest used to identify a disease process and to exclude normal processes, while a “reference standard” is the gold standard for diagnosis and the index test is compared to this gold standard (261). These terms will be used for assessing the quality of primary studies included in the current meta-analysis. For the purposes of this current project, the index test is biomarker, the reference standard is histopathology and/or clinical follow up, the disease process is PDAC and normal process is non-neoplastic pancreas. The performance of the index test is calculated using various statistics such as sensitivity, specificity, and receiver operating characteristic (ROC) curves.

The methodological quality of primary studies included in the meta-analysis was therefore assessed using the QUADAS (**Q**uality **A**ssessment of **D**iagnostic **A**ccuracy

**S**tudies) checklist (261). This list of ten questions is intended to confirm study quality, on topics including: selection criteria; identification of potential sources of bias; description of the index test; suitability of the reference standard; and reporting of indeterminate cases and withdrawal (Table 3.6). These factors can impact on the accuracy of diagnostic tests.

One paper was written in the Chinese language only and was not assessed for quality; this left 44 studies for assessment. Figure 3.6 and 3.7 are graphical presentation of the methodological quality of included studies.

The following is an assessment of the primary papers for ten items in the QUADAS checklist. Representative spectrum and sufficient description regarding the selection of patients were found in 35%, while it was not clear in rest of the papers (item 1). Almost all studies verified all their samples by an acceptable reference standard (histopathology and/or death from pancreatic cancer) (item 2, 3). But in 50% of the included studies differential verification was present, which means that the verification of true disease status was not performed solely by one reference standard. In almost 45% of the papers the true disease status was verified by death from pancreatic cancer (item 4).The incorporation of an index test as an aid to reference standard was avoided in almost all included studies (item 5).

We used the minimum criteria in reporting the antigen retrieval method, primary antibody used and antibody dilution for assessing the reproducibility of IHC methodology for the index test (item 6). We found that sufficient details of IHC methodology was reported in nearly 85% of the papers; while sufficient details were not provided in the rest of the primary papers.

In almost all cases the reference test was interpreted without knowledge of the index test (item 7). However, in 55% of the papers it was not clear whether the index test results were interpreted with or without knowledge of the reference standard (item 8). Finally, reporting of uninterpretable results and withdrawal of specimens was clear in 60% and 75% of the studies respectively (item 9 & 10).

In summary, some of the primary studies are good quality but most of them have issues related to the reporting of selection criteria and representative samples,

differential verification, insufficient details of IHC methodology for assessing the index test for reproducibility, interpretation of the index test without knowledge of the reference standard and the reporting of uninterpretable and withdrawal cases. Addressing these issues in primary studies might greatly facilitate improving the quality of IHC biomarker research in PDAC.

Next I will discuss a review of important biomarkers investigated in only one paper. These biomarkers by definition were not eligible for meta-analysis.

### 3.3.3 Thirty three biomarkers studied by IHC were reported in only one study and were included in the review.

The 33 biomarkers reported in only one study by IHC were found suitable for review and are detailed as follows: 14 biomarkers with reasonable sensitivity and/or specificity; 11 negative biomarkers for PDAC (biomarkers expressed in normal ducts but not in PDAC); additional eight important biomarkers but with low sensitivity and/or specificity.

#### 3.3.3.1 Fourteen biomarkers with reasonable sensitivity and/or specificity were reported in only one study.

Fourteen IHC biomarkers with reasonable sensitivity and/or specificity were reported in only one study and are as follows: 14 3-3 σ (172), annexin A8 (227), B7-H4 (165), bcl2 (262), CK19 (139), claudin 4 (250), CXCL16 (263), Fas-L (264), HMGI(Y) (265), HSP47 (243), ICAM1 (266), PAP (267), RCAS1 (268) and XIAP (159).

Ten biomarkers i.e. annexin A8, bcl2, claudin4, CXCL16, Fas-L, HMGI(Y), HSP47, ICAM1, PAP and RCAS1 have been studied in resection specimens only, two biomarkers i.e. B7-H4 and CK 19 have been studied in cytology specimens only, and two biomarkers i.e. 14 3-3 σ and XIAP have been studied in both. Promising biomarkers in resection specimens with their corresponding sensitivity and specificity values are HMGI(Y) (sensitivity=100% and specificity=100%), XIAP (100% and 100%), 14 3-3 σ (80% and 100%) and Fas-L (100% and 90%). While favourable markers in cytology specimens are 14 3-3 σ (100%/80%) and CK 19 (100% and 100%).

Three studies reporting CEACAM6 (184), COX2 (269) and maspin (270) reported biomarker sensitivity only. Nonetheless, they were included in the review due to the presence of these biomarkers in meta-analysis. In addition, KOC (271) facilitate the differentiation of PDAC from various benign pancreatic lesions (for example inflammatory cyst, mucinous cystadenoma, serous microcystic adenoma) with 88% sensitivity and 100% specificity. However, this study was not included in the meta-analysis because we aimed for biomarkers comparing PDAC with normal pancreas and/or chronic pancreatitis.

All 14 biomarkers with their outcome variables are outlined in Table 3.7.

#### 3.3.3.2 Eleven negative biomarkers for PDAC (biomarkers expressed in normal ducts but not in PDAC) were reported

The expression of "negative biomarkers" is absent or very low in PDAC but their expression is high in normal pancreatic ducts; thereby, helping to exclude PDAC and identify normal ducts. These biomarkers are as follows; amylase (272), Bcl10 (272, 273), CDX2 (274), CEH (carboxyl ester hydrolase) (272), chromogranin A (274), HPK1 (Hematopoietic progenitor kinase 1) (275), lipase (272), MUC2 (276), synaptophysin (274), trypsin (272) and von Hippel-Lindau gene product (pVHL) (166). As an example, 0% sensitivity would mean no staining in PDAC for a biomarker, whereas 0% specificity would mean diffuse positive staining in normal pancreatic ducts for a biomarker.

All 11 biomarkers with their outcome variables are outlined in Table 3.8.

#### 3.3.3.3 Eight important diagnostic biomarkers with low sensitivity and/or specificity were reported

We also included eight biomarkers that were investigated as diagnostic biomarkers for PDAC but their sensitivity and/or specificity was lower than expected. These biomarkers are; CA19-9 (277, 278), CA-50 (278), CD10 (279), CDX2 (276), CK7 (274, 276), CK 17 (276), CK20 (276) and MUC6 (176).

All eight biomarkers with their outcome variables are outlined in Table 3.9.

### 3.3.4 Scoring systems and cut-offs reported in primary studies are variable

The scoring of immunostaining and then establishing a cut-off for diagnostic purposes is an important part of biomarker research. The microscopic interpretation of staining is significantly different between studies as evidenced from a variety of scoring systems reported for biomarkers in this review. After scoring, a cut-off is established to categorise patients into cancer or normal groups. The cut-offs reported for biomarkers in primary studies and compiled in this review clearly show significant variations in cut-offs (Table 3.4, 3.7, 3.8, 3.9). These variations exist sometimes for the same biomarker.

Let us take the example of mesothelin for heterogeneity of scoring systems and cut-offs. Mesothelin has been reported in seven studies and all studies have used different scoring systems and cut-offs. The scoring systems are based on staining intensity (164, 255), proportion of positive cells (238) or a combination of both staining intensity and the proportion of positive cells (173, 236, 237, 239, 254). Now let us take the staining intensity as a scoring system for mesothelin. McCarthy et al (255) have reported two categories (positive and negative) based on staining intensity. The positive staining is regarded as a strong diffuse cytoplasmic staining, while negative staining is regarded as no labelling over background. Agarwal et al (164) have reported four categories based on staining intensity as follows: 0 (no staining); 1 (faint); 2 (definite) 3 (intense) and 4 (very intense). Clearly, this scoring heterogeneity demonstrates the adoption of different and 'novel' style of scoring by different labs that makes clinical translation less straightforward. This variation may also contribute to the variable diagnostic sensitivity and specificity of biomarkers.

Cut-offs or positive thresholds for mesothelin are also heterogeneous and are as follows: simple positive or negative staining (164, 236, 238, 255); focal staining (>1% positive cells) (237); >5% positive cells (239); and >5% positive cells and moderate staining intensity (173). Again these different cut-offs for the same biomarker for the same outcome (PDAC diagnosis) make it difficult to compare and to select a reasonable cut-off. This also leads to the variable diagnostic sensitivity and specificity of biomarkers. The scoring systems and cut-offs for all biomarkers reported in this review are outlined in Table 3.4, 3.7, 3.8 and 3.9.

This heterogeneity of scoring systems and cut-offs is not restricted to mesothelin but applies to most of the biomarkers reported in our systematic review. This inconsistency between different labs makes it more difficult for translational scientists and pathologists to select a scoring system and cut-off for future validation studies. A standardised and well validated scoring system and cut-off is an important factor for clinical translation of biomarkers. It is evident based on the results of this review that a range of scoring systems and cut-offs exist for biomarkers for PDAC diagnosis.

## 3.4 Discussion

**We systematically reviewed the available evidence on diagnostic IHC biomarkers that are able to differentiate PDAC from non-neoplastic pancreas.** 64 studies reporting 49 IHC biomarkers were selected for systematic review. From these, 45 studies reporting 16 biomarkers progressed to meta-analysis. Reporting the methodology and meta-analytical results of the included studies presented some challenges; however, we successfully reported them by a variety of statistical models. These models are well established in the literature (223, 280, 281) but the description is lacking in the diagnostic IHC for PDAC and probably for any cancer literature.

**Meta-analysis was performed separately for biomarkers investigated in resection specimens and cytology specimens.** This grouping was based on the approach adopted by the primary papers included in the meta-analysis. In addition, biomarkers investigated in each specimen type were reported in two different comparisons in primary papers: PDAC compared to normal pancreas; and PDAC compared to chronic pancreatitis. The meta-analysis was thus performed mirroring the primary papers reporting these biomarkers. Clearly, this will make the interpretations from the meta-analysis easy and clinically oriented. Using this approach, the meta-analysis of resection specimens showed four promising diagnostic biomarkers KOC, S100P, maspin and mesothelin. These biomarkers could potentially diagnose PDAC with high sensitivity and specificity. A relative decrease in sensitivity and sharp decrease in specificity for maspin was observed in the PDAC versus chronic pancreatitis category compared to the PDAC versus normal pancreas category in resection specimens. This apparent difference is probably due to a different clone of primary antibody (EAW24) used by Agarwal et l 2008 (172) compared to the clone (G167-70) used in other studies reporting maspin (233-235).

Meta-analytical results of cytology specimens showed five promising biomarkers KOC, SMAD4, S100P, mesothelin and MUC1. SMAD4 has a pooled 80% sensitivity and 100% specificity but the 2 studies reporting SMAD4 were carried out by the same research group (139, 141) introducing subjective bias to the results. MUC5AC has a reasonable sensitivity but it has relatively very low pooled specificity of 58.8%, which means that it might classify benign lesions as

malignant. SMAD4 and MUC5AC are thus not desirable candidates. This left three biomarkers KOC, S100P and MUC1 as better candidates for further validation.

**Importantly, there is an overlap of various markers between different scenarios.** KOC, maspin, mesothelin and S100P have been investigated in both resection and cytology specimens. The diagnostic accuracy indices for KOC are optimal and consistent across all scenarios investigated (174, 175, 180, 186, 217) making it a favourable diagnostic candidate for future clinical validation. Although S100P was 100% sensitive and specific in resection specimens, the sensitivity and specificity dropped to 91% in cytology specimens (159, 166, 177). In addition, the sensitivity of mesothelin reduced by about 20%-30% in cytology compared to resection specimens (173, 236, 238, 254, 255). This probably reflects that some biomarkers (KOC) are consistent in expression pattern between resection and cytology specimens but others (mesothelin, S100P) might show a different expression pattern in the two specimen types. This might be attributed to the differences in the processing of these specimen types or to the very few malignant cells in some cytology samples. The expression level of biomarkers in resection specimen might predict expression level in cytology specimens. A biomarker showing strong and diffuse expression in malignant ducts in tissue sections might likely show higher expression levels in malignant cells in cytology. The reporting of the expression level of a potentially diagnostic biomarker on a continuous scale for example a Histoscore (0-300) (282, 283) may perhaps be helpful. This might provide more information into the potential utility of a biomarker for diagnostic purposes than simply reporting sensitivity and specificity.

**Meta-analysis for each biomarker comprised of coupled Forests plots, bivariate summary estimates and combined SROC curve.** Coupled Forest plots were prepared to provide an overview of the ranges of sensitivity and specificity of biomarkers. Bivariate summary estimates of various biomarkers were estimated by assuming random effects and fixed effect models depending on the number of studies for a biomarker. The bivariate summary estimates of biomarkers were interpreted in pairs for sensitivity and specificity. Finally, the SROC curve helped to compare and rank the diagnostic accuracy of various biomarkers in a given scenario (280, 281). This comprehensive meta-analytical

approach is rare in diagnostic biomarker research and is novel for PDAC literature. This approach might help translational research in other cancer types.

**Methodological limitations in the primary studies have introduced bias and variability to the diagnostic accuracy of the biomarkers under evaluation.** A significant deficiency was observed in the reporting of selection criteria and the representativeness of the included specimens. Almost two third of the included studies did not achieve the minimum criteria of consecutive sampling, age distribution and male to female ratio of the included participants. Moreover, sample size for some of these included studies was low (139, 174, 217, 238, 245, 258) decreasing the power of the studies. Differential verification bias is present in half of the included studies due to two different reference standards (histopathology and/or clinical follow up). The reason for having two reference standards is the fact that not all patients undergoing cytology receive surgical treatment and therefore death from pancreatic cancer due to locally advanced or metastatic disease confirms the diagnosis of pancreatic cancer. However, clinical follow up is an acceptable reference standard (172, 217, 255). Moreover, in most studies it was not clear whether the interpretation of the index test was 'blinded'. Similarly, reporting of un-interpretable results and the withdrawal of samples has not been mentioned for some studies. Both these factors introduce subjective bias into the diagnostic accuracy of a study. The quality of primary studies evaluating potential diagnostic IHC biomarkers could potentially be improved by addressing the issues outlined in the QUADAS assessment (261). Reporting all the QUADAS items in primary research studies is important. Nonetheless, sample size, providing sufficient details of IHC methodology for independent reproducibility and the interpretation of immunostainng without knowledge of true disease state are the most important factors.

**Among the papers included in our study heterogeneity exists in the reporting of IHC and cut-offs for positivity.** Scoring of the immunostaining of tissues needs special attention. There is no uniformity of scoring for the assessed tissues even for a given biomarker. The most commonly reported scoring system is based on the intensity of staining (weak, moderate and strong) (185, 234, 240, 271), a variety of scoring systems exists for the diagnostic biomarkers. Along

with staining intensity some researchers have also taken into account the proportions of positive cells to assess the level of expression of biomarkers in PDAC (159, 177, 237). Moreover, the researchers have divided the proportions of positive cells into various categorical groups (241, 253). Others have generated a quick score (233, 270), the sum score of both staining intensity, the proportion of cells (165) and an immunoreactive score (264) (for detail refer to Table 3.4 and 3.7). However, no consensus based immunostaining scoring system appears to be present for the evaluation of these diagnostic IHC based biomarkers. Interestingly, none of the included papers has used "Histoscore" which is an established and validated scoring system covering both the staining intensity and proportion of positive cells.

After scoring, a cut-off for positivity of a marker is a step ahead of finally calculating the sensitivity and specificity of the biomarker. Unfortunately, like the scoring of tissues there is no standardized cut-off for positivity. A binary categorisation of positive and negative staining for diagnosis is the most frequently reported cut-off (217, 236, 248, 266); nonetheless researchers have used a variety of cut-offs for classifying a tissue as either positive or negative. Similar to scoring of immunostaining, the cut-offs are based on intensity (172, 217, 234, 255) or proportions of positive cells (171, 232, 239) or a combination of both (159, 235, 240).

**The current review demonstrates weaknesses in the reporting of IHC biomarkers for diagnostic purposes.** The standardisation of scoring of immunostaining is of the utmost importance. A consensus based standardisation of scoring by the larger IHC research community might enable us to establish a uniform standard of scoring for reporting diagnostic IHC biomarkers. In addition, the IHC cut-off for positivity in tissue sections and cytology samples requires a uniform interpretative threshold for assigning the case into the positive or negative category. The standardisation of a uniform scoring method and cut-off could directly reflect on the clinical utility of IHC based diagnostic biomarkers.

**The strengths and limitations of our meta-analysis are as follows:** The strengths include a comprehensive search strategy, the first meta-analysis specifically designed to identify diagnostic IHC biomarkers, the quantification and ranking in different statistical models and identifying issues related to

scoring and cut-offs. The limitations include the weakness of primary studies; failure to use bivariate REM for all biomarkers; and failure to demonstrate any significant correlation with covariates that could potentially influence the diagnostic accuracy due to the limited number of articles for biomarkers. It is worth mentioning that some markers out of the ones reported in only one article are promising but they cannot be quantified and ranked in the statistical model.

**Conclusion:** meta-analyses on prognostic IHC biomarkers for PDAC have been published (284, 285). However, to my knowledge this is the first meta-analysis specifically designed to identify potential diagnostic IHC biomarkers for PDAC. The main findings from the review and meta-analysis are presented as follows. Firstly, high ranking potential candidates such as KOC, S100P, maspin, mesothelin and MUC1 were identified for potential clinical validation studies. Secondly, deficiencies in the reporting of diagnostic IHC biomarkers studies were outlined and suggestions made for future IHC biomarker studies. Thirdly, issues with the microscopic assessment of immunostaining and cut-offs were described and a need for standardised scoring system and reliable cut-off was suggested. Fourthly, the evidence of diagnostic accuracy of biomarkers for PDAC is promising for some of the biomarkers. However, none of the widely studied biomarker is completely sensitive to identify the disease at any stage of presentation. Similarly, none is exclusively specific to exclude PDAC. Therefore, a panel of highly sensitive and specific biomarkers is suggested. All these factors influence the translation of biomarkers to clinical practice and are some of the main reasons why we have a compendium of biomarkers (229, 230) but none is yet used in clinical practice to improve diagnosis of PDAC.

Therefore, a panel of better biomarkers in a sufficiently powered study with high methodological quality could help us to achieve appropriate sensitivity/specificity that could possibly be applied as an adjunct to cytology. Based on the available evidence, we might suggest that a panel of KOC, maspin, mesothelin, S100P and MUC1 could be validated by IHC in resection specimen and cytology samples. Afterwards, prospective diagnostic cohort studies to assess their diagnostic significance for possible clinical application are recommended. Finally, this systematic review and meta-analysis could serve as a reference approach for integrating the information and then prioritising the

investigation of candidate diagnostic biomarkers. Moreover, the “meta-analysis” could also serve as a prototype for quantifying and ranking diagnostic biomarkers in statistical models for other tumours.

**What this study adds:** This study identified biomarkers for future validation studies. It also emphasised the importance of good quality validation studies, a uniform scoring system, a standardised and reliable cut-off and panel approach for potentially diagnostic biomarkers.

**What next:** In the next two chapters, 4 and 5, I will now discuss the validation of identified biomarkers KOC, S100P, maspin, mesothelin and MUC1 in resection specimens from patients with PDAC. I will use Histoscore as potential optimal scoring systems and various diagnostic cut-offs. Afterwards in chapter 6 I will investigate cut-offs for observer agreement.

**Table 3.1: Characteristics of 64 included studies**

| | Reference | Biomarkers | Year | Country | Resection Specimens (RS)/Cytology | PDAC vs. NP OR CP OR Both | Study Design |
|---|---|---|---|---|---|---|---|
| **1.** | Abe et al | HMGI(Y) | 2000 | Japan/Italy | RS | Both | Case control |
| **2.** | Agarwal et al | 14 3-3 σ, CEACAM 6, Fascin, Maspin, Mesothelin | 2008 | USA | Both | Both | Cohort |
| **3.** | Akashi et al | RCAS 1 | 2003 | Japan | RS | CP | Case control |
| **4.** | Ali et al | SMAD4 | 2007 | USA | Cytology | NP | Cohort |
| **5.** | Argani et al (a) | Mesothelin | 2001 | USA | RS | NP | Cohort |
| **6.** | Argani et al (b) | PSCA | 2001 | USA | RS | NP | Cohort |
| **7.** | Awadallah et al | B7-H4, P53 | 2008 | USA | Both | Both | Cohort |
| **8.** | Baruch et al | Mesothelin | 2007 | USA | Cytology | NP | Cohort |
| **9.** | Bhardwaj et al | Maspin, MUC4, p53 | 2007 | USA | RS | CP | Case control |
| **10.** | Boltze et al | Fas L | 2002 | Germany | RS | Both | Case control |
| **11.** | Cao et al | Maspin | 2007 | USA | RS | NP | Cohort |
| **12.** | Chhieng et al | MUC1 | 2003 | USA | Cytology | CP | Cohort |
| **13.** | Chu et al | CDX2, CK7, CK17, CK20, MUC2 | 2005 | USA | RS | NA | Cohort |
| **14.** | Deng et al | S100P | 2008 | USA | Cytology | Both | Cohort |
| **15.** | Duxbury et al | CEACAM 6 | 2005 | USA | RS | NP | Cohort |
| **16.** | Erhuma | CD10 | 2007 | Germany | RS | CP | Case control |
| **17.** | Giorgadze et al | MUC1, MUC5AC, MUC6 | 2006 | USA | Cytology | CP | Cohort |
| **18.** | Glass et al | Mesothelin | 2011 | USA | RS | CP | Case control |
| **19.** | Haglund et al | CA19-9 | 1986 | Finland | RS | Both | Case control |
| **20.** | Hassan et al | Mesothelin | 2005 | USA | RS | Both | Case control |
| **21.** | Hosoda et al | CDX2, Chromogranin A, CK7, Synaptophysin | 2010 | Japan | Both | NA | Cohort |
| **22.** | Hosoda et al | BCl10 | 2013 | Japan | Both | NA | Cohort |
| **23.** | Ishimaru et al | P53 | 1996 | Japan | Cytology | CP | Cohort |
| **24.** | Jhala et al | Mesothelin, MUC4 | 2006 | USA | Both | Both | Cohort |
| **25.** | Karanjawala et al | Annexin A8, Claudin 18 | 2008 | USA | RS | Both | Cohort |
| **26.** | Kosarac et al | S100P, XIAP | 2011 | USA | Both | Both | Cohort |
| **27.** | Koshiba et al | COX2 | 1999 | Japan | RS | Both | Cohort |
| **28.** | La Rosa et al | Amylase, BCL10, CEH, Lipase, Trypsin | 2009 | Italy | RS | NA | Cohort |
| **29.** | Lee et al | P53 | 1993 | Australia | Cytology | Both | Cohort |
| **30.** | Li et al | P53 | 2002 | China | Cytology | CP | Cohort |
| **31.** | Ligato et al | KOC, S100A4 | 2008 | USA | Cytology | Both | Cohort |
| **32.** | Lim et al | Maspin | 2004 | Korea | RS | NP | Cohort |
| **33.** | Lin et al | pVHL, S100A4, S100A6, S100P | 2008 | USA | RS | NP | Cohort |
| **34.** | Lu et al | Fascin | 2004 | China/Switzerland | RS | NP | Cohort |
| **35.** | Maass et al | Maspin | 2001 | Germany | RS | NP | Cohort |
| **36.** | Maitra et al (a) | Fascin, HSP47 | 2002 | USA | RS | NP | Cohort |
| **37.** | Maitra et al (b) | COX2 | 2002 | USA | RS | NP | Cohort |
| **38.** | McCarthy et al | Mesothelin, PSCA | 2003 | USA/Nederland | Cytology | Both | Cohort |
| **39.** | Nash et al | Maspin | 2007 | USA | RS | CP | Case control |
| **40.** | Niijima et al | COX2 | 2002 | Japan | RS | Both | Case control |
| **41.** | Oh et al | Maspin | 2002 | Korea | RS | NA | Cohort |
| **42.** | Ohucida et al | S100A6 | 2005 | Japan | RS | NP | Cohort |
| **43.** | Okami et al | COX 2 | 1999 | Japan | RS | NP | Cohort |
| **44.** | Ordonez et al | Mesothelin | 2003 | USA | RS | NP | Cohort |

**Table 3.1: Characteristics of 64 included studies**

| | Reference | Biomarkers | Year | Country | Resection Specimens (RS)/Cytology | PDAC vs. NP OR CP OR Both | Study Design |
|---|---|---|---|---|---|---|---|
| **45.** | Rosty et al | S100A4 | 2002 | USA | RS | Both | Cohort |
| **46.** | Satomura | CA19-9, CA-50 | 1991 | Japan | RS | Both | Case control |
| **47.** | Shimoyama et al | ICAM1 | 1997 | Japan | RS | NP | Case control |
| **48.** | Stewart et al | P53 | 2000 | UK | Cytology | CP | Cohort |
| **49.** | Strickland et al | CEACAM 6 | 2009 | USA/UK | RS | NP | Cohort |
| **50.** | Sun et al | Bcl2 | 2002 | China | RS | Both | Case control |
| **51.** | Swartz et al | MUC4 | 2002 | USA | RS | NP | Cohort |
| **52.** | Tanaka et al | Claudin 18 | 2011 | Japan | RS | NP | Cohort |
| **53.** | Toll et al | KOC | 2009 | USA | Both | CP | Cohort |
| **54.** | Villanacci et al | P53 | 2009 | Italy | Cytology | CP | Cohort |
| **55.** | Wachter et al | KOC | 2011 | Switzerland | RS | Both | Cohort |
| **56.** | Wang et al | MUC1, MUC5AC | 2007 | China | Cytology | CP | Cohort |
| **57.** | Wang et al | HPK1 | 2009 | USA | RS | NP | Case control |
| **58.** | Wen-bin et al | Claudin 4, PSCA | 2008 | China | RS | Both | Case control |
| **59.** | Wente et al | CXCL16 | 2008 | Germany | RS | Both | Case control |
| **60.** | Xie et al | PAP | 2003 | Japan | RS | Both | Cohort |
| **61.** | Yantiss et al | KOC | 2005 | USA | RS | Both | Cohort |
| **62.** | Yantiss et al | KOC | 2008 | USA | Cytology | CP | Cohort |
| **63.** | Zapata et al | CA 19-9, CK19, SMAD4 | 2007 | USA | Cyotlogy | NP | Cohort |
| **64.** | Zhao et al | KOC | 2007 | USA | Cytology | Benign | Case control |

**Note:** *NA=data not available

**Table 3.2: Bivariate summary estimates of various markers in resection specimens**

| No | Biomarkers | % Sensitivity (95% CI) | % Specificity (95% CI) | Studies | Specimens |
|---|---|---|---|---|---|
| **Markers differentiating PDAC vs Normal Pancreas** | | | | | |
| **1.** | Maspin | 91.7 (91.6-91.7) | 97.5 (97.3-97.8) | 4 | 684 |
| **2.** | KOC (IMP3) | 85.2 (44.0-97.7) | 97.8 (59.5-99.9) | 3 | 162 |
| **3.** | Mesothelin | 92.2 (67.2-98.6) | 93.8 (74.6-98.8) | 4 | 231 |
| **4.** | Fascin | 82.8 (57.1-94.6) | 95.8 (65.2-99.7) | 3 | 194 |
| **5.** | COX2 | 88.3 (60.1-97.4) | 88.9 (31.6-99.3) | 3 | 162 |
| **6.** | S100P | 100 (94.9-100) | 100 (94.9-100) | 2 | 140 |
| **7.** | Claudin18 | 83.2 (78.7-87.1) | 98.5 (96.1-99.6) | 2 | 583 |
| **8.** | S100A6 | 98.9 (93.9-100) | 80.6 (71.4-87.9) | 2 | 187 |
| **9.** | S100A4 | 83.8 (75.8-89.9) | 90.0 (82.8-94.9) | 2 | 227 |
| **10.** | PSCA | 74.6 (66.5-81.7) | 95.7 (88.0-99.1) | 2 | 208 |
| **11.** | CEACAM6 | 91.2 (85.4-95.2) | 72.3 (57.4-84.4) | 2 | 194 |
| **Markers differentiating PDAC vs Chronic Pancreatitis** | | | | | |
| **1.** | KOC | 86.7 (79.6-91.5) | 97.6 (87.0-99.6) | 3 | 199 |
| **2.** | Mesothelin | 81.0 (70.1-88.6) | 90.4 (74.1-96.8) | 4 | 190 |
| **3.** | Maspin | 85.2 (71.3-92.9) | 67.9 (12.1-97.0) | 3 | 237 |
| **4.** | MUC4 | 82.0 (73.6-88.6) | 93.1 (83.3-98.1) | 2 | 169 |

**Table 3.3: Bivariate summary estimates of various markers in cytology specimens**

| No | Biomarkers | % Sensitivity (95% CI) | % Specificity (95% CI) | Studies | Specimens |
|---|---|---|---|---|---|
| **Markers differentiating PDAC vs Normal Pancreas** | | | | | |
| **1.** | KOC (IMP3) | 85.3 (68.9-95.0) | 100 (85.2-100) | 2 | 57 |
| **2.** | SMAD4 | 80.0 (71.9-86.6) | 100 (96.5-100) | 2 | 230 |
| **Markers differentiating PDAC vs Chronic Pancreatitis** | | | | | |
| **1.** | KOC | 84.1 (62.2-94.5) | 92.1 (68.3-98.5) | 3 | 77 |
| **2.** | Mesothelin | 63.8 (57.0-74.5) | 92.1 (88.0-99.1) | 4 | 191 |
| **3.** | MUC1 | 82.5 (59.3-93.8) | 77.3 (58.0-89.3) | 3 | 122 |
| **4.** | p53 | 66.2 (52.7-77.5) | 86.3 (73.9-94.4) | 6 | 291 |
| **5.** | S100P | 90.9 (80.0-97.0) | 90.9 (70.8-98.9) | 2 | 77 |
| **6.** | MUC5AC | 87.3 (77.3-94.0) | 58.8 (32.9-81.6) | 2 | 88 |

**Table 3.4: Outcome of studies evaluating candidate biomarkers in more than one study**

| Markers | Reference | Sample Size | Scoring of tissue | Threshold for positivity | Staining pattern | Cytology Sensitivity/ Specificity | Resection Specimens Sensitivity/ Specificity |
|---|---|---|---|---|---|---|---|
| **CEACAM 6** | Duxbury et al 2005 | 89 | 0 (<5% cells positive with no or weak staining), 1(weak), 2(moderate), 3(strong). | >5% with weak staining | Cytoplasmic, Membranous | NA* | Sensitivity 92%<br>Insufficient data for specificity |
| **CEACAM 6** | Agarwal et al 2008 | 61 | 0, 1(faint), 2(definite), 3(intense), 4 (very intense). | No threshold (any staining)<br>Positive/negative | Cytoplasmic | NA | 75%/82% (NP)<br>75%/28% (CP) |
| **CEACAM 6** | Strickland et al 2009 | 151 | 0-3+ and proportion of cells positive. | ≥10% positivity | Cytoplasmic, Membranous | NA | 96%/70% |
| **Claudin 18** | Karanjawala et al 2008 | 168 | Percentage of cells positive<br>Diffuse= >80% cells with 2 or 3staining.<br>0-3 staining | Only cores having ≥30% of carcinoma were scored.<br>Positive/negative | Cytoplasmic, Membranous | NA | 96%/96% (NP & CP) |
| **Claudin 18** | Tanaka et al 2011 | 156 | Only cores having ≥30% of carcinoma were scored<br>0, weak(1+, 1%-25%cells), , strong(3+, >80% cells) | Positive/negative | Cytoplasmic | NA | 70% /100% (NP) |
| **COX2** | Okami et al 1999 | 43 | 0-3 scale | Positive/negative | Cytoplasm | NA | 100%/100% (NP) |
| **COX2** | Koshiba et al 1999 | 50 | 0, 1(1%-33%), 2(34%-66%), 3 (67%-100%)<br>0-3 staining intensity<br>Score=Intensity* proportion | Score >4 positive | Cytoplasm, nuclear membrane | NA | 72% sensitivity<br>Data for specificity is not available |
| **COX2** | Niijima 2001 | 26 | 300 epithelial cells in major ductal lesions<br>Negative (<10%), +(10-30%), ++(30-50%), +++(>50%) | ≥10% cells | Cytoplasm | NA | 80%/100% (NP)<br>80%/100% (CP) |
| **COX2** | Maitra 2002 | 36 | Automated image analysis<br>Score= intensity*percentage of cells positive | Cut-off= >20% of cells positive | Cytoplasm | NA | 77%/60% (NP) |
| **Fascin** | Maitra 2002 | 57 | Focal (1-25%), Diffuse (26-100%)<br>0-4 Intensity | Positive and negative staining. | Cytoplasmic granular | NA | 95%/94% (NP) |
| **Fascin** | Lu et al 2004 | 21 | Positive/negative | Positive/negative | Cytoplasm | NA | 62%/100% |
| **Fascin** | Agarwal et al 2008 | 61 | 0, 1(faint), 2(definite), 3(intense), 4(very intense). | No threshold (any staining)<br>Positive/negative. | Cytoplasmic | NA | 84%/100% (NP)<br>84%/62% (CP) |
| **KOC (IMP3)** | Yantiss et al 2005 | 45 | 0%, <25%, 25–49%, 50–79%, 80–100%<br>0-3+ | >25% positive cells with 2+ staining | Cytoplasm | NA | 95%/100% (NP)<br>95%/100% (CP) |
| **KOC** | Zhao et al 2007 | 48 | 0-3+ | Positive/negative | Cytoplasm and/or membrane | 88%/100% (benign) | NA |
| **KOC** | Ligato et al 2008 | 44 | 0, Focal (>5%<25% cells), diffuse (>25% cells)<br>0-3+ | >5% cells with 1+ staining | Cytoplasm | 92%/95% (both) | NA |
| **KOC (IMP3)** | Yantiss et al 2008 | 25 | 0-3+ staining intensity* | Positive/negative | Cytoplasm | 92%/100% (CP) | NA |

**Table 3.4: Outcome of studies evaluating candidate biomarkers in more than one study**

| Markers | Reference | Sample Size | Scoring of tissue | Threshold for positivity | Staining pattern | Cytology Sensitivity/ Specificity | Resection Specimens Sensitivity/ Specificity |
|---|---|---|---|---|---|---|---|
| **KOC** | Toll et al 2009 | 21 | 0-3+ | >75% positive cells with 3+ staining | Cytoplasm | 71%/100% (CP and benign ductal epithelium) | 79%/100% (CP and benign ductal epithelium) |
| **KOC (IMP3)** | Wachter et al 2011 | 45 RS | 0, 1+(10% weak), 2+(10% moderate), 3+(>10% strong) | ≥10% positive cells with 2+ staining | Cytoplasm | 88%/100% (NP)<br>88%/95% (CP) | 85%/98% (NP)<br>85%/98% (CP) |
| **Maspin** | Maass et al 2001 | 27 | 0-3+ | 1+ staining | Cytoplasm | NA | 96%/100% (NP) |
| **Maspin** | Oh et al 2002 | 38 | Quick Score. 0, 1(<1/100tumor cells), 2(1/100-1/10), 3(1/10-1/3), 4(1/3-2/3), 5(>2/3)<br>0-3+ | Quick score of ≤5 | Cytoplasmic and/or nuclear | NA | 100% sensitivity |
| **Maspin** | Lim et al 2004 | 72 | Quick Score. 0, 1(<1/100tumor cells), 2(1/100-1/10), 3(1/10-1/3), 4(1/3-2/3), 5(>2/3)<br>0-3+ | Quick Score ≤5 | Cytoplasmic and/or nuclear | NA | 100%/100% |
| **Maspin** | Cao et al 2007 | 223 | Negative (<5%), focal (5%-50%), diffuse (>50%) | ≥5% positive cells | Cytoplasmic and/or nuclear | NA | 94%/100% |
| **Maspin** | Nash et al 2007 | 96 | 0-3+ | ≥5% positive cells and 2+ staining | Nuclear and/or Cytoplasmic staining | NA | 74%/96% (CP) |
| **Maspin** | Bhardwaj et al 2007 | 93 | Intensity=0, 1(weak), 2(strong)<br>Extent= <5%, 5%-50%, >50% | ≥5% cells | Cytoplasmic and Nuclear | NA | 90%/67% (CP) |
| **Maspin** | Agarwal et al 2008 | 61 RS | 0, 1(faint), 2(definite), 3(intense), 4(very intense). | No threshold (any staining)<br>Positive/negative | Cytoplasmic | NA | 88%/55% (NP)<br>88%/16% (CP) |
| **Mesothelin** | Argani et al 2001 | 60 | Focal (1-25%), Diffuse (26-100%)<br>0-4 Intensity | Positive and negative staining. | Granular Cytoplasmic | NA | 100%/86% (NP) |
| **Mesothelin** | Ordonez et al 2003 | 14 | ±<1%, +1 (1-25%), +2 (26-50%), +3 (51-75%), +4 (>75%) | Positive/negative | Cytoplasmic, membranous | NA | 86%/100% (NP) |
| **Mesothelin** | McCarthy et al 2003 | 30 | Positive (strong, diffuse Cytoplasmic staining)<br>Negative (no labelling over background) | Positive/negative | Cytoplasmic | 68%/91% (non-neoplastic cases) | NA |
| **Mesothelin** | Hassan et al 2005 | 53 | Negative (<1% staining), Focal (1-25%), extensively positive (>25%)<br>0-3+ staining intensity | ≥1% cells | Apical membranous and Cytoplasm | NA | 100%/100% (NP)<br>100%/94% (CP) |
| **Mesothelin** | Jhala et al 2006 | 40 (RS)<br>65 FNA | 0-4+ and percentage of cells | >5% cells and moderate staining | Cytoplasm | 62%/100% (CP) | 80%/90% (CP) |
| **Mesothelin** | Baruch et al 2007 | 28 | <10%, 10%-50%, >50%<br>0-3 intensity | Positive/negative | Cytoplasmic and/or membranous | 68%/90% (Non-neoplastic Pancreatic tissue) | NA |
| **Mesothelin** | Agarwal et al 2008 | 61 RS<br>27 EUS-FNA<br>56 Direct smears | 0, 1(faint), 2(definite), 3(intense), 4(very intense). | No threshold (any staining)<br>Positive/negative | Cytoplasmic | 68%/100% (EUS-FNA, Benign)<br>69%/96% (Direct smears, Benign) | 84%/100% (NP & CP) |

**Table 3.4: Outcome of studies evaluating candidate biomarkers in more than one study**

| Markers | Reference | Sample Size | Scoring of tissue | Threshold for positivity | Staining pattern | Cytology Sensitivity/ Specificity | Resection Specimens Sensitivity/ Specificity |
|---|---|---|---|---|---|---|---|
| **Mesothelin** | Glass et al 2011 | 25 | 0(0%-4% cells), 1(5%-15%), 2(16%-50%), 3(51%-100%)<br>0-3+ staining intensity | ≥5% cells | Cytoplasm and apical staining | NA | 78%/100% (CP) |
| **MUC1** | Chhieng et al 2003 | 35 | Any proportion of cells positive regardless of intensity | Any proportion of cells positive regardless of intensity | Membranous and cytoplasmic | 96%/94% (Reactive changes) | NA |
| **MUC1** | Giorgadze et al 2006 | 67 | Any proportion of cells positive regardless of intensity | Any proportion of cells positive regardless of intensity | Membranous and cytoplasmic | 83%/40% (NP)<br>83%/100% (CP) | 83%/0% (NP)<br>83%/66% (CP) |
| **MUC1** | Wang et al 2007 | 56 | Negative (<5%), Positive (>5%) | >5% positive cells | Cytomembranous and cytoplasmic | 78%/75% (CP & Benign) | NA |
| **MUC4** | Swartz et al 2002 | 40 | Focal (1-50%)<br>Strong (>50%) | Positive/negative | Cytoplasm | NA | 89%/100% (NP) |
| **MUC4** | Jhala et al 2006 | 40 (RS)<br>65 FNA | 0-4+ and percentage of cells | >5% cells and moderate staining | Cytoplasm | 91%/100% (CP) | 89%/99% (CP) |
| **MUC4** | Bhardwaj et al 2007 | 93 | Intensity=0, 1(weak), 2(strong)<br>Extent= <5%, 5%-50%, >50% | ≥5% cells | Cytoplasmic and membranous | NA | 77%/78% (CP) |
| **MUC5AC** | Giorgadze et al 2006 | 67 | Any proportion of cells positive regardless of intensity | Any proportion of cells positive regardless of intensity | Membranous and cytoplasmic | 97%/50% (CP) | 100%/66% (CP) |
| **MUC5AC** | Wang et al 2007 | 56 | Negative (<5%), Positive (>5%) | >5% positive cells | Membranous and cytoplasmic | 80%/56% (CP & Benign) | NA |
| **p53** | Lee et al 1993 | 71 | 0-3 staining intensity | Positive/negative | Nuclear | 65%/100% (NP)<br>65%/81% (CP) | NA |
| **p53** | Ishimaru 1996 | 28 | Any degree of nuclear staining | Positive/negative | Nuclear | 90%/100% (CP) | NA |
| **p53** | Stewart et al 2000 | 161 | Nuclear staining observed at (×250) | Positive/negative | Nuclear | 53%/100% (NP)<br>53%/98% (CP) | NA |
| **p53** | Li et al 2002 | 22 | One or more brown colour nuclei was considered positive | Positive/negative | Nuclear | 59%/100% (CP) | NA |
| **p53** | Bhardwaj et al 2007 | 93 | Intensity=0, 1(weak), 2(strong)<br>Extent= <5%, 5%-50%, >50% | ≥5% cells | Nuclear | NA | 60%/88% (CP) |
| **p53** | Awadallah et al 2008 | 15 | Staining= 0, 1(very light), 2(moderate to strong). Proportion of cells 0% (score0), 1%-50% (score1), >50%-100% (score2)<br>Sum score= Proportion+staining (range 0-4) | Positive/negative | Nuclear | 50%/80% (benign) | 83% sensitivity insufficient data to calculate specificity |
| **p53** | Villanacci et al 2009 | 24 | Any degree of nuclear staining | Positive/negative | Nuclear | 87%/78% (CP) | NA |
| **PSCA** | Argani et al 2001 | 60 | Focal (1-25%), Diffuse (26-100%)<br>0-4 Intensity | Positive and negative staining. | Cytoplasmic | NA | 60%/98% (NP) |

**Table 3.4: Outcome of studies evaluating candidate biomarkers in more than one study**

| Markers | Reference | Sample Size | Scoring of tissue | Threshold for positivity | Staining pattern | Cytology Sensitivity/ Specificity | Resection Specimens Sensitivity/ Specificity |
|---|---|---|---|---|---|---|---|
| **PSCA** | McCarthy et al 2003 | 30 | Positive (strong, diffuse Cytoplasmic staining) Negative (no labelling over background) | Positive/negative | Cytoplasmic | 84%/91% (non-neoplastic cases) | NA |
| **PSCA** | Wen-bin et al 2008 | 100 | Postive/negative | Positive/ negative | Cytoplasm | NA | 80%/80% (NP) 80%/67% (CP) |
| **S100A4** | Rosty et al 2002 | 66 | Focal (1-25%), Diffuse (26-100%) | Positive and negative staining | Cytoplasmic and nuclear | NA | 93%/100% (NP) 93%/100% (CP) |
| **S100A4** | Lin et al 2008 | 56 | 0 (<10%), +1 (11-25%), +2 (26-50%), +3 (50%-75%), +4 (≥75%) | >10% cells positive | Nuclear OR nuclear and cytoplasmic was considered positive | NA | 73%/80% (NP) |
| **S100A4** | Ligato et al 2008 | 44 | 0, Focal (>5%<25% cells), diffuse (>25% cells) 0-3+ | >5% positive cells with 1+ staining | Nuclear | 79%/95% (Both NP and CP) | NA |
| **S100A6** | Ohucida et al 2005 | 75 | 0, 1(<20%), 2(20%-75%), 3(>75%) 0-3 intensity Final score= intensity*proportion [0, Weak(1-3), Moderate (4-6), strong (>6) | Positive/negative | Nuclear and cytoplasmic staining | NA | 100%/81% (NP) |
| **S100A6** | Lin et al 2008 | 56 | 0 (<10%), +1 (11-25%), +2 (26-50%), +3 (50%-75%), +4 (≥75%) | >10% cells positive | Nuclear OR nuclear and cytoplasmic was considered positive | NA | 98%/80% (NP) |
| **S100P** | Lin et al 2008 | 56 | 0 (<10%), +1 (11-25%), +2 (26-50%), +3 (50%-75%), +4 (≥75%) | >10% of cells positive | Nuclear OR nuclear and cytoplasmic was considered positive | NA | 100%/100% (NP) |
| **S100P** | Deng et al 2008 | 52 | 0, +1(<25%), +2(26-50%), +3 (51-75%), +4 (>75%) 0-3 staining | Positive/Negative | Nuclear OR nuclear and cytoplasmic was considered positive | 100%/93% (Benign or reactive) All 6 suspicious cases by cytology stained for S100P and finally diagnosed as Adenocacrcinoma | NA |
| **S100P** | Kosarac et al 2011 | 14 RS 31 EUS-FNA | Negative (<10%), +1(10%-25%), +2(26%-75%), +3(>75%) 0-3 Intensity | ≥10% positive cells with +1 staining intensity. | Nuclear OR nuclear and cytoplasmic was considered positive | 78%/88% (benign) | 100%/100% (non-neoplastic including CP) |
| **SMAD4** | Zapata et al 2007 | 25 | Negative, Low positive (moderate staining), high positive (strong intense) | Positive/negative | Nuclear | 80%/100% (NP) | NA |
| **SMAD4** | Ali et al 2007 | 100 | Negative, 1+( Low intensity positivity, 2+ (high intensity positivity) | Positive/negative | Nuclear | 80%/100% (NP) | |

**Note:** *NA= not available

**Table 3.5: Immunohistochemistry details of all biomarkers progressing to meta-analysis.**

| Protein Assayed | Study | Primary Antibody | Antibody Dilution | Antigen retrieval |
|---|---|---|---|---|
| CEACAM6 | Agarwal 2008 | Santa Cruz | 1:200 | NA* |
| CEACAM6 | Strickland 2009 | In house anti-CEACAM6 Monoclonal | 10 µg/ml | NA |
| Claudin 18 | Karajawala 2008 | Rabbit monoclonal, clone ZMD 395, Invitrogen | 1:1000 | HIER, pH 6.0 |
| Claudin 18 | Tanaka 2011 | ZMD395, Zymed | 1:1000 | NA |
| COX2 | Okami 1999 | Rabbit polyclonal, Immuno-Biological Laboratories, Co. | 5 µg/ml | HIER |
| COX2 | Niijima 2001 | Polyclonal antibodies, Santa Cruz | 1:100 | NA |
| COX2 | Maitra (a) 2002 | Santa Cruz | 1:100 | HIER, pH 6.0 |
| Fascin | Lu 2002 | M3567, DAKO | 1:50 | NA |
| Fascin | Maitra (b) 2002 | Mouse monoclonal, 55K-2, DAKO | 1:500 | HIER, pH 6.0 |
| Fascin | Agarwal 2008 | DAKO | 1:1000 | NA |
| KOC | Yantiss 2005 | Mouse anti-L523S/KOC antibody | 2µg/ml | HIER, pH 6.0 |
| KOC | Yantiss 2008 | Mouse Monoclonal, L523S, Corixa Corporation | 2 µg/ml | HIER, pH 6.0 |
| KOC | Ligato 2008 | Monoclonal, L523S, clone 69.1, DAKO | NA | NA |
| KOC | Toll 2009 | Monoclonal, KOC/L523S, clone 69.1, DAKO | 1:500 | HIER, pH 9.5 |
| KOC | Wachter 2011 | Mouse monoclonal, clone 69.1, DAKO | 1:500 | HIER, pH 6.0 |
| Maspin | Maass 2001 | Monoclonal, BD Biosciences PharMingen | 1:50 | HIER, pH 6.0 |
| Maspin | Lim 2004 | Monoclonal, BD Biosciences Pharmingen | 1:3,000 | HIER, pH 6.0 |
| Maspin | Nash 2007 | Monoclonal, clone G167-70, PharMingen | 1:800 | HIER, pH 6.1 |
| Maspin | Cao 2007 | Monoclonal, BD Biosciences Pharmingen | 1:50 | HIER, pH 6.0 |
| Maspin | Bhardwaj 2007 | Monoclonal, BD Biosciences Pharmingen | 1:800 | HIER, pH 6.1 |
| Maspin | Agarwal 2008 | Novocastra | 1:20 | NA |
| Mesothelin | Argani (a) 2001 | Mouse monoclonal, clone 5B2, Novocastra | 1:20 | HIER, Sodium citrate buffer |
| Mesothelin | McCarthy 2003 | Mouse monoclonal, clone 5B2, Novocastra | 1:20 | HIER, Sodium citrate buffer |
| Mesothelin | Ordonez 2003 | Mouse monoclonal, clone 5B2; Novocastra. | 1:30 | HIER, Tris-EDTA buffer, pH 8.0 |
| Mesothelin | Hassan 2005 | Mouse monoclonal, 5B2 | 1:20 | HIER, pH 6.0 |
| Mesothelin | Jhala 2006 | 5B2 (MESO1), Lab Vision | 1:20 | HIER |
| Mesothelin | Baruch 2007 | Novocastra | 1:30 | NA |
| Mesothelin | Agarwal 2008 | NovaCastra | 1:10 | NA |
| Mesothelin | Glass 2011 | Clone, 5B2 Vector | 1:15 | NA |
| MUC1 | Chhieng 2003 | Clone DF3, Novocastra | 1:200 | HIER, pH 6.0 |
| MUC1 | Giorgadze 2005 | Mouse monoclonal, clone Ma695, Novocastra | 1:75 | HIER, pH 6.0 |
| MUC1 | Wang 2007 | Mouse monoclonal, clone Mh1, Labvision | 1:100 | HIER, pH 6.0 |
| MUC4 | Swartz 2002 | Mouse monoclonal, clone 8G7 | 1:3,0000 | HIER, pH 6.0 |
| MUC4 | Jhala 2006 | 8G7, S.K.B. | 1:3,000 | HIER |
| MUC4 | Bhardwaj 2007 | Zymed Laboratories | 1:200 | HIER, pH 6.1 |
| MUC5AC | Giorgadze 2005 | Clone CLH2, Novocastra | 1:100 | NA |
| MUC5AC | Wang 2007 | Clone 45M1, Labvision | 1:100 | NA |
| p53 | Lee 1993 | Polyclonal rabbit, NCL-p53-CMl Novocastra | N/A | NA |
| p53 | Ishimaru 1999 | Mouse monoclonal, Novocastra | 1:30 | NA |
| p53 | Li 2002 | The En Vision method, Shanghai Zhong Da Corporation | NA | NA |
| P53 | Stewart 2002 | Mouse monoclonal, DO-7, DAKO | 1:2,000 | HIER, EDTA at pH 8 |
| P53 | Bhardwaj 2007 | DAKO | 1:50 | HIER, pH 6.1 |
| p53 | Villanacci 2009 | Mouse monoclonal, clone DO-7, Thermo Scientific. | 1:50 | HIER |

**Table3.5: Immunohistochemistry details of all biomarkers progressing to meta- analysis.**

| Protein Assayed | Study | Primary Antibody | Antibody Dilution | Antigen retrieval |
|---|---|---|---|---|
| P53 | Awadallah 2008 | Monoclonal, DO-7, DAKO | 1 µg/mL | NA |
| PSCA | Argani (b) 2001 | Mouse monoclonal, Clone 1G8, obtained from R. E. R. | 1:200 | HIER, Sodium citrate buffer |
| PSCA | McCarthy 2003 | Monoclonal, clone 1G8, obtained from, Department Of Urology, University of California, Los Angeles, CA. | 1:200 | HIER, Sodium citrate buffer |
| PSCA | Wen-bin 2008 (Article in Chinese) | N/A | N/A | N/A |
| S100A4 | Rosty 2002 | Rabbit polyclonal, DAKO | 2 µg/ml | HIER, pH 6.0 |
| S100A4 | Ligato 2008 | Polyclonal, DAKO | NA | NA |
| S100A4 | Lin 2008 | Rabbit polyclonal, Neo Markers | 1:200 | HIER, pH 6.0 |
| S100A6 | Ohucida 2005 | Mouse monoclonal, Sigma | N/A | NA |
| S100A6 | Lin 2008 | Mouse monoclonal, Clone: CACY-100, Sigma-Aldrich | 1:1000 | Proteinase K |
| S100P | Lin 2008 | Mouse monoclonal, Clone 16, BD Biosciences Pharmingen | 1:100 | Proteinase K |
| S100P | Deng 2008 | Mouse monoclonal, BD Biosciences Pharmingen | 1:100 | Proteinase K |
| S100P | Kosarac 2010 | Monoclonal, clone 16/S100P, BD Biosciences Pharmingen | 1:100 | NA |
| SMAD4 | Zapata 2007 | Novacastra | 1:40 | HIER, pH 6.0 |
| SMAD4 | Ali 2007 | Novacastra | 1:40 | HIER, pH 6.0 |

**Note:** *NA= not available

**Table 3.6: QUADAS tool for assessing the methodological qualities of included studies.**

| No | Item | Description |
|---|---|---|
| **1** | Representative spectrum and selection criteria? | Yes, if the paper fulfils the following criteria: the selection of patients had been consecutive; the male to female ratio and the age distribution of patients; and the number of samples with pancreatic cancer and non-neoplastic disease reported. |
| **2** | Acceptable reference standard*? | Yes, if histopathology and/or death from pancreatic cancer was used as a reference standard. |
| **3** | Partial verification avoided? | Yes, if all the samples included in the paper received verification by either histopathology and/or death from pancreatic cancer. |
| **4** | Differential verification avoided? | Yes if the same reference standard (either histopathology OR death from pancreatic cancer but not both) was used for all samples included in the paper. |
| **5** | Incorporation avoided? | Yes, if the reference standard was not part of the index test**. |
| **6** | Details of index test? | Yes, if the details of immunohistochemistry (Antigen retrieval method, Primary antibody (clone) and antibody concentration) have been discussed in sufficient details to allow replication of test. |
| **7** | Reference standard results blinded? | Yes, if the histopathology results for pancreatic cancer and non-neoplastic tissue were interpreted without knowledge of the results of the index test. |
| **8** | Index test results blinded? | Yes, if the index test results were interpreted without knowledge of the histopathology results for pancreatic cancer and non- neoplastic tissue. |
| **9** | Uninterpretable results reported? | Yes, if uninterpretable or intermediate test results were reported. In addition, we would call it "Yes", if there was no uninterpretable or intermediate test result. |
| **10** | Withdrawals explained? | Yes, if withdrawals from the study were explained or there was no withdrawal from the included participants. |

**Note:** *Reference standard is a test that correctly diagnoses the pancreatic cancer (Histopathology or death from pancreatic cancer). **Index tests are the IHC biomarkers. Note: Item 6 has been added purposefully to QUADAS tool (Both QUADAS and Cochrane database suggest adding items necessary for a particular question).

**Table 3.7: Outcome of studies evaluating candidate biomarkers with reasonable sensitivity and/or specificity reported in one study**

| Markers | Reference | Sample Size | Scoring of tissue | Threshold for positivity | Staining pattern | Cytology Sensitivity/ Specificity | Resection Specimens Sensitivity/ Specificity |
|---|---|---|---|---|---|---|---|
| **14 3-3 σ** | Agarwal et al 2008 | 61 RS<br>27 EUS-FNA<br>56 Direct smears | 0, 1(faint), 2(definite), 3(intense), 4(very intense). | No threshold (any staining) Positive/negative | membranous | 100%/80% (EUS-FNA, Benign)<br>96%/78% (Direct smears, Benign) | 88%/100% (NP & CP) |
| **Annexin A8** | Karanjawala et al 2008 | 154 | Percentage of cells positive<br>Focal staining<br>Diffuse= >80% cells with 2 or 3, staining 0-3 | Only cores having ≥30% of carcinoma were scored.<br>Focal or diffuse positive. | Nuclear, Cytoplasmic | NA | 98%/88% (NP & CP) |
| **Bcl2** | Sun et al 2002 | 150 | 0 (<10%), Positive (≥10%) | ≥10% of cells positive | Cytomembrane and /or cytoplasmic | NA | 72%/95% (NP)<br>72%/90% (CP) |
| **B7-H4** | Awadallah et al 2008 | 25 | Staining= 0, 1(very light), 2(moderate to strong). Proportion of cells 0% (score0), 1%-50% (score1), >50%-100% (score2)<br>Sum score= Proportion+staining (range 0-4) | Positive/negative | Membranous/ cytoplasmic | 90%/80% (CP) | 92% sensitivity, insufficient data to calculate specificity |
| **CK 19** | Zapata et al 2007 | 25 | Negative, Low positive (moderate staining), high positive (strong intense) | Positive/negative | Cytoplasm | 100%/100% (NP) | NA |
| **Claudin 4** | Wen-bin et al 2008 | 100 | Positive/negative | Positive/ negative | Not Available | | 88%/70% (NP)<br>88%/67% (CP) |
| **CXCL16** | Wente et al 2008 | 31 | 0-3+ staining | Positive/negative | Cytomembranous | | 85%/45% (NP)<br>85%/0% (CP) |
| **Fas-L** | Boltze et al 2002 | 126 | 0-3+ staining intensity<br>0, 1(<10%cells), 2(10%-50%cells), 3(51%-80%cells), 4(>80%cells)<br>Immunoreactive score (0 to 12) = Intensity×Proportion of cells | .<br>Cut-off= Score>2 | Cytoplasm | NA | 100%/100% (NP)<br>100%/90% (CP) |
| **HMGI(Y)** | Abe et al 2000 | 21 | 1000 cells in 3 random fields. | Positivity= ≥20% of cells | Nuclear | NA | 100%/100% (NP)<br>100%/100% (CP) |
| **HSP 47** | Maitra 2002 | 57 | Focal (1-25%), Diffuse (26-100%)<br>0-4 Intensity | Positive and negative staining. | PDAC desmoplastic stroma vs. stroma of non-neoplastic pancreatic parenchyma. | NA | 100%/88% (NP) |
| **ICAM1** | Shimoyama et al 1997 | 19 | 0-3 staining<br>0, 1(<33%), 2(34%-66%), 3(<67%) | Positive/negative | Cytoplasm | NA* | 100%/100% (NP) |
| **PAP** | Xie et al 2003 | 87 | 0(<10%), 1(10%-30%), 2(30%-50%), 3(>50%). | Cut-Off= 2 | Cytoplasm | NA | 79%/100% (NP)<br>79%/81% (CP) |
| **RCAS 1** | Akashi et al 2003 | 25 | <5%= negative, >5%=positive | Cut-Off= >5% cells positive | Cytoplasm | NA | 100%/60% (CP) |
| **XIAP** | Kosarac et al 2011 | 14 RS<br>31 EUS-FNA | 0-3 Intensity<br>Negative (<10%), +1(10%-25%), +2(26%-75%), +3(>75%) | ≥10% positive cells with +1 intensity, | Granular Cytoplasmic | 83%/50% (benign) | 100%/100% (non-neoplastic including CP) |

**Note:** *NA= not applicable

**Table 3.8: Outcome of studies evaluating negative biomarkers for PDAC (biomarkers expressed in normal ducts but not in PDAC)**

| Markers | Reference | Sample Size | Scoring of tissue | Threshold for positivity | Staining pattern | Cytology Sensitivity/ Specificity | Resection Specimens Sensitivity/ Specificity |
|---|---|---|---|---|---|---|---|
| **Amylase** | La Rosa et al 2009 | 14 | Weak (1+) & <50% of tumour cells; intense (3+) & diffuse reaction; moderate (2+) showing features between scores 1+ and 3+. | Positive/Negative | N/A | | 0%/no data to calculate specificity but positive in normal acini. |
| **BCL10** | La Rosa et al 2009 | 14 | Weak (1+) & <50% of tumour cells; intense (3+) & diffuse reaction; moderate (2+) showing features between scores 1+ and 3+. | Positive/Negative | Granular Cytoplasm | | 0%/ no data to calculate specificity but authors state and diagram shows positivity of acinar cells of normal pancrease, while ductal and islet cells were negative. |
| **BCL10** | Hosoda et al 2013 | 23 (RS), 18 (EUS-FNA) | 4-tiered scoring system: negative; 1+, faint and focal staining (<50% of the total area); 2+, faint and diffuse staining (>50% of the total area) or strong but focal staining (< 50% of the total area); and 3+, strong and diffuse staining (>50% of the total area). | Tumour cells showing moderate or greater intensity in 50% of the area (2+, or 3+). | Granular Cytoplasm | 0%/ no data to calculate specificity | 0%/ no data to calculate specificity but authors state and diagram shows positivity of acinar cells of normal pancrease, while ductal and islet cells were negative. |
| **CDX2** | Hosoda et al 2010 | 25 | 0: 0–5%, faint staining; 1+: 6–75%, variable staining intensity; and 2+: 76–100%, mostly strong staining. | 2+ | Nuclear | 4% /no data to calculate specificity but authors state and diagram shows the expression of CDX2 in normal cells | 0% /no data to calculate specificity but authors state and diagram shows the expression of CDX2 in normal ducts |
| **CEH (carboxyl ester hydrolase)** | La Rosa et al 2009 | 14 | Weak (1+) & <50% of tumour cells; intense (3+) & diffuse reaction; moderate (2+) showing features between scores 1+ and 3+. | Positive/Negative | N/A | | 0%/no data to calculate specificity but positive in normal acini. |
| **Chromogranin A** | Hosoda et al 2010 | 25 | 0: 0–5%, faint staining; 1+: 6–75%, variable staining intensity; and 2+: 76–100%, mostly strong staining. | 2+ | Cytoplasmic | 4%/ no data to calculate specificity | 0% /no data to calculate specificity |
| **HPK1** | Wang et al 2009 | 79 | Negative= No or only weak focal staining (<5% of cells); Positive= Strong staining (≥5% of the cells. | Strong cytoplasmic staining for HPK1 (≥5% of the cells) | Cytoplasmic | | 4%/0% (PDAC vs Benign pancreatic tissue) |
| **Lipase** | La Rosa et al 2009 | 14 | Weak (1+) & <50% of tumour cells; intense (3+) & diffuse reaction; moderate (2+) showing features between scores 1+ and 3+. | Positive/Negative | N/A | | 0%/no data to calculate specificity but positive in normal acini. |
| **MUC2** | Chu et al 2005 | 46 | Proportion of positive cells | >5% Positive cells | Cytoplasmic and membranous | | 2% /no data to calculate specificity. |

**Table 3.8: Outcome of studies evaluating negative biomarkers for PDAC (biomarkers expressed in normal ducts but not in PDAC)**

| Markers | Reference | Sample Size | Scoring of tissue | Threshold for positivity | Staining pattern | Cytology Sensitivity/ Specificity | Resection Specimens Sensitivity/ Specificity |
|---|---|---|---|---|---|---|---|
| **Synaptophysin** | Hosoda et al 2010 | 25 | 0: 0–5%, faint staining; 1+: 6–75%, variable staining intensity; and 2+: 76–100%, mostly strong staining. | 2+ | NA* | 0%/no data to calculate specificity | 0% /no data to calculate specificity |
| **Trypsin** | La Rosa et al 2009 | 14 | Weak (1+) & <50% of tumour cells; intense (3+) & diffuse reaction; moderate (2+) showing features between scores 1+ and 3+. | Positive/Negative | Cytoplasm | | 0%/no data to calculate specificity but positive in normal acini. |
| **von Hippel-Lindau gene product (pVHL)** | Lin et al 2008 | 56 | 0 (<10%), +1 (11-25%), +2 (26-50%), +3 (50%-75%), +4 (≥75%) | >10% positive cells | Granular cytoplasmic/membranous staining | | 0% /0% |

**Note:** *NA= not available

**Table 3.9: Outcome of studies evaluating biomarkers with low sensitivity and/or specificity**

| Markers | Reference | Sample Size | Scoring of tissue | Threshold for positivity | Staining pattern | Cytology Sensitivity/ Specificity | Resection Specimens Sensitivity/ Specificity |
|---|---|---|---|---|---|---|---|
| **CA19-9** | Satomura et al 1991 | 80 | (-); <1/3 epithelial cells stained, (+); >1/3 but <2/3 of the epithelial cells stained, (+ +); >2/3 epithelial cells stained, (+ + +). | <1/3 epithelial cells stained, (+) | Apical and cytoplasmic | | 83% /24% (NP)<br>83% /24% (CP) |
| **CA19-9** | Haglund et al 1986 | 112 | Positive/negative | Positive/negative | Apical and cytoplasmic | | 80% / 21% (NP)<br>80% / 4% (CP) |
| **CA-50** | Satomura 1991 | 80 | (-); <1/3 epithelial cells stained, (+); >1/3 but <2/3 of the epithelial cells stained, (+ +); >2/3 epithelial cells stained, (+ + +). | <1/3 epithelial cells stained, (+) | Apical and cytoplasmic | | 98%/ 8% (NP)<br>98%/ 4% (CP) |
| **CD10** | Erhuma et al 2007 | 30 | Proportion of positive cells. | >25% stained cells. | Membranous and cytoplasmic | | 25% /50% (CP) |
| **CDX2** | Chu et al 2005 | 46 | Proportion of positive cells | >5% Positive cells | Cytoplasmic and membranous | | 21% / no data to calculate specificity. |
| **CK7** | Hosoda et al 2010 | 25 | 0: 0–5%, faint staining; 1+: 6–75%, variable staining intensity; and 2+: 76–100%, mostly strong staining. | 2+ | Cytoplasmic | 92%/no data to calculate specificity but author state and diagram shows the expression of CK7 in normal cells | 88% /no data to calculate specificity but authors state and diagram shows the expression of CK7 in normal ducts |
| **CK7** | Chu et al 2005 | 46 | Proportion of positive cells. | >5% Positive cells | Cytoplasmic and membranous | | 98%/no data to calculate specificity. |
| **CK 17** | Chu et al 2005 | 46 | Proportion of positive cells. | >5% Positive cells | Cytoplasmic and membranous | | 83% / no data to calculate specificity. |
| **CK20** | Chu et al 2005 | 46 | Proportion of positive cells. | >5% Positive cells | Cytoplasmic and membranous | | 46% / no data to calculate specificity. |
| **MUC6** | Giorgadze et al 2006 | 67 | Any proportion of cells positive regardless of intensity | Any proportion of cells positive regardless of intensity | Membranous and cytoplasmic | 35%/0% (NP)<br>35%/0% (CP) | 37%/60% (NP)<br>37%/100% (CP) |

**Note:** *NA= not available

**Figure 3.1: Prisma flow chart of selection of papers for systematic review and meta-analysis.**

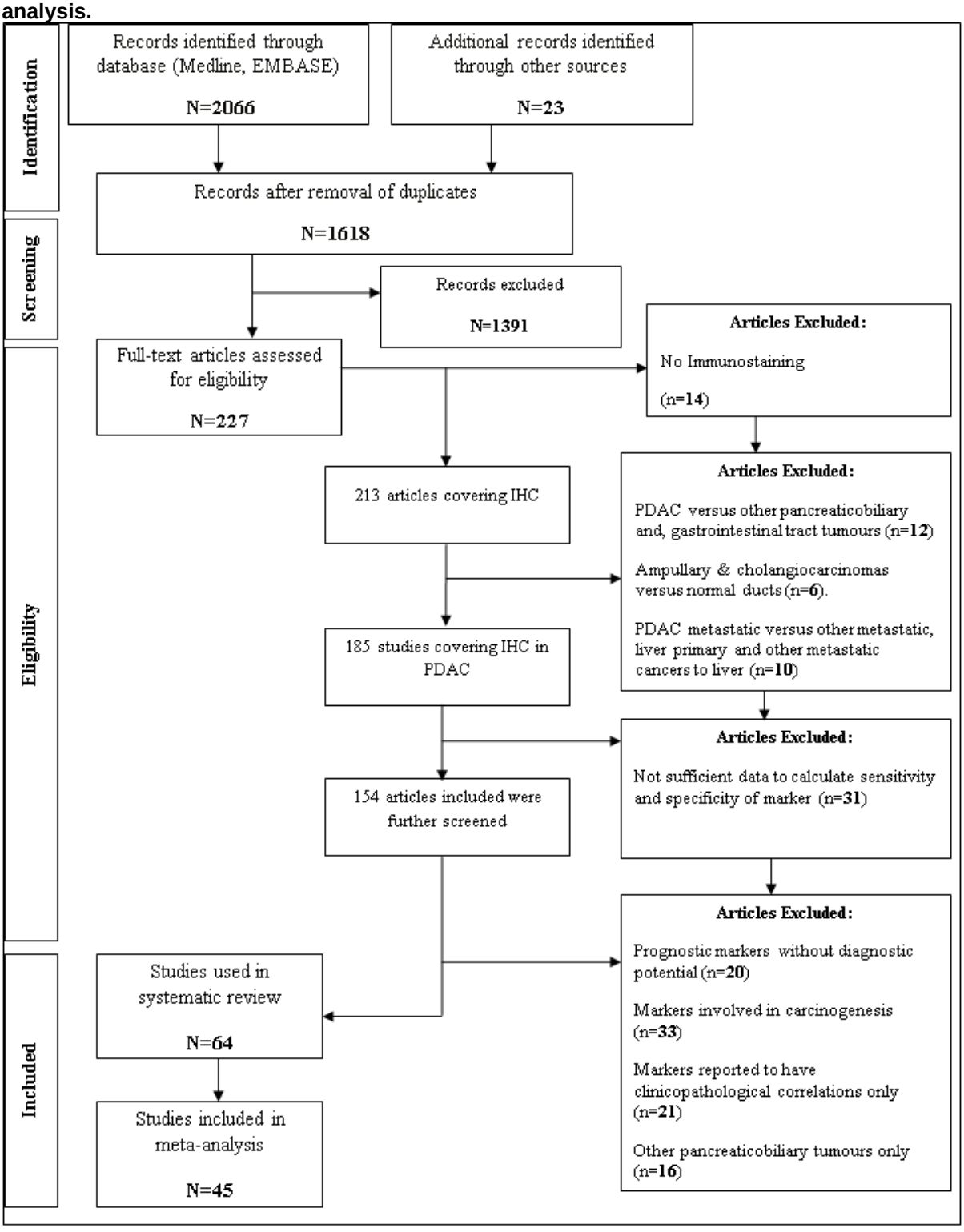


**Figure 3.2: Coupled Forest plots of sensitivity and specificity of included studies for various markers in resection specimens arranged by biomarkers.**

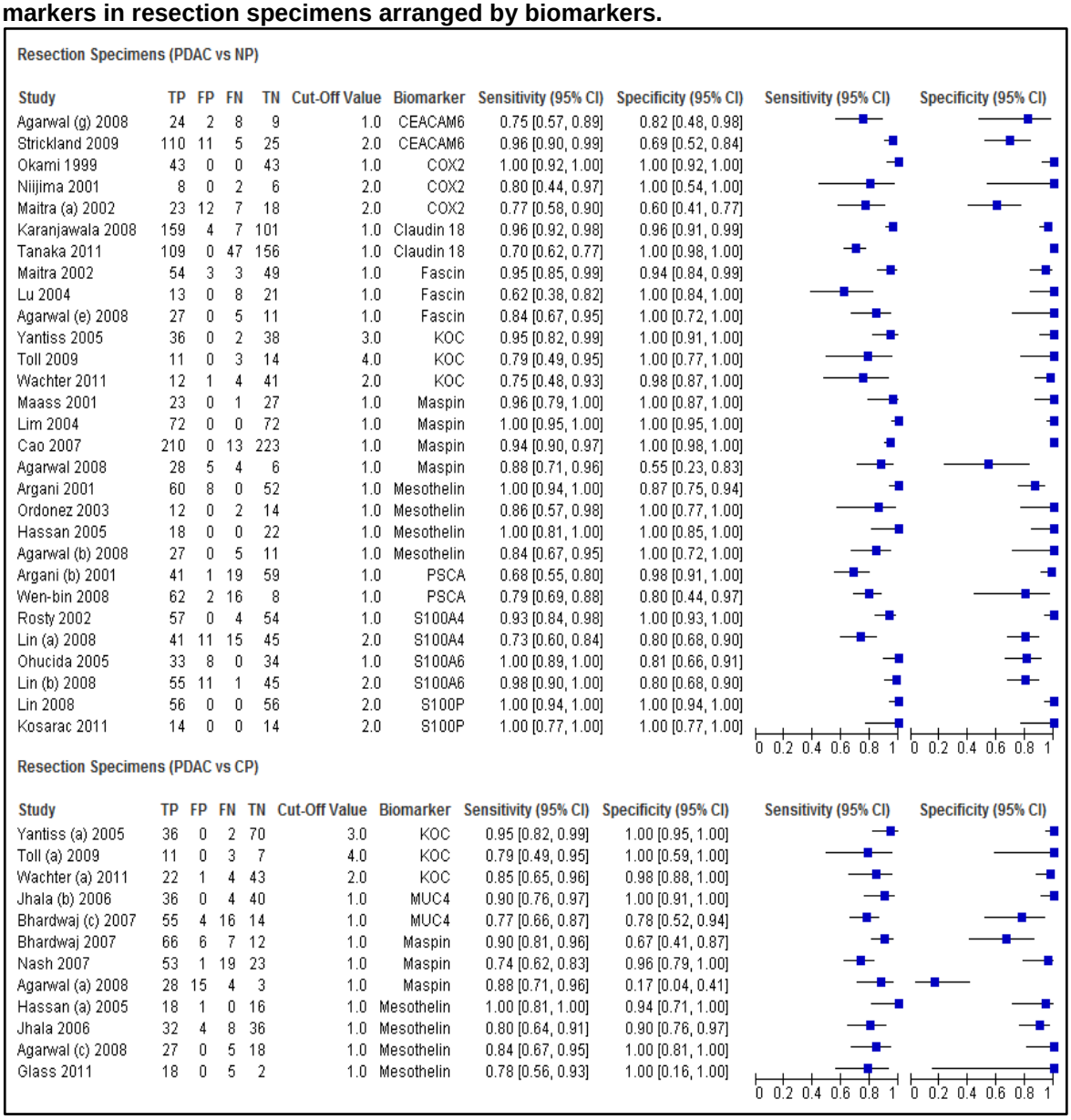



**Figure Legend:** The squares represent the sensitivity and specificity for a given candidate biomarker, the black line its confidence interval. Studies are grouped by the type of biomarkers. Cut-Off values based on proportion of positive cells (1= 1%-9%, 2= 10%-24%, 3= 25%-49%, 4= ≥50%) for biomarkers are also shown in the Forest plot. Some papers reported sensitivity & specificity data for more than one biomarker and/or for more than one scenario and they were included more than once in Forest plot.

Abbreviations: PDAC, Pancreatic ductal adenocarcinoma; NP, Normal Pancreatic duct; CP, Chronic Pancreatitis; TP, true positive; FP, false positive; FN, false negative; TN, true negative.

**Figure 3.3: Combined SROC curves of various biomarkers studied in resection specimens.**

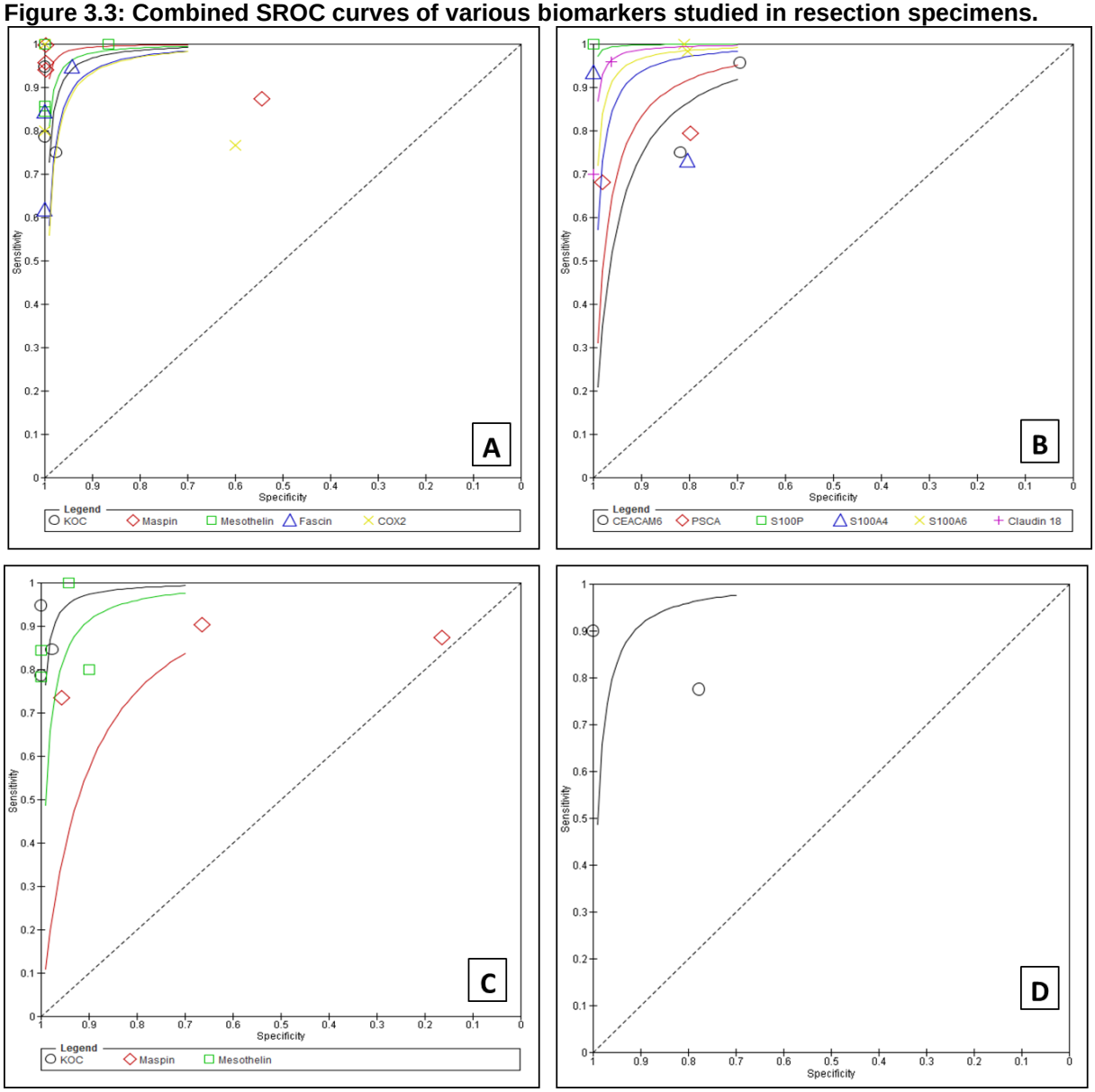



**Figure Legend:** Each SROC curve consists of a summary line of diagnostic accuracy resulting from various studies for a biomarker and individual study points around the summary line. In addition, SROC curves from various biomarkers have been compared in a combined SROC curve for each scenario **A)** PDAC vs normal pancreas- five biomarkers (maspin, mesothelin, KOC, fascin and COX2) each reported in at least 3 studies **B)** PDAC vs normal pancreas- six biomarkers (S100P, claudin18, S100A6, S100A4, PSCA and CEACAM6) each reported in 2 studies **C)** PDAC vs chronic pancreatitis-three biomarkers (KOC, mesothelin and maspin) each reported in at least 3 studies **D)** PDAC vs chronic pancreatitis-One biomarker (MUC4) reported in 2 studies.

**Figure 3.4: Coupled Forest plots of sensitivity and specificity of included studies for various markers in cytology specimens arranged by biomarkers.**

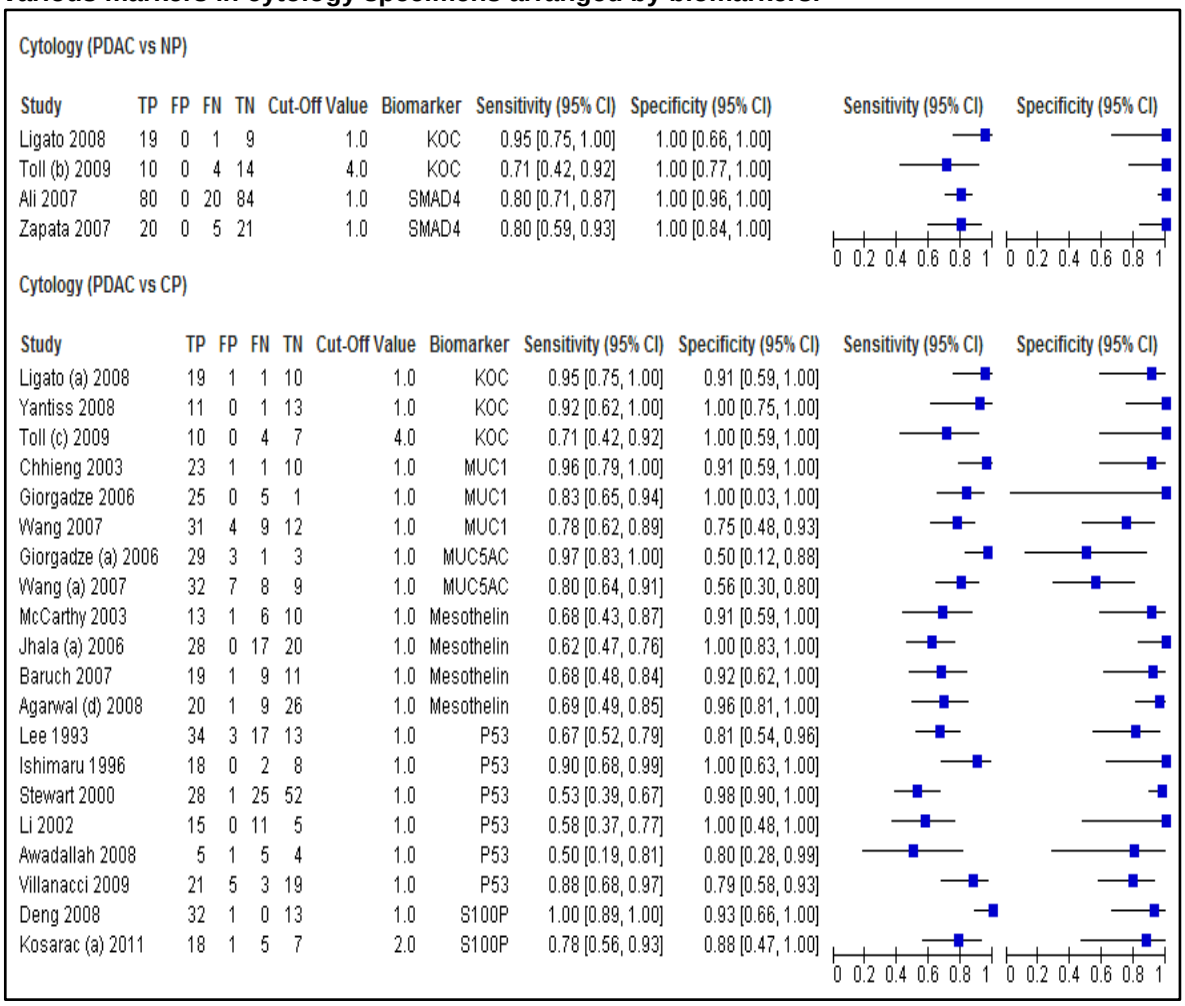



**Figure Legend:** The squares represent the sensitivity and specificity for a given candidate biomarker, the black line its confidence interval. Studies are grouped by the type of biomarkers. Cut-Off values based on proportion of positive cells (1= 1%-9%, 2= 10%-24%, 3= 25%-49%, 4= ≥50%) for biomarkers are also shown in the Forest plot. Some papers reported sensitivity & specificity data for more than one biomarker and/or for more than one scenario and they were included more than once in Forest plot.

***Abbreviations:*** PDAC, Pancreatic ductal adenocarcinoma; NP, Normal Pancreatic duct; CP, Chronic Pancreatitis; TP, true positive; FP, false positive; FN, false negative; TN, true negative.

**Figure 3.5: Combined SROC curves of various biomarkers studied in cytology specimens.**

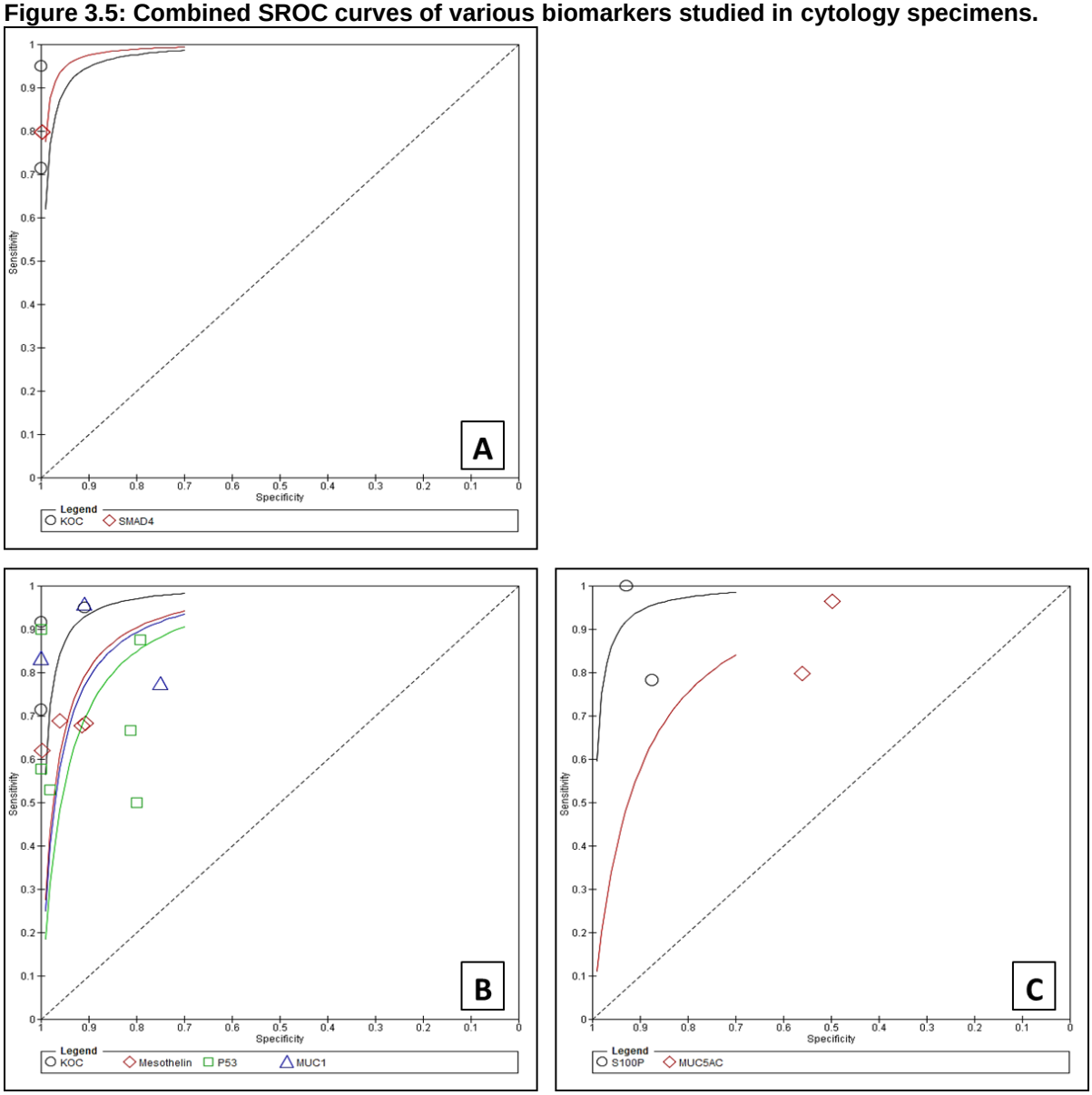



**Figure legend:** Each SROC curve consists of a summary line of diagnostic accuracy resulting from various studies for a biomarker and individual study points around the summary line. In addition, SROC curves from various biomarkers have been compared in a combined SROC curve for each scenario **A)** PDAC vs normal pancreas-Two biomarkers (KOC and SMAD4) each reported in 2 studies **B)** PDAC vs chronic pancreatitis-Four biomarkers (KOC, mesothelin, MUC1 and p53) each reported in at least 3 studies **C)** PDAC vs chronic pancreatitis-Two biomarkers (S100P and MUC5AC) each reported in 2 studies.

**Figure 3.6: Methodological quality graph: review authors' judgements about each methodological quality item presented as percentages across all included studies.**

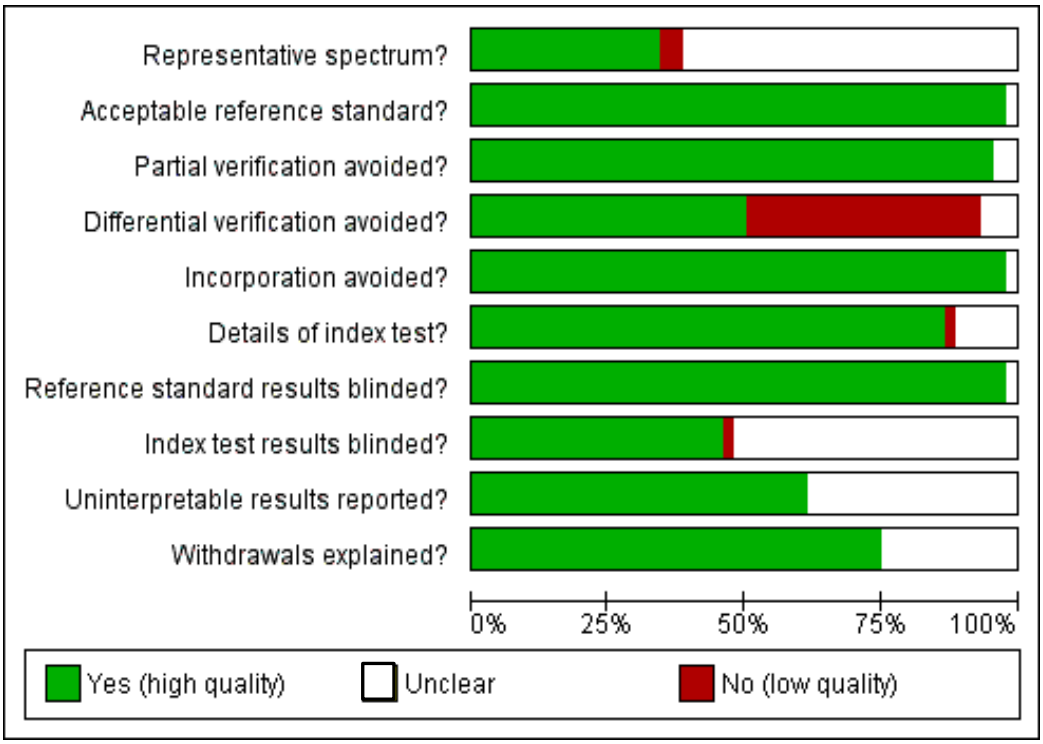
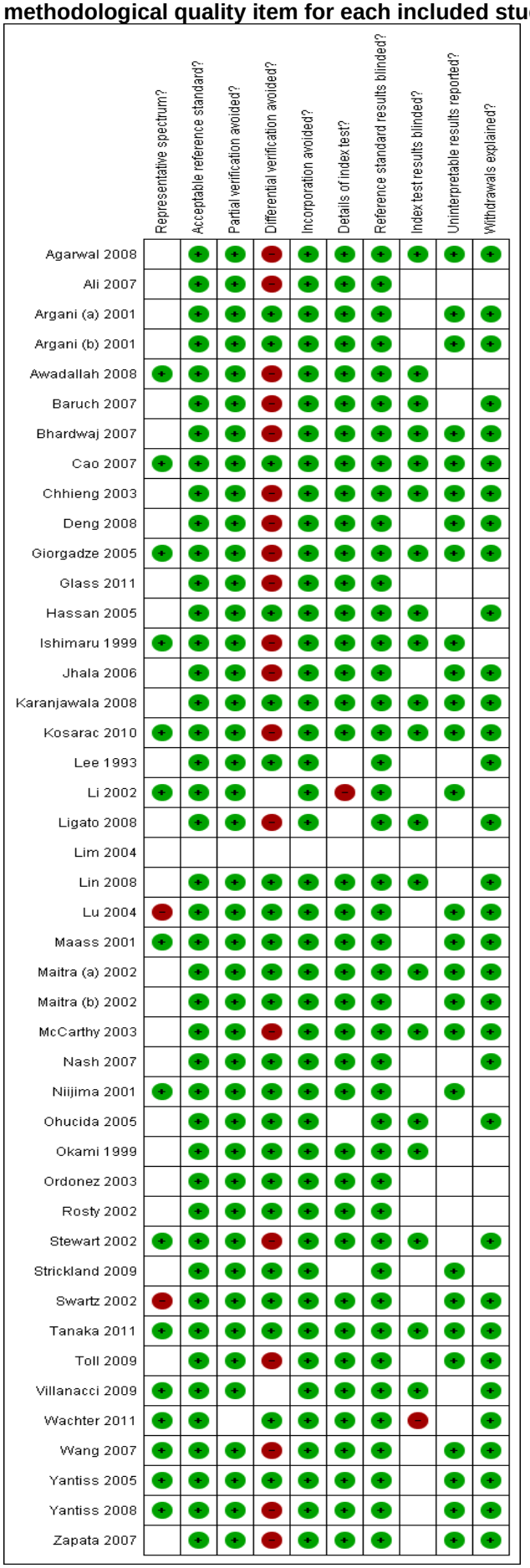


**Figure 3.7: Methodological quality summary: review authors' judgements about each methodological quality item for each included study.**



# 4 Expression of KOC, S100P, mesothelin, and MUC1 in pancreatico-biliary adenocarcinomas: development and utility of a potential diagnostic immunohistochemistry panel.

## 4.1 Chapter Summary

Pancreatico-biliary adenocarcinomas (PBA) have a poor prognosis. Diagnosis is usually achieved by imaging and/or endoscopy with confirmatory cytology. Cytological interpretation can be difficult especially in the setting of chronic pancreatitis/cholangitis. Immunohistochemistry (IHC) biomarkers could act as an adjunct to cytology to improve the diagnosis. Thus, we performed a meta-analysis of diagnostic IHC biomarkers for PDAC and selected KOC, S100P, mesothelin and MUC1 for further validation in PBA resection specimens.

Tissue microarrays (TMAs) containing tumour and normal cores in a ratio of 3:2, from 99 surgically resected PBA patients, were used. IHC was performed on the TMAs using an automated platform using antibodies against KOC, S100P, mesothelin and MUC1. Tissue cores were scored for staining intensity and the proportion of tissue stained using a Histoscore method (range, 0–300). Sensitivity and specificity for individual biomarkers, as well as biomarker panels, were determined with different cut-offs for positivity and compared by summary receiver operating characteristic (ROC) curves.

The expression of all four biomarkers was high in PBA versus normal ducts, with a mean Histoscore of 150 vs. 0.4 for KOC, 165 vs. 0.3 for S100P, 115 vs. 0.5 for mesothelin and 200 vs. 14 for MUC1 ($p<0.0001$ for all comparisons). Five cut-offs were carefully chosen for sensitivity/specificity analysis. Four of these cut-offs, namely 5%, 10% or 20% positive cells and Histoscore 20 were identified using ROC curve analysis and the fifth cut-off was moderate-strong staining intensity. Using 20% positive cells as a cut-off achieved higher sensitivity/specificity values: KOC 84%/100%; S100P 83%/100%; mesothelin 88%/92%; and MUC1 89%/63%. Analysis of a panel of KOC, S100P and mesothelin achieved 100% sensitivity and 99% specificity if at least 2 biomarkers were positive at the 10% cut-off; and 100% sensitivity and specificity at the 20% cut-off.

We examined biomarkers using various cut-offs to identify a suitable panel and cut-off based on diagnostic accuracy. A biomarker panel of KOC, S100P and mesothelin with at least 2 biomarkers positive was found to be a good panel with both 10% and 20% cut-offs in resection specimens from patients with PBA.

## 4.2 Introduction

Adenocarcinomas of the head of the pancreas and extra-hepatic cholangiocarcinomas (CCC) present similarly most often with jaundice, pain or weight loss (286). Morphological similarities in addition to generally poor prognosis for both diseases enable PDAC to be grouped with extra-hepatic CCC to form so-called pancreatico-biliary adenocarcinomas (PBA).

Diagnosis of PBA relies upon a combination of radiological and cytology or pathology findings (22, 287-289). Confirmatory tissue diagnosis is necessary before chemotherapy or radiotherapy treatment, however a biopsy specimen is not always required for resection when the suspicion of cancer is high; as generally, the resection will provide therapeutic benefit, and substantially delaying surgery to confirm a diagnosis may deny potentially curative treatment (28, 58, 59, 61, 113, 288). Cytological samples obtained at EUS-FNA and ERCP with brushing requires the distinction of malignant PB epithelial cells from reactive pancreatic and bile duct cells as well as other gastrointestinal contaminants. This task requires tremendous expertise and can be difficult (102).

We performed a meta-analysis of potential PDAC IHC diagnostic biomarkers (290) aiming to generate a list of biomarkers assessed in either surgical or cytology specimens, where PDAC was compared with normal pancreas and/or chronic pancreatitis. Meta-analytical results showed KOC, S100P, mesothelin and MUC1 to be high-ranking candidates. These biomarkers have not entered into routine clinical practice partly because they were investigated in separate studies with relatively small sample sizes and without uniform and clinically appropriate thresholds for positivity.

We sought to investigate the utility of these four candidate biomarkers in the characterisation of PBA, including both PDAC and CCC. CCC has been included because it often enters the clinical and pathological differential diagnosis; and its positive biomarkers are generally shared with PDAC (237, 276, 291, 292).

The following is a brief description of the biomarkers investigated here; found in previous chapter. KOC (K homology domain containing protein overexpressed in

cancer) is a promising biomarker investigated in several different neoplasms. It is also called IMP3 (Insulin like growth factor mRNA binding protein). It is expressed in the developing epithelium and placenta during embryogenesis but the expression is limited in the adult tissues (293). IMP3 has been shown to promote proliferation in leukemia cells, pointing to its role in regulating cellular proliferation (294). It is overexpressed in different malignancies including pancreatic, lung, breast, colon and endometrial carcinomas (180, 295-298).

S100P first identified from placenta, belongs to the S100 family of calcium binding proteins (299). S100P expression is associated with cell proliferation, survival and invasiveness (300). S100P is overexpressed in pancreatic, lung, breast, gastric and colorectal cancer (159, 177, 301).

Mesothelin is a glycoprotein that is attached to the cell membrane by a glycosylphosphatidyl-inositol anchor. It is a differentiation antigen and in normal human tissues is expressed only in mesothelial cells lining pericardium, pleura and peritoneum (302). The precise biological function of mesothelin is largely unknown but it may play an important role in cellular proliferation and tumour progression (303). Mesothelin is overexpressed in several malignancies including mesotheliomas, pancreatic, lung and ovarian carcinoma (236, 238, 304).

MUC 1 (Mucin 1; also known as PEM) is a transmembrane glycoprotein with a heavily glycosylated extracellular domain. It is normally expressed in the epithelial cells of breast, esophagus, stomach, duodenum, pancreas, uterus and lung tissue (305). Tumour associated MUC1 is different from MUC1 expressed in normal cells. MUC1 is aberrantly glycosylated and overexpressed in a range of epithelial tumours. MUC1 mediates the production of growth factors that promote cellular proliferation and survival (306). MUC1 is overexpressed in a variety of tumours including pancreatic, ampullary, gastric, esophageal and breast (170, 176, 307-310).

### 4.2.1 Aims

Having identified better biomarkers from the meta-analysis, the aims of this chapter were as follows: to validate biomarkers by IHC and to assess the diagnostic accuracy (sensitivity and specificity) of identified biomarkers singly

and in panels in a local pancreatico-biliary resection cohort; to investigate and find appropriate diagnostic cut-offs resulting in suitable diagnostic sensitivity and specificity for the categorisation of PBA versus normal. These aims should help us in determining a clinically useful diagnostic biomarker or panel of biomarkers with a robust cut-off for positivity that could potentially be taken forward for validation in PBA cytology samples.

KOC, mesothelin, S100P and MUC1 were investigated in one cohort (SD TMAs) and the results are presented in this chapter. Due to loss of tissue cores in SD TMAs, maspin was investigated in another cohort (NBJ TMAs) and the results are presented in next chapter (chapter 5). However, both cohorts originate from the same surgical and pathology departments.

## 4.3 Results

IHC was performed on an automated platform for KOC, mesothelin, S100P and MUC1. Three tissue microarrays (TMAs) from patients with PBA were used for IHC. The TMAs are composed of both tumour and normal cores in a ratio of 3:2. After staining, TMAs were scanned and uploaded in a digital microscope (Distiller 2.2) for scoring the immunostaining. The scoring was performed manually using a scoring sheet and in a blinded fashion to the final diagnosis (tumour or normal). All scoring of immunostaining was performed by the author (AA) and a consensus was reached on difficult cores with a consultant histopathologist (KAO).

A semiquantitative "Histoscore" with a range (0-300) taking into account both staining intensity and the proportion of positive cells was used for evaluating protein expression. Various cut-offs were then established from Histoscores and further analysis of sensitivity and specificity was then performed.

The results are subdivided into four main categories: expression of biomarkers in tumour and normal tissue using "positivity" (proportion of positive cells) and "Histoscore" (both intensity and proportion); comparing the expression of biomarkers in PDAC and CCC; establishing cut-offs from receiver operating characteristics (ROC) curves; and calculating sensitivity and specificity based on various cut-offs singly and as a panel of biomarkers. I hypothesised that by combining various biomarkers in panels we would be able to determine the combination of biomarkers that would deliver the best diagnostic sensitivity and specificity.

### 4.3.1 Expression of all biomarkers was epithelial and was higher in tumour than normal tissue

For each marker assessed in the PB TMAs, IHC staining was seen only in epithelial cells. As expected, KOC expression was observed in the cytoplasm; S100P was expressed in the cytoplasm and nucleus, while mesothelin and MUC1 expression was cytoplasmic and membranous (Figure 4.1). In general, moderate to strong intensity of staining for KOC, mesothelin, S100P and MUC1 in PBA was observed. Moreover, for all four biomarkers, significantly higher expression in tumour

versus normal tissue (non-neoplastic ducts or pancreatic acinar tissue) was observed as expected (Table 4.1, $p<0.0001$, Independent sample t test).

### 4.3.2 The percentage positivity in tumour is similar for all biomarkers but Histoscores show variation in expression levels

When scored simply as the percentage of positive staining cells, similar results for all four biomarkers in tumour tissue were observed. As shown in Table 4.1 and Figure 4.2, the mean percentage of positive carcinoma cells in tumour cores was 74% for KOC, 75% for S100P, 73% for mesothelin and 75% for MUC1. It was important to quantify not only the extent of biomarker expression, but also the level of expression. By employing the Histoscore scoring method, which takes into account both the staining intensity and the extent of expression across the tissue, we were able to perform a more comprehensive analysis of our biomarkers. Utilizing this method revealed variance of expression of the different biomarkers. As shown in Table 4.1 and Figure 4.3, the mean tumour Histoscore for MUC1 was 193, while for S100P, KOC and mesothelin, the mean tumour Histoscores were 165, 150 and 115 respectively.

A biomarker highly expressed in tumour tissue is a better candidate for diagnostic purposes as it achieves better sensitivity values especially in samples with less tissue such as cytology samples. All four biomarkers investigated in this project are highly expressed in adenocarcinoma and thus they are better diagnostic candidates. Next, I assessed the expression of biomarkers in normal ducts and pancreatic acini to evaluate biomarker specificity.

### 4.3.3 Expression of all biomarkers in normal tissue was very low except MUC1

Although one biomarker, MUC1, was expressed in normal ducts as evidenced by the mean percentage positivity of 18% of normal cells in normal duct cores, the expression of the other three biomarkers was very low in normal pancreatico-biliary ducts (Table 4.1, Figure 4.2 & 4.3). Thus, in a mixed population of tumour and normal cells, only positive staining would need to be taken into account for diagnostic purposes, as positivity is essentially associated only with tumour cells. Furthermore, there were no significant differences in biomarker

expression between normal ducts only and normal ducts and acini together (Table 4.2). Again this means that in a mixed population of tumour and normal cells, only positive staining needs to be taken into account and is essentially associated only with tumour cells. Thus, IHC staining using these markers could greatly facilitate interpretation of cytology samples. The cytology samples from patients with PDAC are composed of adenocarcinoma cells, normal ductal and acinar epithelial cells and gastro-intestinal contaminants amongst other inflammatory and degenerating cells.

A biomarker with very low expression in normal tissue is a better candidate for diagnostic purposes due to high specificity and not classifying the benign disease process as malignant. Thus KOC, mesothelin and S100P are more specific candidates for PBA than MUC1 for diagnostic purposes. The expression of all three biomarkers is limited to adenocarcinoma cells with very low expression in benign tissue. These biomarkers might therefore have optimum diagnostic sensitivity and specificity. I next analysed the expression of these biomarkers in PDAC compared to CCC and then calculated the sensitivity and specificity of biomarkers.

### 4.3.4 Expression of all biomarkers was similar in PDAC and CCC

Biomarkers expression was also assessed in PDAC compared to CCC. Generally, the expression of all four biomarkers is comparable between PDAC and CCC and this is supported by the lack of any statistically significant difference in the mean expression of biomarkers between these two tumour types ($p>0.05$, independent sample t test). The mean positivity in PDAC compared to CCC for KOC, S100P, mesothelin and MUC1 is 74% vs. 60%, 75% vs. 80%, 72% vs. 75% and 76% vs. 72% respectively. In addition, the mean Histoscore in PDAC compared to CCC for KOC, S100P, mesothelin and MUC1 is 155 vs. 131, 160 vs. 180, 120 vs. 135 and 195 vs. 198 respectively (Table 4.3).

Therefore, for sensitivity and specificity analyses PDAC and CCC were grouped as pancreatico-biliary adenocarcinomas (PBA). Next I assessed the sensitivity and specificity of biomarkers using various cut-offs calculated with the help of receiver operating characteristics (ROC) curve analyses.

### 4.3.5 Cut-offs were established using ROC curve analysis

The sensitivity and specificity of these four biomarkers were evaluated using five cut-offs (thresholds) for positivity as follows: 5% positive cells of any staining intensity (5% cut-off); 10% positive cells of any staining intensity (10% cut-off); 20% positive cells of any staining intensity (20% cut-off); moderate or strong staining of any cells (+2/+3 cut-off); and Histoscore equal to and more than 20 (HS20). Three of these cut-offs were based on percentage of positive cells and identified by ROC curve analysis. The sensitivity of each biomarker was plotted against (1 − specificity), and ROC curves with coordinates were generated for all four biomarkers. The area under the curve was 0.93 (0.88-0.97, 95% CI) for KOC, 0.92 (0.85-0.99, 95% CI) for S100P, 0.95 (0.92-0.99, 95% CI) for mesothelin, and 0.87 (0.81-0.93, 95% CI) for MUC1.

Based on the percentage of positive cells in the tumour compared with normal cores, ROC curve analysis allowed us to assess potential cut-offs, from 5% positive cells to 95% positive cells, with their corresponding sensitivity and specificity values for all four biomarkers (Table 4.4, Figure 4.4). Three suitable cut-offs; 5%, 10% or 20% of positive cells of any staining intensity, were selected based on their sensitivity and specificity values. The fourth cut-off was based on moderate to strong staining intensity (+2/+3 staining) in any of the cells. This was selected as moderate to strong staining was expected to be easily interpreted by pathologists. Interestingly, cases with +2/+3 staining for all four biomarkers have more than 20% cells positive for each of the four biomarkers. Indeed patients with +2/+3 staining are represented only 5 cases with less than 50% of cells positive for MUC1, 2 cases in which KOC was expressed in fewer than 50% of cells, and only 1 case each for mesothelin and S100P staining with less than 50% positivity. The fifth cut-off was based on Histoscore value of 20 (HS20), and was determined from ROC curve analysis (Figure 4.5).

### 4.3.6 KOC, S100P and mesothelin were good candidate biomarkers based on sensitivity and specificity values

There are two approaches to utilising cut-offs to calculate sensitivity and specificity values. One is to score the biomarker expression as a continuous variable in tumour and normal tissue for example each patient receive a

Histoscore (range, 0-300) and percentage positivity (0%-100%). Then identify appropriate cut-offs from the scoring results using ROC curves that achieve optimum diagnostic sensitivity and specificity for the biomarker. The other is to select cut-off(s) at the start of scoring the tissue for example selecting 10% positive cells as cut-off, then ask the scorer to categorise the tissue as a binary variable i.e. positive or negative for staining based on 10% positive cells. The former approach to cut-off identification was used for all subsequent sensitivity and specificity analysis in this chapter and chapter 5. The latter approach of cut-offs was used in chapter 6 to validate the cut-offs identified in the current chapter for observer agreement. As an example, a 10% cut-off for PDAC would mean the presence of at least 10% positively stained cells for a biomarker in the ductal epithelium. Thus a tissue with more than 10% positive cells in ductal epithelium will be categorised as tumour based on biomarker staining. On the contrary, staining in less than 10% positive cells in ductal epithelium will be categorised as being benign or a normal condition.

The sensitivities and specificities of all four biomarkers were calculated using the five cut-offs namely 5%, 10%, 20%, +2/+3 and HS20. KOC expression appears to show reasonably high sensitivity and specificity for all cut-offs except for the cut-off based on +2/+3 staining, which resulted in low sensitivity of only 67%. The 20% cut-off achieves marginally better sensitivity (84%) and specificity (100%) values compared with other cut-offs for KOC (Figure 4.6A). S100P appears to have similar sensitivity and specificity values for all five cut-offs with the 20% cut-off again achieving better combination of specificity and sensitivity, with values of 83% sensitivity and 100% specificity (Figure 4.6B). Applying the five cuts-offs to the analysis of mesothelin expression resulted in significantly different sensitivity and specificity values, however, the best combination was again achieved using the 20% cut-off, with 88% sensitivity and 92% specificity (Figure 4.6C). Although the sensitivity of MUC1 is high for all cut-offs, its specificity is too low with a range of 18%-63% for various cut-offs (Figure 4.6D). Thus MUC1 might not be a good diagnostic candidate.

In summary KOC, meosthelin and S100P achieved reasonable combinations of sensitivity and specificity due to their high expression in tumour tissue but very low expression in normal tissue. The sensitivity of MUC1 is high due to

overexpression in tumour but its specificity is low and hence it is not a suitable candidate for diagnostic purposes comparing PBA with benign tissue. KOC, mesothelin and S100P are highly specific biomarkers but their sensitivity of detecting malignancy is not perfect. Thus as a single candidate biomarker they might miss malignancy and could potentially lead to false negative diagnosis. One way of improving the diagnostic sensitivity is to use panel of biomarkers. Hence in the next section I will present panels of biomarkers to determine a highly sensitive and specific panel for future validation.

### 4.3.7 A biomarker panel of KOC, S100P and mesothelin with at least 2 biomarkers positive was found to be a good working panel

Due to the biological variability between individual tumours it is unlikely that one biomarker will stain all tumours. A panel approach using biomarker combinations can thus be an adequate diagnostic solution. I next wanted to assess the sensitivity and specificity achieved using panels of biomarkers. The 10% and 20% cut-offs were selected for this investigation. These two cut-offs achieve good combination of sensitivity and specificity in the single biomarker analysis as shown in the previous section (section 4.3.6 and Figure 4.6).

The sensitivity and specificity was assessed when one biomarker in a panel is positive. For example for a 10% cut-off, a positivity of 10% epithelial cells for one or more biomarker in a panel is sufficient to categorise the patient into tumour category. Using this approach four panels of biomarkers were tested as shown in Table 4.5. These panels were: a panel comprising all four biomarkers; a panel of three biomarkers (KOC, S100P and mesothelin); and two panels of two biomarkers (KOC and mesothelin, KOC and S100P). A panel of KOC and S100P achieved better combinations of sensitivity/specificity values of 98%/96% for the 10% cut-off and 99%/99% for the 20% cut-off. In addition, these panels were compared by combined SROC curve, using both the 10% cut-off (Figure 4.7A) and 20% cut-offs (Figure 4.7B). The combined SROC curve showed that a panel of KOC and S100P is superior to the other panels for both 10% and 20% cut-offs.

Lastly, one biomarker panel comprising KOC, S100P and mesothelin was tested for sensitivity and specificity when at least two biomarkers in the panel are

positive. For example for a 10% cut-off, a positivity of 10% epithelial cells for two or more biomarkers in a panel will be sufficient to categorise the patient into tumour category. Analysis of this panel achieved 100% sensitivity and 99% specificity for 10% cut-off; and 100% sensitivity and specificity for 20% cut-off (Table 4.5). Taken together, our results show that a panel of KOC, mesothelin and S100P could be used to improve the diagnosis of PBA in difficult to diagnose cases.

## 4.4 Discussion:

**Four potentially diagnostic biomarkers, KOC, S100P, mesothelin and MUC1, were investigated in a relatively large cohort of PB patients (n=99).** The expression levels of KOC, S100P and mesothelin were high in tumour tissue compared with normal tissue. The diagnostic accuracy (sensitivity and specificity) of KOC and S100P individually was greater than that of mesothelin and MUC1. A panel of KOC, S100P and mesothelin with at least two biomarkers positive achieved optimal diagnostic accuracy in the differentiation of adenocarcinoma from normal tissue.

**IHC biomarkers have been investigated to improve the diagnosis of PBA but none is yet routinely used.** They have been investigated in both the surgical and cytological cohorts for potentially improving the diagnosis of PBA (159, 217, 233, 234, 237, 240, 253). There are six significant reasons delaying the clinical translation of diagnostic biomarkers in PBA and probably in other cancers (outlined in detail in chapter 1). These reasons and our approach to address them are outlined below.

**First,** a plethora of research exists on diagnostic IHC biomarkers coming from the bench assessed in pilot studies. There are many excellent papers but fewer validation studies for biomarkers that have shown promising results. Clearly, validation is important for future clinical application. Therefore, we performed a meta-analysis on diagnostic IHC biomarkers for PDAC (290), to review, quantify and assess the performance of already existing biomarkers and to try and identify superior candidate biomarkers.

The biomarkers derived from the meta-analysis in PDAC were applied in our study to both PDAC and CCC samples. Separate meta-analysis was not performed for CCC, because there are relatively few published papers on biomarkers in CCC (approximately 20-fold fewer than for PDAC; PubMed search in June 2014, unpublished data). However, those papers which are available for CCC suggest that the biomarker expression profile is similar to PDAC. To our knowledge, all of the known positive biomarkers for PDAC (versus corresponding normal tissue), including MUC1, P53, CK17, mesothelin, fascin, MUC4, 14-3-3σ and prostate stem

cell antigen, show similar IHC expression in CCC (versus corresponding normal tissue) (237, 276, 291, 292).

For these reasons, we focused on PDAC for the identification of potential diagnostic biomarkers then tested the resulting candidates in TMAs containing tissue from both PDAC and CCC using IHC. From our meta-analysis, we selected KOC (217, 240, 241, 253), S100P (159, 166, 177), mesothelin (172, 236, 255) and MUC1 (170, 188, 276) for investigation. We found that expression of these biomarkers was similar in PDAC and CCC (Table 4.3): our results therefore agree with the previous literature (237, 276, 291, 292).

**Second,** the sample size for studies investigating diagnostic biomarkers for PDAC is relatively small (median sample size, n=48 from 57 articles). Moreover, matched normal tissue for most of the carcinoma case was not always available, leading to even smaller sample sizes for calculating biomarker specificity. Therefore, statistical power is relatively low and subsequently potentially useful biomarkers may be ignored. Our relatively larger sample size of 99 PBA cases (n=99 adenocarcinomas and n=99 matched normal tissue for each case; total n=198) provided a solid platform for investigating these diagnostic IHC biomarkers.

**Third,** the lack of a standardised scoring system and absence of a uniform cut-off (threshold) for the interpretation of IHC remains problematic. Thus, researchers use a variety of traditional and novel scoring systems and diverse cut-offs, making the adoption of scoring systems and cut-offs potentially challenging for pathologists (159, 165, 172, 232, 234, 236, 237, 240, 264). We systematically chose cut-offs from ROC curve analysis to fully explore the diagnostic potential of all four biomarkers. These cut-offs provide an opportunity for the pathologists to select the best threshold that is more clinically applicable and has the potential to be routinely used in pathology. Three of these cut-offs are based on the proportion of positive cells (5%, 10% and 20%) with staining of any intensity. The fourth cut-off is based on any proportion of cells exhibiting moderate and strong staining intensity, and the fifth cut-off is based on a Histoscore of 20. Notably, the 20% cut-off and Histoscore 20 provide reasonable sensitivity and specificity values for PBA diagnosis. A higher Histoscore value could potentially lead to more false negatives in tumour cases, therefore, a low

cut-off value of 20 was chosen. Clearly, this cut-off will remove the probability of false negative and should increase the diagnostic confidence of pathologists for higher Histoscore values. For example, a Histoscore value of 200 for a biomarker in a suspicious case might help the pathologist to diagnose a tumour with confidence and with a much higher specificity.

**Fourth,** most of the IHC diagnostic biomarkers have been investigated individually (177, 234, 241, 252, 254), with few studies reporting the utility of biomarker panels. (172, 253). We carefully selected candidate biomarkers reported in different studies (KOC, mesothelin, S100P and MUC1) for investigation in a single experimental setting. Investigation of these biomarkers in a single cohort gave us the opportunity to compare biomarkers, and then further explore their diagnostic accuracy in a panel. Expectation from an ideal diagnostic biomarker is its ability to identify the diseased population (sensitivity) and exclude the normal population (specificity) in 100% cases. However, no single biomarker is 100% perfect; therefore these biomarkers were investigated in various combinations, to select an optimum panel for potential clinical application. For example, the individual sensitivity/specificity of KOC and S100P at a cut-off of 20% positive cells was 84%/100% and 83%/100% respectively. However, using a panel of KOC and S100P improved sensitivity to 99% without compromising the specificity (99%).

Furthermore, using a panel of KOC, S100P and mesothelin with at least two positive biomarkers achieved almost 100% sensitivity and specificity for both 10% and 20% cut-offs. This approach would assign a patient into the tumour positive category if two or more biomarkers are positive, possibly giving more assurance to the pathologist before assigning the patient into positive category. Moreover, a combination of KOC, S100P and mesothelin antibodies should stain all major cellular compartments (cell membrane, nucleus and cytoplasm). Clinically, a cytology sample comprises a mixed population of cells and this panel will stain malignant cells more intensely making the interpretation of IHC convenient for the pathologist. The possible additional advantage of KOC is that it is not expressed in the contaminating gastrointestinal epithelial cells that are usually present in cytological samples (217, 271). Our data also confirm the lack of expression of KOC in normal duodenum. Taken together, our results reinforce

the reported sensitivity/specificity values for KOC, S100P and mesothelin (159, 172, 240, 271) and further explores their utility as a panel.

**Fifth,** is that different research groups use different IHC experimental conditions, primary antibodies, clones, antigen retrieval, antibody dilutions and manual/automated platforms that could potentially lead to a diverse range of sensitivity and specificity values for biomarkers (171, 172, 232, 233, 238, 249). We thoroughly searched the literature for IHC parameters for KOC, S100P, mesothelin and MUC1. Those IHC parameters that achieved superior diagnostic accuracy were selected and further optimised in our histology laboratory before staining our cohort.

**Sixth**, an important requirement for biomarker translation to the clinic is independent validation with the aim of improving already existing diagnosis. Purposeful validation in surgical and cytological tissue from PBA cohorts and subsequent prospective clinical study on cytological samples is deficient. Therefore, as an important step for potential clinical translation we investigated KOC, S100P, mesothelin and MUC1 in a surgical cohort of PBA patients with promising results for KOC, S100P and mesothelin as a biomarker panel.

The next step forward is to possibly investigate these biomarkers in a retrospective study and then in a prospective cohort of cytology samples. This chapter systematically attempted to answer all six major reasons hindering the clinical translation of diagnostic IHC biomarkers for pancreatic cancer. It also provides future direction and work packages to be performed before these diagnostic biomarkers can be used in day-to-day pathology practice.

**The strengths and limitations are:** The strengths include relatively large sample size, using semi-quantitative Histoscores for scoring, using multiple cut-offs for biomarkers and panel approaches. The main limitation of this project is the use of a separate TMA set for maspin.

**Conclusion:** Our results demonstrate that a biomarker panel of KOC, S100P and mesothelin is capable of categorising PB malignancy with high diagnostic accuracy in resection specimens. We plan to investigate this panel in archival cytological samples. As an adjunct to cytology, this panel has the potential to

augment the categorisation for challenging diagnostic cases in routine clinical practice.To our knowledge, this is the first study of PB literature that identified cut-offs systematically for diagnostic purposes and used stringent panels to identify an optimum biomarker panel.

**What this adds:** This study validated the biomarkers identified in meta-analysis. Histoscore was used as a scoring system and five potential diagnostic cut-offs were investigated. Biomarkers showed similar expression in PDAC and CCC. Biomarkers were validated singly and in panels for PBA diagnosis.

**What Next:** I will now investigate maspin as a potential diagnostic biomarker in the next chapter, which will be followed by investigation of cut-offs for observer agreement and finally investigation of panel of biomarkers in cytology samples.

**Table 4.1: Summary statistics of KOC, S100P, mesothelin and MUC1 expression comparing PBA with normal tissue.**

| Biomarkers | | Pancreaticobiliary Adenocarcinoma | Normal tissue | P value |
|---|---|---|---|---|
| **KOC** | | | | |
| ***Positivity**** | Mean | 74% | 0.4% | <0.0001 |
| | Median | 100% | 0% | |
| ***Histoscore*** | Mean | 150 | 0.5 | <0.0001 |
| | Median | 180 | 0 | |
| **S100P** | | | | |
| ***Positivity*** | Mean | 75% | 0.30% | <0.0001 |
| | Median | 100% | 0% | |
| ***Histoscore*** | Mean | 165 | 0.3 | <0.0001 |
| | Median | 180 | 0 | |
| **Mesothelin** | | | | |
| ***Positivity*** | Mean | 73% | 4% | <0.0001 |
| | Median | 90% | 0% | |
| ***Histoscore*** | Mean | 115 | 4 | <0.0001 |
| | Median | 110 | 0 | |
| **MUC1** | | | | |
| ***Positivity*** | Mean | 75% | 18% | <0.0001 |
| | Median | 90% | 10% | |
| ***Histoscore*** | Mean | 193 | 48 | <0.0001 |
| | Median | 200 | 30 | |

**Note:** *Positivity (percentage of positive cells with any staining intensity in tumour and normal tissue); P value (shows the statistical significance of the difference in expression of a biomarker in tumour vs. normal tissue); Positivity range (0%-100%), Histoscore range (0-300).

**Table 4.2: Summary statistics of KOC, S100P, mesothelin and MUC1 expression comparing normal ducts with normal ducts & acini together**

| Biomarkers | | Normal ducts | Normal ducts & acini |
|---|---|---|---|
| **KOC** | | | |
| ***Positivity**** | Mean | 0.2% | 0.4% |
| | Median | 0% | 0% |
| ***Histoscore*** | Mean | 0.4 | 0.5 |
| | Median | 0 | 0 |
| **S100P** | | | |
| ***Positivity*** | Mean | 0.3% | 0.3% |
| | Median | 0% | 0% |
| ***Histoscore*** | Mean | 0.3 | 0.3 |
| | Median | 0 | 0 |
| **Mesothelin** | | | |
| ***Positivity*** | Mean | 5% | 4% |
| | Median | 0% | 0% |
| ***Histoscore*** | Mean | 5 | 4 |
| | Median | 0 | 0 |
| **MUC1** | | | |
| ***Positivity*** | Mean | 16% | 18% |
| | Median | 5% | 10% |
| ***Histoscore*** | Mean | 37 | 48 |
| | Median | 14 | 30 |

**Note:** *Positivity (percentage of positive cells of any staining intensity in tumour and normal tissue); Positivity range (0%-100%), Histoscore range (0-300).

**Table 4.3: Summary statistics of KOC, S100P, mesothelin and MUC1 expression comparing PDAC with CCC**

| Biomarkers | | Pancreatic ductal adenocarcinoma | Cholangiocarcinoma | P value |
|---|---|---|---|---|
| **KOC** | | | | |
| ***Positivity**** | Mean | 74% | 60% | 0.09 |
| | Median | 100% | 90% | |
| ***Histoscore*** | Mean | 155 | 131 | 0.33 |
| | Median | 180 | 120 | |
| **S100P** | | | | |
| ***Positivity*** | Mean | 75% | 80% | 0.38 |
| | Median | 100% | 100% | |
| ***Histoscore*** | Mean | 160 | 180 | 0.15 |
| | Median | 150 | 190 | |
| **Mesothelin** | | | | |
| ***Positivity*** | Mean | 72% | 75% | 0.65 |
| | Median | 90% | 90% | |
| ***Histoscore*** | Mean | 120 | 135 | 0.37 |
| | Median | 105 | 140 | |
| **MUC1** | | | | |
| ***Positivity*** | Mean | 76% | 72% | 0.58 |
| | Median | 90% | 100% | |
| ***Histoscore*** | Mean | 195 | 198 | 0.91 |
| | Median | 220 | 260 | |

**Note:** *Positivity (percentage of positive cells with any staining intensity in tumour tissue); P value, independent sample t test (shows the statistical significance of the difference in expression of these biomarkers in pancreatic ductal adenocarcinoma vs. cholangiocarcinoma); Positivity range (0%-100%), Histoscore range (0-300).

**Table 4.4: Cut-offs resulting from ROC curve analysis based on the percentage of positive cells in tumour and normal cases for KOC, S100P, mesothelin and MUC1.**

| Cut-offs* | KOC | | S100P | | Mesothelin | | MUC1 | |
|---|---|---|---|---|---|---|---|---|
| | Sensitivity | Specificity | Sensitivity | Specificity | Sensitivity | Specificity | Sensitivity | Specificity |
| **5%** | 87 | 96 | 86 | 97 | 95 | 77 | 92 | 18 |
| **10%** | 87 | 98 | 86 | 97 | 94 | 88 | 90 | 35 |
| **20%** | 84 | 100 | 83 | 100 | 88 | 92 | 89 | 63 |
| **35%** | 81 | 100 | 80 | 100 | 82 | 97 | 87 | 85 |
| **45%** | 77 | 100 | 80 | 100 | 79 | 97 | 84 | 90 |
| **55%** | 72 | 100 | 78 | 100 | 73 | 100 | 77 | 93 |
| **65%** | 71 | 100 | 71 | 100 | 70 | 100 | 71 | 95 |
| **70%** | 70 | 100 | 66 | 100 | 67 | 100 | 70 | 96 |
| **80%** | 68 | 100 | 59 | 100 | 56 | 100 | 65 | 96 |
| **90%** | 63 | 100 | 56 | 100 | 53 | 100 | 53 | 98 |
| **95%** | 56 | 100 | 54 | 100 | 42 | 100 | 40 | 98 |

**Note: *** These cut-offs are based on percentage positivity [percentage of positive cells (0%-100%) of any staining intensity (weak, moderate and strong) in tumour and normal tissues].

**Table 4.5: Panels of biomarkers used for sensitivity and specificity analyses, using 10% and 20% positive cells as cut-offs for positivity**

| **10% positive cells as cut-off** | | |
|---|---|---|
| **Panels** | **Sensitivity** | **Specificity** |
| **KOC, S100P, Mesothelin, MUC1** | 100% | 40% |
| **KOC, S100P, Mesothelin** | 100% | 88% |
| **KOC, Mesothelin** | 97% | 87% |
| **KOC, S100P** | 98% | 96% |
| **KOC, S100P, Mesothelin*** | 100% | 99% |
| **20% positive cells as cut-off** | | |
| **Panels** | **Sensitivity** | **Specificity** |
| **KOC, S100P, Mesothelin, MUC1** | 100% | 65% |
| **KOC, S100P, Mesothelin** | 99% | 94% |
| **KOC, Mesothelin** | 96% | 93% |
| **KOC, S100P** | 99% | 99% |
| **KOC, S100P, Mesothelin*** | 100% | 100% |

**Note: *** At least 2 biomarkers required to be positive in this panel. In the rest of the panels only one biomarker was required to be positive in a panel.

**Figure 4.1: Immunostaining of KOC, S100P, mesothelin and MUC1.**

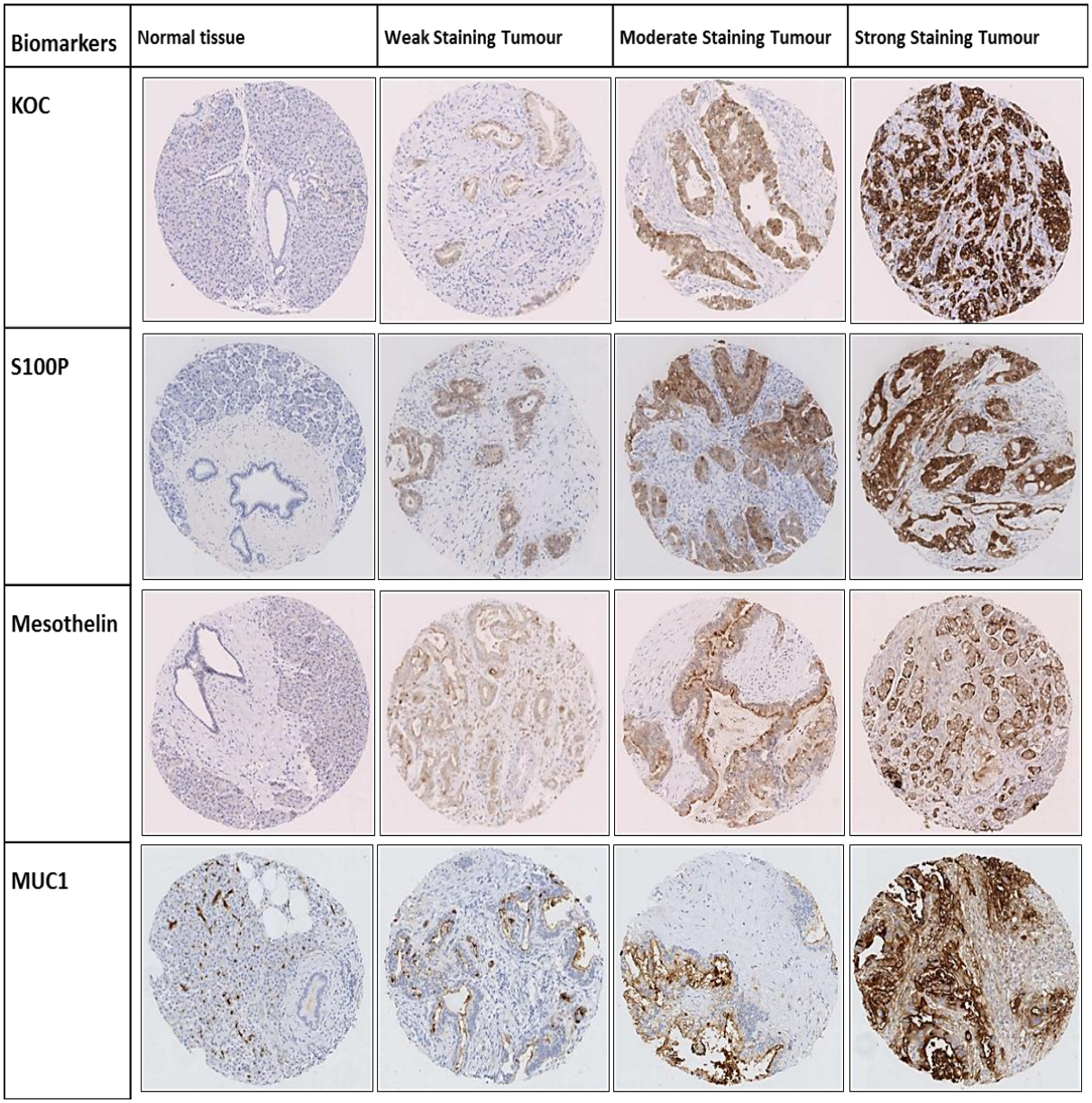
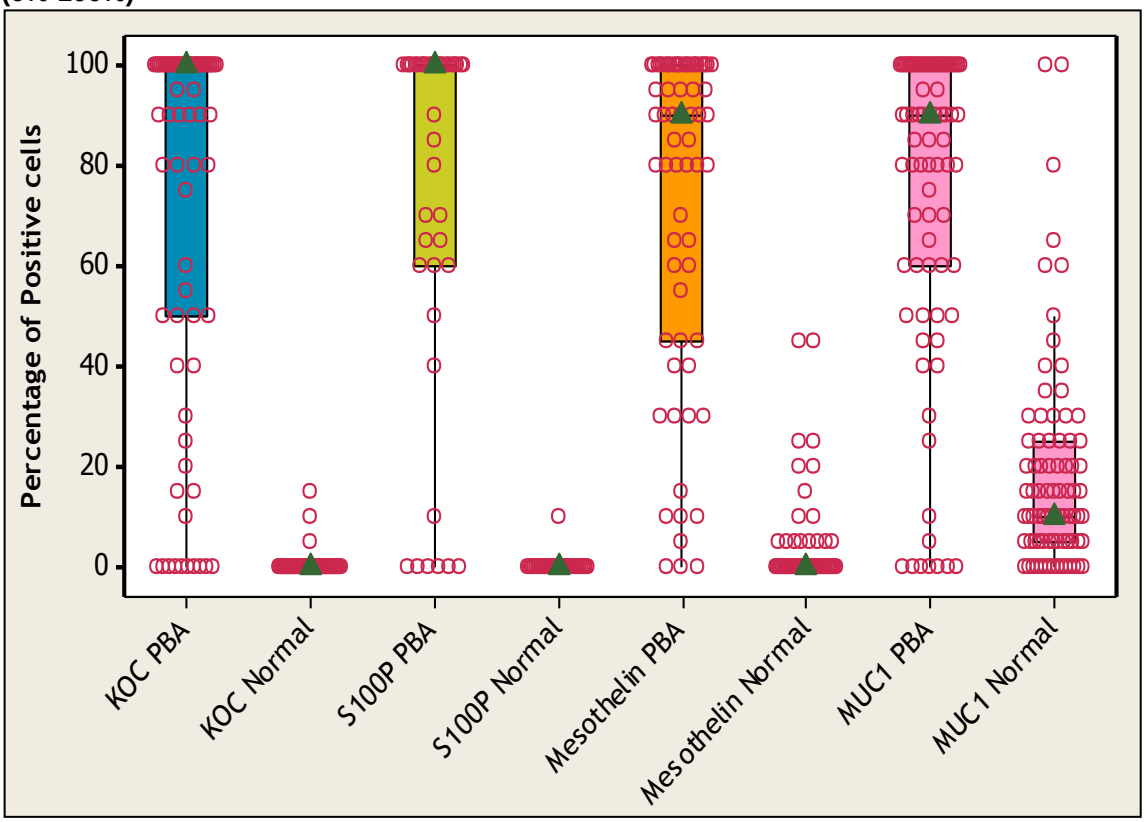


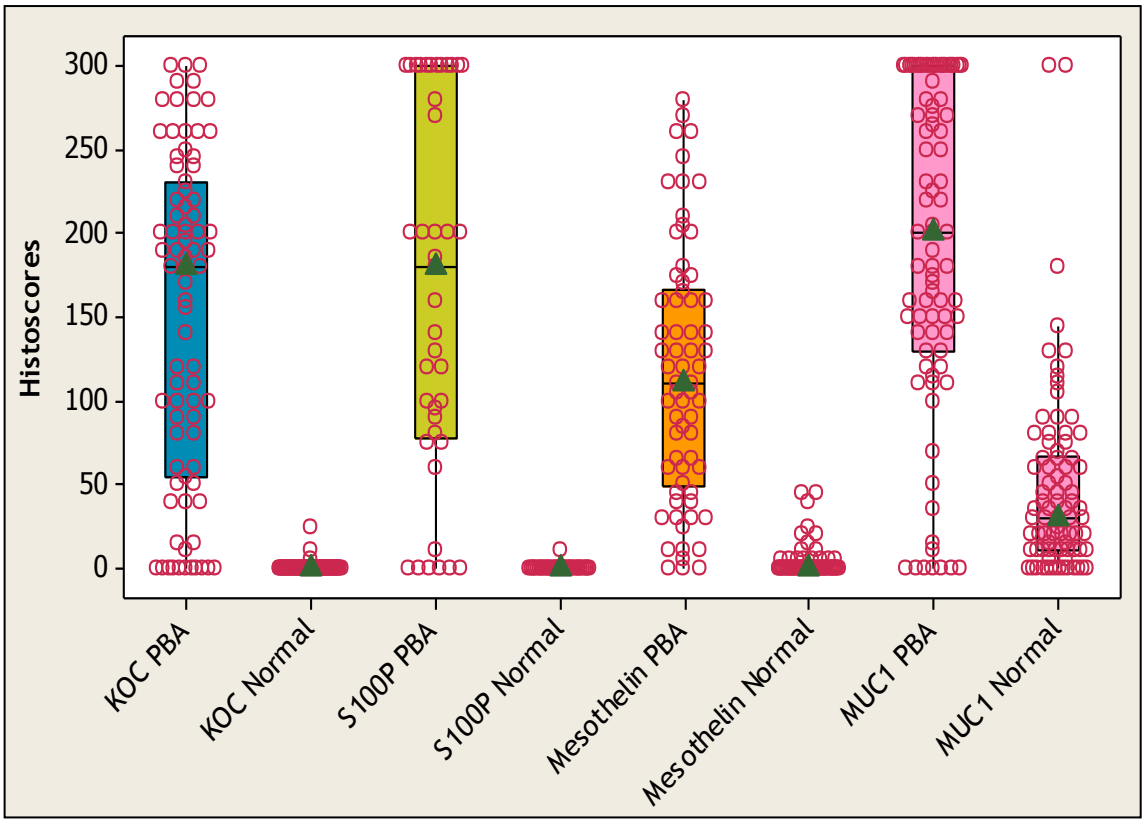

**Figure Legend:** Representative images of staining of all four biomarkers in normal tissue (normal pancreatic tissue) and range of staining intensities (weak, moderate and strong) in tumour tissue from tissue microarray cores. The staining for all biomarkers is epithelial. The staining for KOC is cytoplasmic, staining for S100P is cytoplasmic and/or nuclear, staining for mesothelin and MUC1 is cytoplasmic and/or membranous. Normal ducts and acini and negative for KOC. S100P and mesothelin but some staining in normal acini is present for MUC1.

**Figure 4.2: Boxplots comparing the expression of biomarkers (KOC, mesothelin, S100P and MUC1) in PBA compared to normal tissue based on percentage of positively stained cells (0%-100%)**




**Symbols explained**: Red Circles are individual data symbol. Green triangles are median data symbols.

**Figure 4.3: Boxplots comparing the expression of biomarkers (KOC, mesothelin, S100P and MUC1) in PBA compared to normal tissue based on Histoscores (0-300)**




**Symbols explained:** Red Circles are individual data symbols. Green triangles are median data symbols.

**Figure 4.4: Receiver operating characteristics (ROC) curves for biomarkers based on positive percentage of cells**

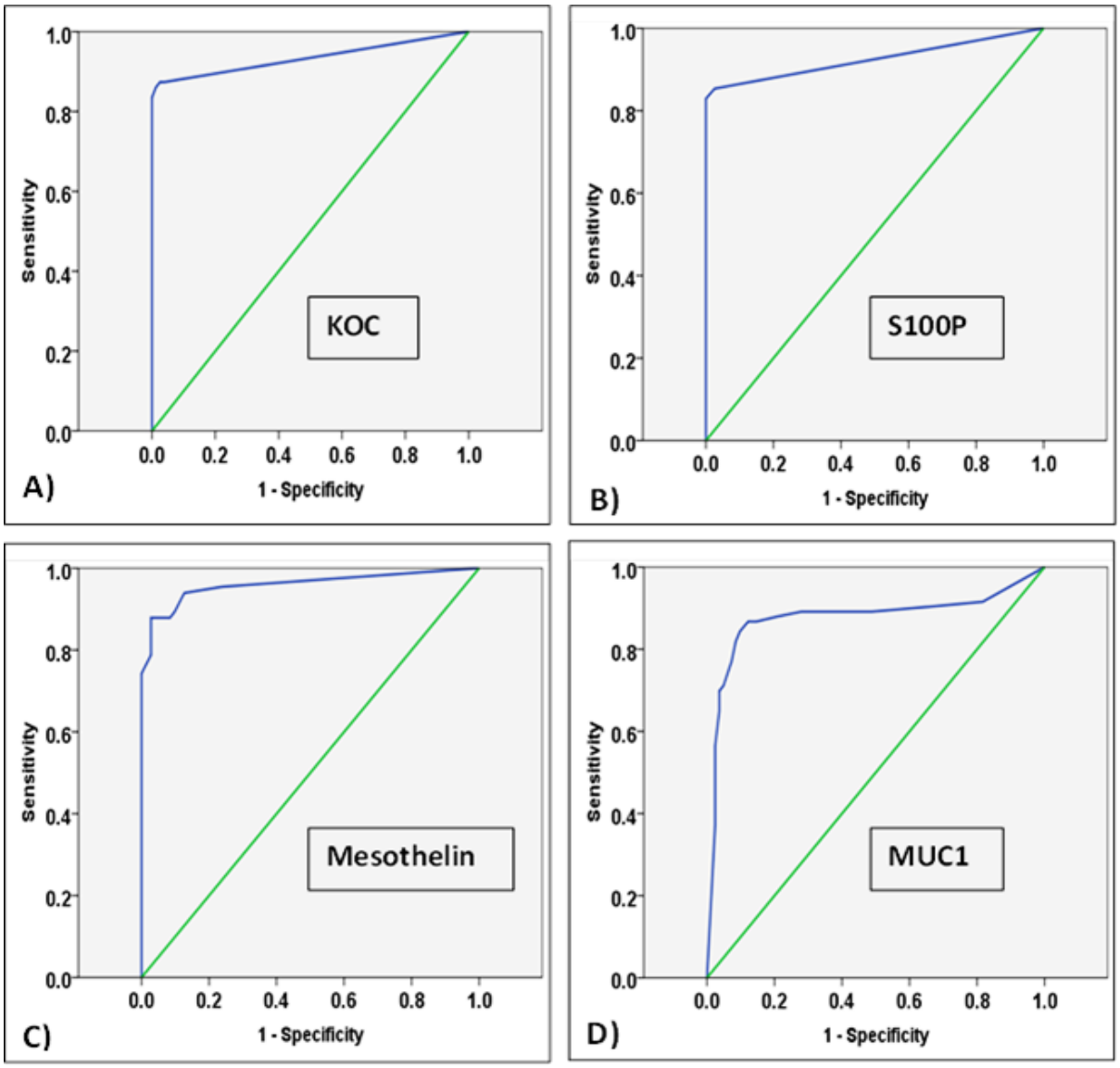



**Figure Legend:** ROC curves based on percentage of positive cells (0%-100%) for any staining intensity (weak, moderate or strong), in tumour and normal cases, for four biomarkers **(A)** KOC, **(B)** S100P, **C)** mesothelin and **D)** MUC1.

**Figure 4.5: Receiver operating characteristics (ROC) curves for biomarkers based on Histoscore**

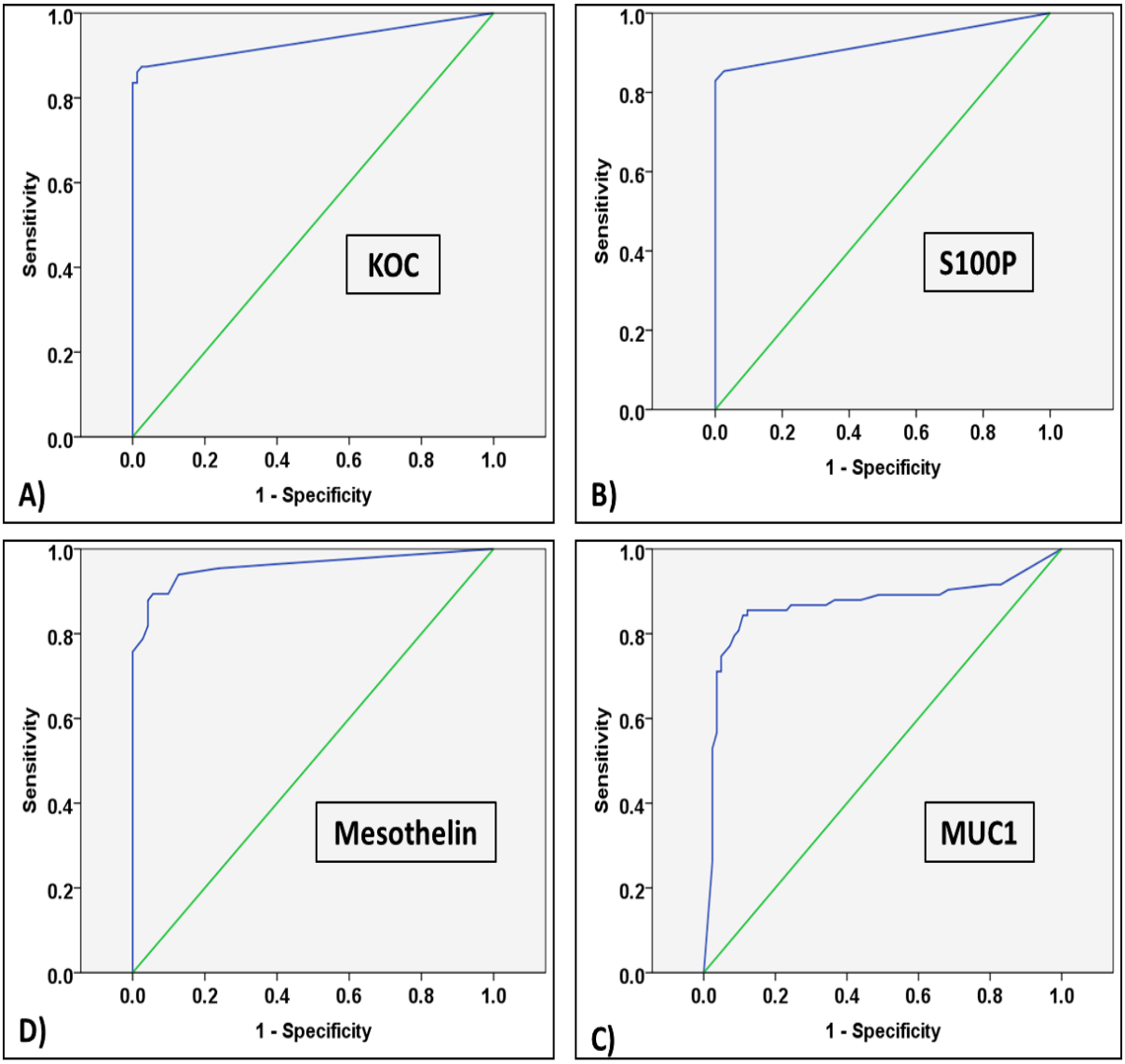



**Figure Legend:** ROC curves based on Histoscore (0-300), in tumour and normal cases, for four biomarkers **(A)** KOC, **(B)** S100P, **(C)** mesothelin and **(D)** MUC1.

**Figure 4.6: Sensitivity and specificity analysis based on five cut-offs for biomarkers.**

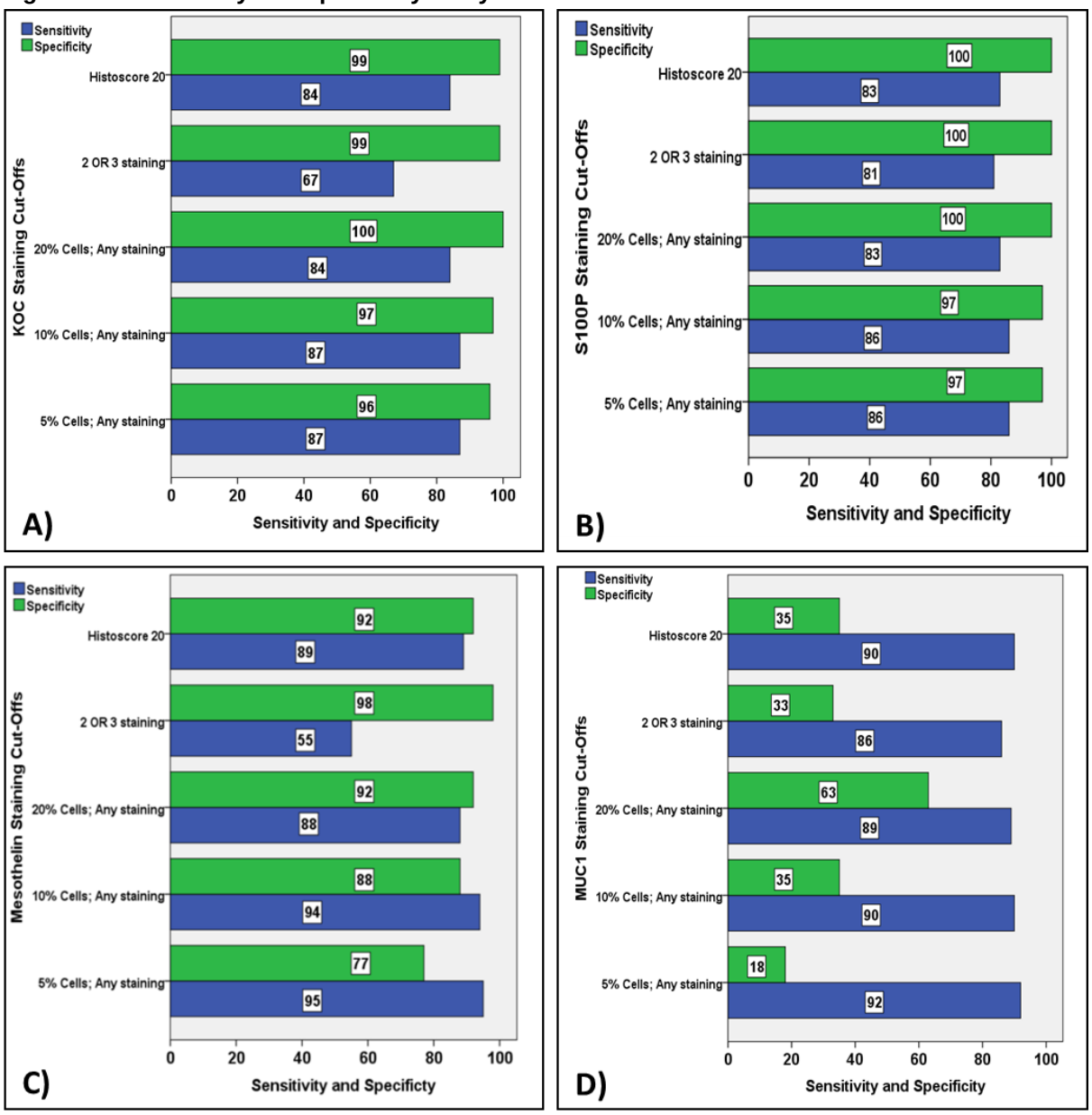



**Figure Legend:** Sensitivity and specificity analysis of biomarkers for the diagnosis of pancreatico-biliary adenocarcinoma compared to normal tissue, based on five cut-offs for positivity: 5% positive cells of any staining intensity; 10% positive cells of any staining intensity; 20% positive cells of any staining intensity; 2 OR 3 intensity i.e. moderate or strong staining of cells; and Histoscore 20. Analysis is presented for A) KOC, B) S100P, C) mesothelin and D) MUC1.

**Figure 4.7: Combined Summary ROC curves for comparing panels of biomarkers.**

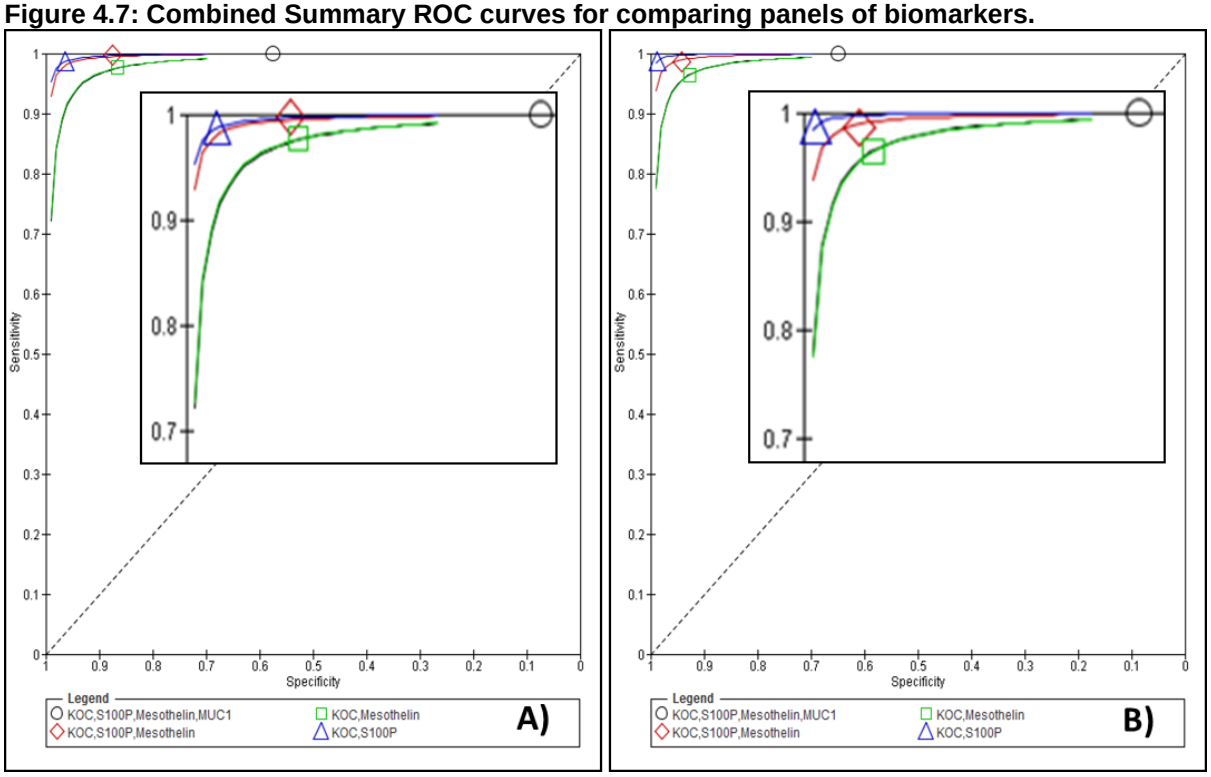



**Figure Legend:** *Combined Summary ROC curves for 10% (A) and 20% (B) cut-offs if only one biomarker was required to be positive in a panel. Four panels of biomarkers were compared. Panel 1 - KOC, S100P, Mesothelin and MUC1; Panel 2 - KOC, S100P, Mesothelin; Panel 3 - KOC, S100P; Panel 4 - KOC, Mesothelin.

*Summary ROC curves plot sensitivity against specificity and draw a summary line depicting combined sensitivity and specificity of a panel. Combined Summary ROC curves compare different panels to show the most “accurate” panel. The summary line at the top left corner shows the biomarker which is most accurate compared to others lying lower and further to the right. This enables the most accurate panel to be identified.

# 5 Expression of maspin in pancreatico-biliary adenocarcinomas as a potential diagnostic immunohistochemistry biomarker

## 5.1 Chapter Summary

Maspin as a potential diagnostic biomarker was also identified from our meta-analysis and is investigated in this chapter. The aim was to independently validate the diagnostic significance of maspin in our cohort and to assess its suitability for diagnostic panel investigation in cytology.

Tissue microarrays containing tumour and normal cores in a ratio of 6:2, from 137 surgically resected PBA patients, were used for immunohistochemistry (IHC). IHC was performed on an automated platform using an antibody against maspin. Tissue cores were scored for staining intensity and the proportion of tissue stained using a Histoscore method (range, 0–300). Sensitivity and specificity was determined with different cut-offs for positivity and compared by summary receiver operating characteristic (ROC) curve.

The expression of maspin was high in PBA versus normal ducts, with a mean Histoscore of 215 vs. 0.4 ($p<.0001$). Five cut-offs were carefully chosen for sensitivity/specificity analysis namely 5%, 10% or 20% positive cells, +2/+3 and Histoscore 20. Using 20% positive cells as a cut-off achieved 96% sensitivity and 99% specificity for maspin.

We examined the potential diagnostic utility of maspin using various cut-offs and found that maspin achieves optimum diagnostic sensitivity and specificity in resection specimens from patients with PBA. The high specificity (98%-100%) might enable maspin to be used as a potential candidate in a diagnostic panel with other biomarkers.

## 5.2 Introduction

Five candidate biomarkers KOC, S100P, mesothelin, MUC1 and maspin were identified from the meta-analysis. The first four were investigated using 'SD TMAs' and the results are presented in chapter 4. In this chapter I will now discuss maspin as a potential diagnostic biomarker for pancreatico-biliary adenocarcinomas (PBA). 'NBJ TMAs' were used instead of 'SD TMAs' for investigating maspin expression, because a significant loss of tissue cores in the later sections of 'SD TMAs' was noticed and thus considered inappropriate for biomarker studies. The loss of tissue cores was due to the intervening use of the TMA resource for biomarker studies by other researchers.

Maspin (mammary serine protease inhibitor) was first identified in normal mammary epithelial cells and has been characterised as a tumour suppressor in many cancer types (311). The anti-tumour effects of maspin are related to inhibiting cell invasion, promoting apoptosis and inhibiting angiogenesis (234, 312). Experimental data suggest the role of maspin as a tumour suppressor but clinical data have conflicting results regarding its prognostic significance in different tumour types (311, 313). Maspin expression is associated with a favourable prognosis in lung, prostate and colorectal cancer, while in pancreatic and breast cancer it predicts poor prognosis (232, 314-317). These conflicting results can partly be explained by the distinct subcellular localisation (cytoplasm, nuclear or both) of maspin (313).

### 5.2.1 Aims

The aims of this chapter are as follows: to validate maspin by IHC and to assess the diagnostic accuracy (sensitivity and specificity) in the local pancreatico-biliary resection cohort and; to investigate and identify appropriate diagnostic cut-offs resulting in good sensitivity and specificity for the distinction of PBA from normal. These two aims will help us in determining the suitability of maspin as a candidate diagnostic biomarker and its potential role in the panel of biomarkers identified in the previous chapter.

## 5.3 Results

### 5.3.1 Expression of maspin was epithelial and was higher in tumour than normal tissue

IHC staining was seen only in epithelial cells. As expected, maspin expression was observed in the cytoplasm and nucleus but only cytoplasmic staining was scored for diagnostic purposes (Figure 5.1). Generally, staining intensity was moderate to strong and significantly higher expression in tumour versus normal tissue (non-neoplastic ducts or pancreatic acinar tissue) was observed (Table 5.1, $p<0.0001$, Independent sample t-test).

The mean percentage of positive carcinoma cells in tumour cores was 90% for maspin (Table 5.1 and Figure 5.2). By employing a Histoscore scoring method, the mean tumour Histoscore for maspin was 215 (Table 5.1 and Figure 5.3). The mean expression of maspin (90% positivity and 215 histoscore) is higher than the biomarkers studied in the previous chapter.

As discussed early in this thesis an ideal diagnostic biomarker should be highly expressed in the tumour tissue compared to normal and maspin appears to have a very high expression in PBA. This makes maspin a desirable diagnostic candidate for investigation in cytology.

### 5.3.2 Expression of maspin in normal tissue was very low

The expression of maspin was very low in normal tissue as shown by the mean percentage of positive cells (0.4%) and mean Histoscore (0.4) in normal tissue (Table 5.1, Figure 5.2). Thus, in a mixed population of tumour and normal cells, only positive staining would need to be taken into account for diagnostic purposes, as positivity is essentially associated only with tumour cells. Furthermore, there were no significant differences in biomarker expression between normal ducts only and normal ducts and acini together (Table 5.2). Again this means that in a mixed population of tumour and normal cells, only positive staining needs to be taken into account and is essentially associated only with tumour cells.

An ideal diagnostic biomarker should have very low expression in normal tissue compared to tumour and maspin appears to have very low expression in normal tissue. This makes maspin a desirable diagnostic candidate for investigation in cytology. Thus, IHC staining using maspin could greatly facilitate interpretation of cytology samples with indeterminate or difficult diagnosis.

### 5.3.3 Expression of maspin was similar in PDAC and cholangiocarcinomas (CCC)

Maspin expression was also assessed in PDAC compared to CCC as shown in Table 5.3. The expression of maspin is comparable between PDAC and cholangiocarcinoma and this is supported by no statistically significant difference in the mean expression of biomarkers between these two tumour types ($p>0.05$, independent sample t-test). The mean positivity in PDAC compared to CCC for maspin is 90% vs. 92% and the mean Histoscore is 215 vs. 220.

Therefore, for sensitivity and specificity analyses PDAC and CCC were grouped as pancreatico-biliary adenocarcinomas (PBA). Next I assessed the sensitivity and specificity of biomarkers using various cut-offs calculated with the help of receiver operating characteristics (ROC) curve analysis.

### 5.3.4 Cut-offs were established using ROC curve analysis

The sensitivity and specificity of maspin was evaluated using five cut-offs (thresholds) for positivity as follows: 5% positive cells of any staining intensity (5% cut-off); 10% positive cells of any staining intensity (10% cut-off); 20% positive cells of any staining intensity (20% cut-off); moderate or strong staining of any cells (+2/+3); and Histoscore equal to or more than 20 (HS20). Three of these cut-offs were based on percentage of positive cells and identified by ROC curve analysis. The sensitivity of maspin was plotted against (1 – specificity), and corresponding ROC curves with coordinates were generated. The area under the curve for maspin was 0.99 (0.97-1.00, 95% CI).

Based on percentage of positive cells in the tumour compared with normal cores, ROC curve analysis allowed us to assess potential cut-offs, from 5% positive cells to 95% positive cells, with their corresponding sensitivity and

specificity values for maspin (Table 5.4, Figure 5.4). Three optimal cut-offs; 5%; 10% or 20% of positive cells of any staining intensity were selected based on their sensitivity and specificity values. The fourth cut-off was based on moderate to strong staining intensity (+2/+3 staining) in any of the cells. Interestingly, cases with +2/+3 staining for maspin have more than 20% cells positive. Indeed patients with +2/+3 staining have only 2 cases with less than 50% of cells positive for maspin. The fifth cut-off was based on a Histoscore value of 20 (HS20), and was selected from ROC curve analysis (Figure 5.5).

### 5.3.5 Maspin achieves optimum sensitivity and specificity for all five cut-offs

The sensitivities and specificities of maspin were calculated using five cut-offs, as shown in Figure 5.6. Maspin expression appears to show reasonably high sensitivity and specificity for all cut-offs including the cut-off based on +2/+3 staining, which resulted in sensitivity of 91% and 100% specificity. The 20% positive cells cut-off achieves marginally better sensitivity (96%) and specificity (99%) values compared with other cut-offs. The sensitivity and specificity values achieved for maspin across all cut-offs are high and are better than the biomarkers investigated in the previous chapters for some cut-offs. For example for +2/+3 cut-off the sensitivity of KOC is 67%, mesothelin is 55%, S100P is 81% but for maspin it is 91%. Maspin might potentially be used along with the biomarkers validated in previous chapter due to its high specificity and sensitivity.

## 5.4 Discussion

**Maspin was investigated in a relatively large cohort of PB patients (n=137).** The expression level of maspin is high in tumour tissue compared to normal tissue. The diagnostic accuracy (sensitivity and specificity) of maspin is high and is comparable to the biomarkers investigated in the previous chapter. Maspin as a diagnostic biomarker for improving the diagnosis of PDAC has been investigated in surgical (164, 167) and cytology specimens (167, 171, 235) but like other biomarkers further clinical translation has not yet occurred.

**Let us apply the six main reasons delaying clinical translation of biomarkers in PDAC (discussed in chapter 1) to maspin and our approach to address them. First,** maspin has been previously investigated and the diagnostic evidence of maspin was identified in the meta-analysis. It was identified as a high ranking candidate and hence it was investigated in our surgical and cytology cohort. **Second,** the sample size of 'NBJ TMAs' used for maspin staining is very good compared to published diagnostic IHC literature for PBA (164, 167, 171, 237). **Third,** the cut-offs determined in the previous chapters namely 5%, 10%, 20%, +2/+3 and HS20 were identified by ROC curve analysis for maspin. These cut-offs were then validated for maspin with largely similar diagnostic patterns to the other biomarkers. The use of 5% cut-off for maspin has been reported (171, 232) but other cut-offs are novel for maspin in PBA literature. **Fourth,** maspin has been investigated singly in this chapter but it is highly specific (99% for 20% cut-off) and sensitive (96% for 20% cut-off). Therefore, hypothetically it can enter into the panel of biomarkers identified in the previous chapter as a potential candidate in future validation studies. **Fifth,** the optimum IHC conditions from the literature were identified and then optimised in our laboratory as outlined in chapter 2 (section 2.2.4.2). The IHC conditions that we finally used to stain the TMAs achieved optimum staining in tumour tissue but very less staining in the normal tissue. **Sixth,** maspin was independently validated in this chapter as a diagnostic IHC biomarker and could potentially be used with other biomarkers as a panel in further validation studies. Therefore, a panel comprising of KOC, mesothelin, S100P and maspin might potentially be used for investigation and validation in cytology samples. Finally, the expression of maspin was similar in PDAC and CCC (Table 5.3): our results therefore agree with the previous literature (237, 276, 291, 292).

**Conclusion:** Our results demonstrate that maspin is capable of categorising PB malignancy with high diagnostic accuracy in resection specimens. We plan to investigate maspin along-with KOC, S100P and mesothelin (from chapter 4) as a panel in archival cytology samples. As part of a future plan, this panel will then be investigated in prospectively collected cytology samples. As an adjunct to cytology, this panel has the potential to augment the categorisation for challenging diagnostic cases in routine clinical practice.

**What this study adds:** This study independently validated maspin identified in the meta-analysis in a surgical cohort of patients with PBA. Histoscore was used as a scoring system and high diagnostic sensitivity and specificity was observed for all five potential diagnostic cut-offs. Maspin showed similar expression in PDAC and CCC.

**What Next:** We now have identified a potential panel of markers, namely KOC, S100P, mesothelin and maspin for investigation in cytology samples. Diagnostic cut-offs identified through ROC curve analysis achieved very good sensitivity and specificity for the diagnosis of PBA. However, the question of how good these cut-offs are in achieving observer agreement between pathologists is still to be answered. The project thus evolved and a need to investigate observer variations for cut-offs emerged. Therefore, in the next chapter diagnostic cut-offs will be investigated for observer agreement between practising pathologists.

**Table 5.1: Summary statistics of maspin expression comparing PBA with normal tissue.**

| Biomarkers | | Pancreaticobiliary Adenocarcinoma | Normal tissue | P value |
|---|---|---|---|---|
| **Maspin** | | | | |
| ***Positivity**** | Mean | 90% | 0.4% | <0.0001 |
| | Median | 100% | 0% | |
| ***Histoscore*** | Mean | 215 | 0.4 | <0.0001 |
| | Median | 240 | 0 | |

**Note:** *Positivity (percentage of positive cells of any staining intensity in tumour and normal tissue); P value (Shows the statistical significance of the difference in expression of maspin in tumour vs. normal tissue); Positivity range (0%-100%), Histoscore range (0-300).

**Table 5.2: Summary statistics of maspin expression comparing PBA with normal ducts and normal ducts & acini together**

| Biomarkers | | Normal ducts | Normal ducts & acini |
|---|---|---|---|
| **Maspin** | | | |
| ***Positivity**** | Mean | 0.4% | 0.4% |
| | Median | 0% | 0% |
| ***Histoscore*** | Mean | 0.4 | 0.4 |
| | Median | 0 | 0 |

**Note: ***Positivity (percentage of positive cells of any staining intensity in tumour and normal tissue); Positivity range (0%-100%), Histoscore range (0-300).

**Table 5.3: Summary statistics of maspin expression comparing PDAC with CCC**

| Biomarkers | | Pancreatic ductal adenocarcinoma | Cholangiocarcinoma | P value |
|---|---|---|---|---|
| **Maspin** | | | | |
| ***Positivity**** | Mean | 90% | 92% | 0.18 |
| | Median | 100% | 100% | |
| ***Histoscore*** | Mean | 215 | 220 | 0.53 |
| | Median | 235 | 240 | |

**Note:** *Positivity (percentage of positive cells with any staining intensity in tumour tissue); P value, independent sample t test (shows the statistical significance of the difference in expression of maspin in pancreatic ductal adenocarcinoma vs. cholangiocarcinoma); Positivity range (0%-100%), Histoscore range (0-300).

**Table 5.4: Cut-offs resulting from ROC curve analysis based on the percentage of positive cells in tumour and normal cases for maspin.**

| Cut-offs* | Maspin | |
|---|---|---|
| | **Sensitivity** | **Specificity** |
| 5% | 98 | 95 |
| 10% | 98 | 98 |
| 20% | 96 | 99 |
| 35% | 95 | 100 |
| 45% | 94 | 100 |
| 55% | 92 | 100 |
| 65% | 91 | 100 |
| 70% | 90 | 100 |
| 80% | 85 | 100 |
| 90% | 74 | 100 |
| 95% | 65 | 100 |

**Note: *** These cut-offs are based on percentage positivity [percentage of positive cells (0%-100%) of any staining intensity (weak, moderate and strong) in tumour and normal tissues].

**Figure 5.1: Maspin immunostaining**

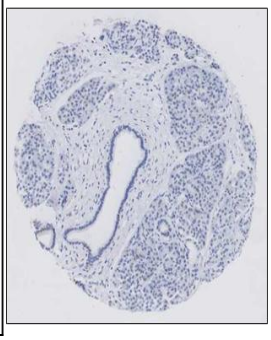
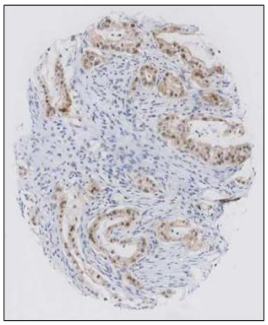
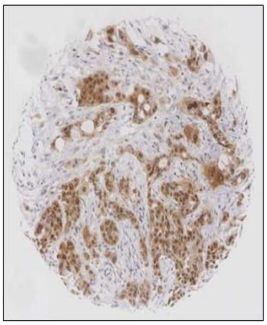
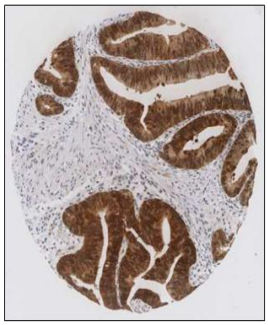
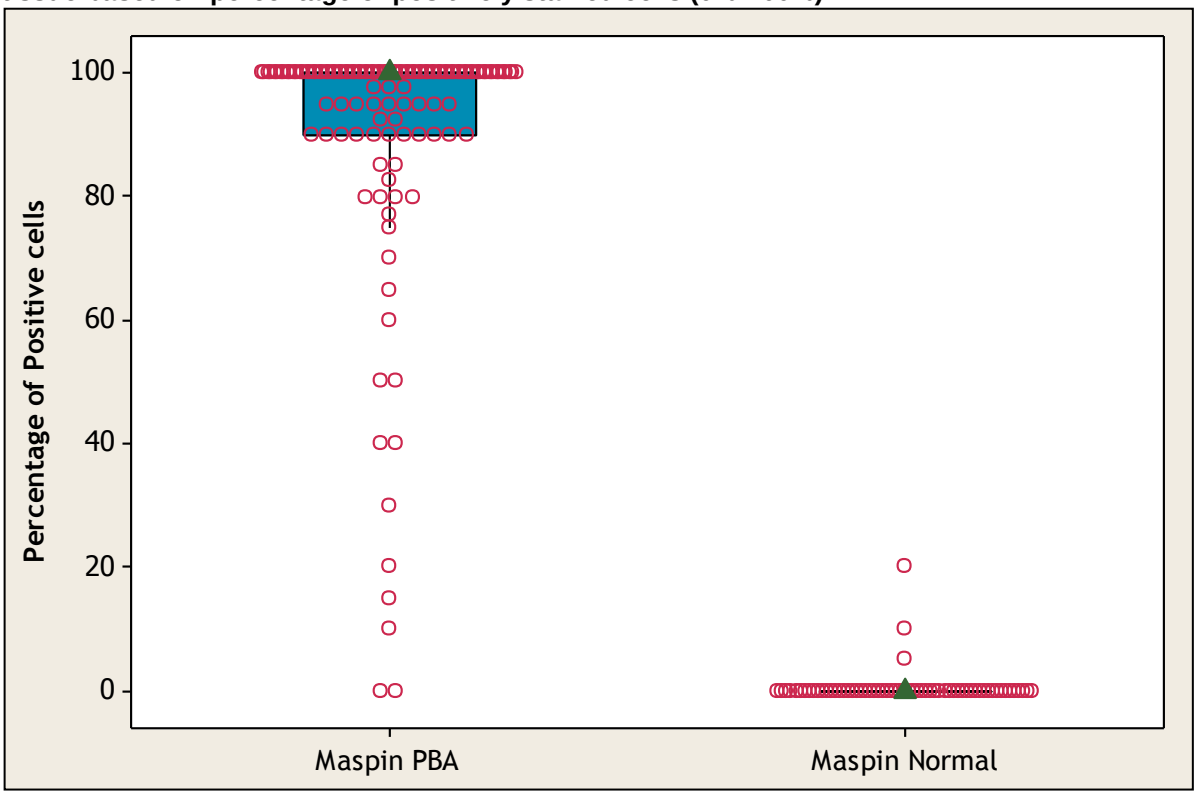
| Biomarker | Normal tissue | Weak Staining Tumour | Moderate Staining Tumour | Strong Staining Tumour |
|---|---|---|---|---|
| Maspin |  |  |  |  |

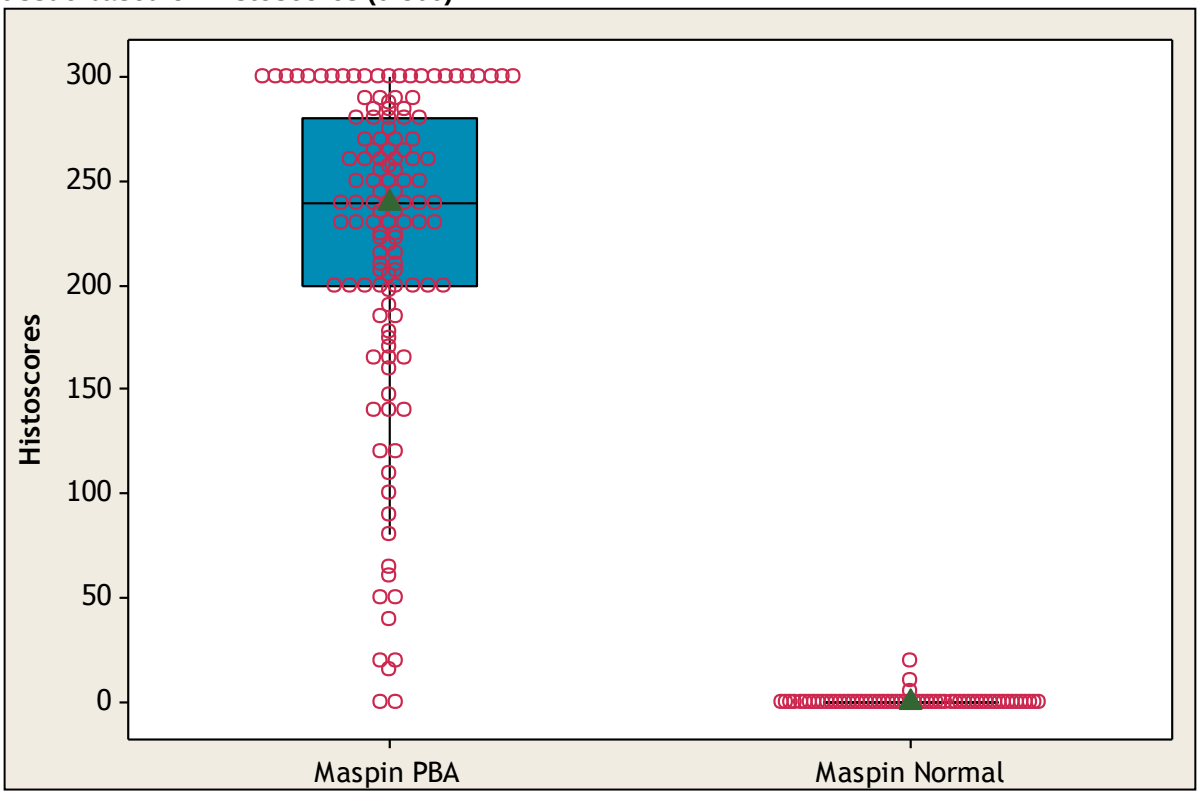
**Figure Legend:** Representative images of staining for maspin in normal tissue and range of staining intensities (weak, moderate and strong) in tumour tissue from tissue microarray cores. The staining is epithelial and is both cytoplasmic and nuclear (but for diagnostic purposes only cytoplasmic staining was considered. Normal ducts and acini are negative for maspin but varying proportion of positivity is noticed in tumour tissue.

**Figure 5.2: Boxplots comparing the expression of maspin in PBA compared to normal tissue based on percentage of positively stained cells (0%-100%)**




**Symbols explained:** Red Circles are individual data symbol. Green triangle is median data symbol

**Figure 5.3: Boxplots comparing the expression of maspin in PBA compared to normal tissue based on histoscores (0-300)**




**Symbols explained:** Red Circles are individual data symbol. Green triangle is median data symbol

**Figure 5.4: Receiver operating characteristics (ROC) curves for maspin based on positive percentage of cells**

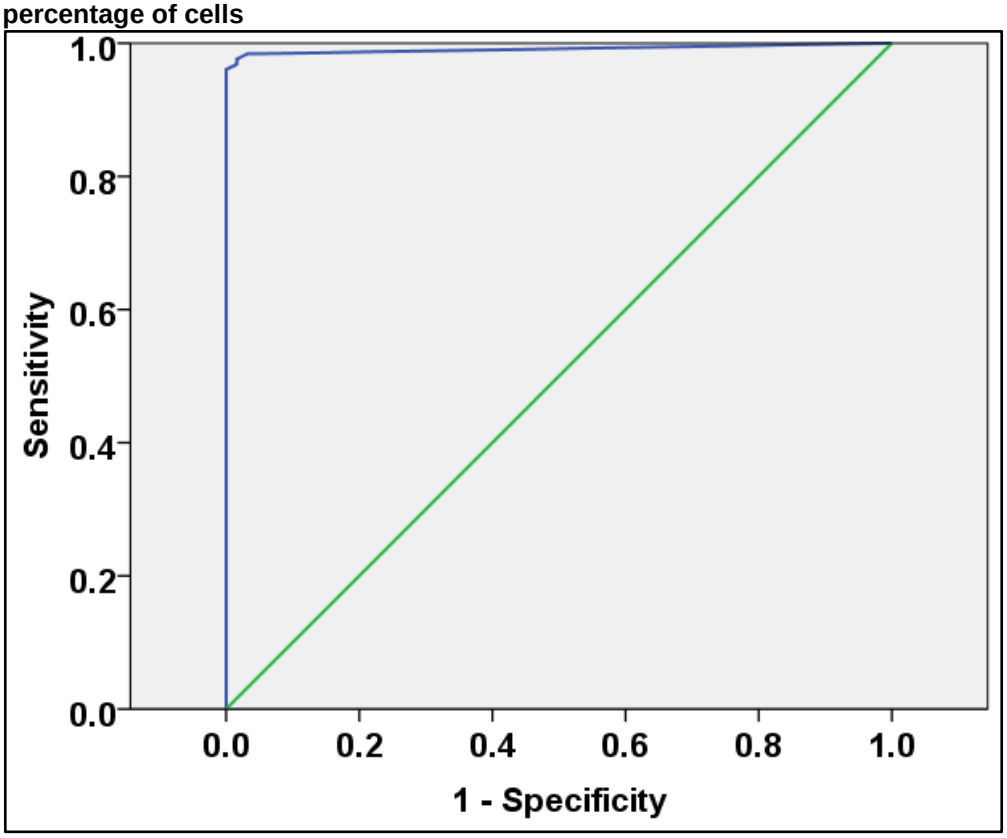



**Figure Legend:** ROC curves based on percentage of positive cells (0%-100%) for any staining intensity (weak, moderate or strong), in tumour and normal cases, for maspin.

**Figure 5.5: Receiver operating characteristic (ROC) curves for maspin based on Histoscores**

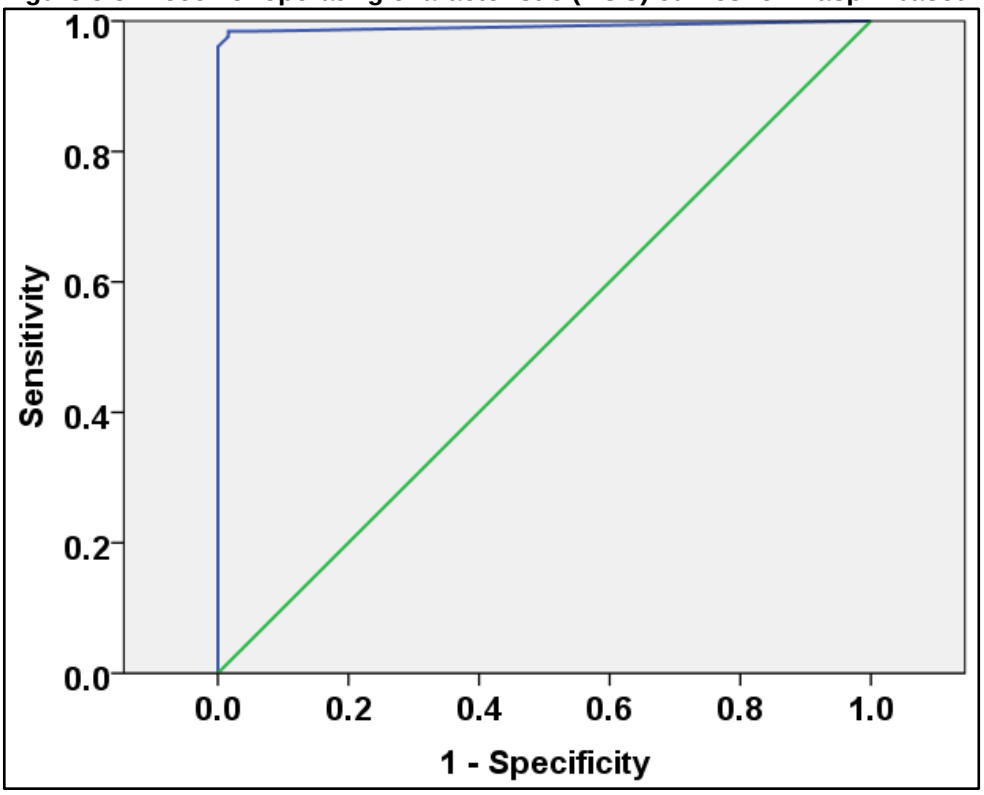



**Figure Legend:** ROC curves based on Histoscore (0-300), in tumour and normal cases, for maspin.

**Figure 5.6: Sensitivity and specificity analysis of maspin based on five cut-offs**

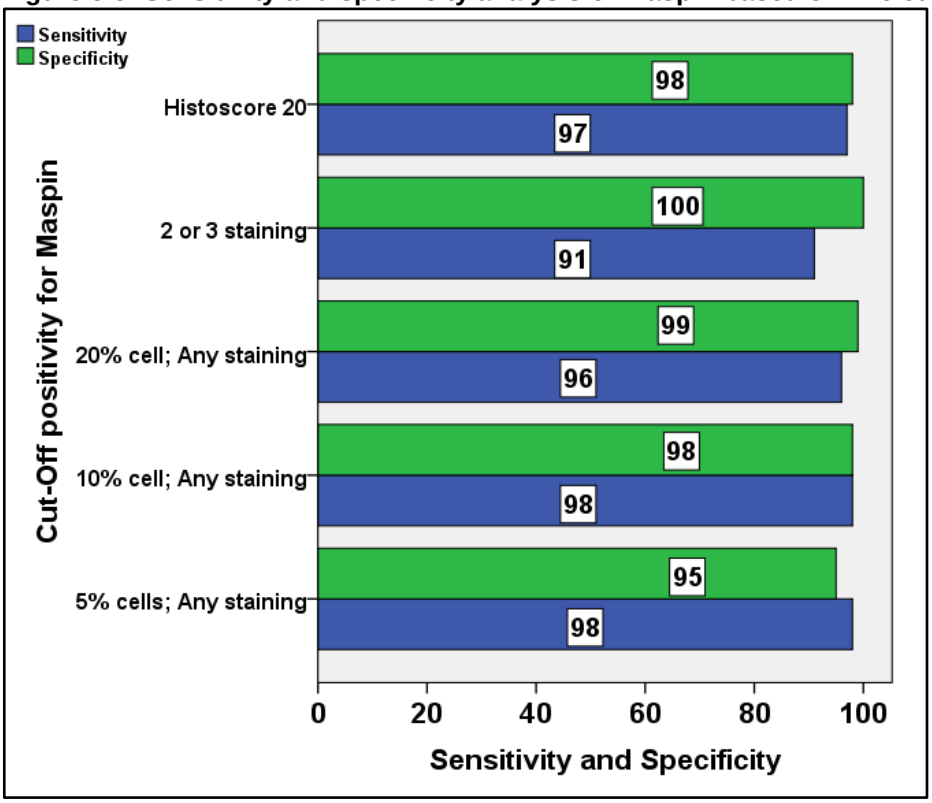



**Figure Legend:** Sensitivity and specificity analysis of maspin for the diagnosis of pancreatico-biliary adenocarcinoma compared to normal tissue, based on five cut-offs for positivity: 5% positive cells of any staining intensity; 10% positive cells of any staining intensity; 20% positive cells of any staining intensity; 2 OR 3 intensity i.e. moderate or strong staining of cells; and Histoscore 20.

# 6 Investigating various thresholds as immunohistochemistry cut-offs for observer agreement

## 6.1 Chapter Summary

Clinical translation of immunohistochemistry (IHC) biomarkers requires a reliable and reproducible cut-off for interpretation of immunostaining. Most of the IHC biomarker research focuses on the clinical relevance (diagnostic or prognostic utility) of cut-offs with less emphasis on observer agreement using these cut-offs. We identified three cut-offs from our TMA work and literature namely, 10% positive epithelial cells (10% hereafter), 20% positive epithelial cells (20% hereafter) and moderate to strong staining intensity (+2/+3 hereafter) for investigating observer agreement. The aim was to establish consensus based cut-off(s) that could potentially be used by pathologists.

A series of 36 IHC images of microarray cores for four IHC biomarkers with variable staining intensity and percentage of positive cells was used for investigating inter- and intra-observer agreement. Seven pathologists participated in the study and they scored the immunostaining of each image for the three cut-offs. Kappa statistic was used to assess the strength of agreement for each cut-off.

The inter-observer agreement between all seven pathologists using the three cut-offs was reasonably good. A good agreement was observed for experienced pathologists using 10% cut-offs and the agreement was statistically higher than junior pathologists ($p=0.02$). In addition, the mean intra-observer agreement for all seven pathologists using the three cut-offs was reasonably good. For all three cut-offs a positive correlation was observed with perceived ease of interpretations ($p<0.0001$ for 10% cut-off, $p=0.001$ for +2/+3 cut-off and $p=0.004$ for 20% cut-off). Finally, cytoplasmic only staining achieved higher agreement using all three cut-offs than cytoplasmic/nuclear staining and cytoplasmic/membranous staining.

All three cut-offs investigated achieve reasonable strength of agreement modestly decreasing inter and intra-observer variability in IHC interpretation but 10% is slightly better than 20% and +2/+3 cut-offs. These cut-offs have previously been used in cancer pathology and we have provided evidence that they are reproducible between practising pathologists.

## 6.2 Introduction

The use of immunohistochemistry (IHC) biomarkers for clinical decision making is an important research field with a significant translational potential. A multitude of biomarkers for a variety of cancers is available for clinical translation and an enormous literature exists on novel biomarker discovery, but only a minority is used for patient care. Amongst other reasons, one barrier for clinical translation of biomarkers is the lack of a standardised cut-off or threshold for interpretation of IHC (189, 190).

Evaluation of immunostaining is important in translational studies assessing biomarker expression for diagnostic, prognostic or predictive purposes. Biomarker expression reported in the literature is usually on a continuous or ordinal scale but for meaningful clinical use it is usually dichotomised and a cut-off is established for assigning a patient into either positive/negative category or high/low expression category (318). In addition, for some biomarkers two or more categories may be required for example the use of the 'Allred score' for estrogen receptor positivity (161). For clinical translation two issues revolve around a standardised cut-off for IHC biomarkers. One is the identification of an appropriate cut-off that provides suitable sensitivity/specificity for diagnostic biomarkers and stratifies patients based on survival and response to treatment for prognostic and predictive biomarkers respectively. The other issue is to assess the inter- and intra-observer agreement in the interpretation of cut-off. The former can be addressed using a receiver operating characteristics (ROC) curve that could help to identify an appropriate cut-off (189, 319). The latter issue can be answered by assessing the level of agreement between practising pathologists (320-322).

There is no standardized cut-off for diagnostic IHC biomarkers. Most of the reported cut-offs are purposive that best fit cancer or normal groups. These cut-offs are based on the intensity of staining or percentage of positive cells or combination of both intensity and percentage in terms of immunoreactive scores, H scores and "quick" scores (264, 270, 321, 323-325). Two widely used cut-offs reported in the literature for IHC diagnostic biomarkers are positive/negative staining (e.g. p16/Ki-67 staining for the diagnosis of cervical intra-epithelial neoplasia III) and 10% positive cells (e.g. CEACAM6 differentiating

PDAC from non-neoplastic pancreas) (326-330). Other reported cut-offs are: 5% positive epithelial cells for maspin (differentiating PDAC from chronic pancreatitis) (171); more than 20% positive epithelial cells for HMGI(Y) (differentiating PDAC from non-neoplastic pancreas) (265); and more than 30% cells with uniform, intense membranous staining of invasive tumour cells for human epidermal growth factor receptor 2 (HER2) (positive HER2 staining in breast cancer) (331).

These scoring systems and cut-offs have been adopted for research purposes and some of them are in clinical practice but studies looking at their reproducibility between pathologists are minimal. A cut-off should both be clinically relevant and easily interpretable by pathologists. There is a tendency to focus more on the clinical relevance of the cut-off for a biomarker with less focus on the level of agreement between pathologists when they use it for scoring purposes (320, 332). Inter- and intra-observer variation of a cut-off is infrequently analysed despite the fact that it is recognised as a potential barrier to clinical translation.

We selected three cut-offs for investigation based on our TMA work and literature. These cut-offs are 10% positive epithelial cells (10% hereafter), 20% positive epithelial cells (20% hereafter) and moderate to strong staining intensity with any proportion of positive cells (+2/+3 hereafter) (156, 210, 322, 326, 329, 333-335). These cut-offs are clinically relevant achieving appropriate diagnostic and prognostic utility and we postulated that they are easily interpretable and reproducible amongst pathologists.

### 6.2.1 Aims

We have now recognized from the systematic review (chapter 3) that a variety of cut-offs exist and there is a lack of standard cut-offs for diagnostic IHC biomarkers investigated in PDAC. This project thus evolved and a need for a standardised cut-off for IHC biomarkers was appreciated. A standard cut-off has two main components. One is the ability to categorise patients into distinct groups and the second is providing evidence of observer agreement. The former, i.e. the ability of these cut-offs (10%, 20% and +2/+3) in the diagnosis of adenocarcinoma and excluding normal was assessed in chapters 4 and 5. The latter, i.e. observer agreement, is addressed in the current chapter. The aims

therefore were: to investigate the cut-offs (10%, 20% and +2/+3) for inter and intra-observer agreements; and to explore the factors influencing agreement between pathologists for IHC cut-offs.

## 6.3 Results

Seven practising pathologists (3 experienced and 4 junior) participated in this study and they were coded as A, B, C, D, E, F and G. A series of 36 IHC images from four biomarkers was carefully selected. A scoring sheet was provided to the pathologists with instructions to score images for the three cut-offs namely 10%, 20% and +2/+3. Inter-observer agreement of IHC interpretation was assessed in one session which was followed after three weeks by a second similar session for evaluating intra-observer agreement. All seven pathologists thus scored all 36 images for the three cut-offs in each of the two sessions. Taken together, 1512 evaluations were made in both sessions by the pathologists. The average time for interpretation of an image was roughly 30-45 seconds. The strength of agreement was assessed using kappa (κ) scores (Table 6.1) which are based on the Landis and Koch model (226). The κ score between 0.21-0.40 is regarded as 'fair' agreement, κ score between 0.41-0.60 is regarded as 'moderate' agreement, κ score between 0.61-0.80 is regarded as 'substantial agreement' and κ score between 0.81-1.00 is regarded as 'almost perfect' agreement (226).

Results are divided into four parts: inter-observer agreement; perceived ease of scoring; intra-observer agreement; and staining of the different cellular compartments. The inter-observer part is further subdivided based on the pairwise comparison of pathologists into: all seven pathologists; experienced pathologists versus junior pathologists.

### 6.3.1 Inter-observer agreement

The aim was to determine the level of agreement between the pathologists for the three cut-offs and to assess whether one cut-off achieves better agreement than the others.

#### 6.3.1.1 The inter-observer agreement between all seven pathologists is reasonably good for all three cut-offs

This analysis was carried out from the scoring results of all seven pathologists. Pairwise comparisons of the interpretation of immunostaining among pathologists were used to generate κ scores. As an example, the κ score inter-observer agreement between pathologist A and B using the 10% cut-off was 0.80 (substantial agreement, P<0.001) (Table 6.2). Subsequently, 21 such inter-observer (AB, AC, AD and so on…) κ scores were generated for each of the three cut-offs. Finally a mean inter-observer κ score for each cut-off was used as a measure of strength of agreement between pathologists (Table 6.2).

The mean inter-observer κ scores were 0.64, 0.59 and 0.62 for 10%, 20% and +2/+3 cut-offs respectively (Table 6.2). The mean κ score agreement for 10% and +2/+3 cut-offs is in the 'substantial' agreement category and for 20% cut-off it is the 'moderate' agreement category. However, the κ score agreements between the three cut-offs were not statistically different from each other (Figure 6.1).

Figure 6.2 shows examples of IHC images used in this study. Images with low observer agreement have either weak staining intensity or the proportion of positively stained cells is lower compared to images with high level agreement. In fact, tissues with both strong staining intensity and a higher percentage of positive cells have higher agreement regardless of the biomarker and staining pattern.

In summary, the inter-observer agreements between all seven pathologists for the three cut-offs were reasonably good. In addition, the agreements for the cut-offs were not statistically different from each other. A good observer agreement using these cut-offs might facilitate their use for IHC biomarkers. Next I compared the agreements of experienced and junior pathologists.

#### 6.3.1.2 Experienced pathologists have statistically higher agreement for 10% cut-off than junior pathologists

The aim of this analysis was to compare the level of agreement amongst experienced pathologists with that of junior pathologists.

The mean inter-observer κ scores for experienced pathologists were 0.81, 0.70 and 0.55 for 10%, 20% and +2/+3 cut-offs respectively (Table 6.3A). The mean inter-observer κ scores for junior pathologists were 0.61, 0.60 and 0.73 for 10%, 20% and +2/+3 cut-offs respectively (Table 6.3B). The agreement on 10% cut-off is statistically higher for the experienced pathologists than the junior pathologists (P=0.02, Mann-Whitney U test). However, no statistically significant difference between experienced and junior pathologists was observed for 20% and +2/+3 cut-offs.

In summary, a higher level of agreement was observed for experienced pathologists using 10% cut-off and this was statistically higher than junior pathologists.

### 6.3.2 A positive correlation was observed between all three cut-offs and the perceived ease of scoring

In order to determine the perceived ease of scoring using the three cut-offs, the pathologists were asked to categorise each individual core as easy or challenging to score on the scoring sheet. A positive correlation was observed between all three cut-offs and perceived ease of scoring ($p<.0001$). However, in a multivariate analysis the 10% cut-off ($\beta=0.41$, $p<0.001$) was more easily scored as compared to +2/+3 cut-off ($\beta=0.38$, $p=0.001$) or the 20% cut-off ($\beta=0.34$, $p=0.004$) (Table 6.4).

Interestingly, the pattern emerging from this correlation, that 10% is relatively more easily scored, followed by +2/+3 and 20%, supports the mean inter- and intra-observer κ scores for these cut-offs (Table 6.2 and Table 6.5).

### 6.3.3 Intra-observer agreements

The aim was to investigate the intra-observer reproducibility of these cut-offs and to determine if the reproducibility of one cut-off was better.

#### 6.3.3.1 The intra-observer agreement for all three cut-offs was high confirming the reproducibility of cut-offs

A pairwise comparison of the scoring and then re-scoring of the IHC images by the same pathologist was made at two sessions three weeks apart. Kappa scores

were generated for all seven pathologists (A-A, B-B, C-C, D-D, E-E, F-F and G-G) using the three cut-offs (Table 6.5). As an example, the κ score intra-observer agreement for pathologist A using the 10% cut-off was 0.76 ('substantial' agreement, $P<0.001$) (Table 6.5). Subsequently, seven such intra-observer agreements (AA, BB, CC, DD, EE, FF and GG) were generated for each cut-off. A mean intra-observer κ score for each cut-off was then used as a measure of strength of agreement (Table 6.5).

The mean intra-observer κ scores were 0.71, 0.60 and 0.73 for 10%, 20% and +2/+3 cut-offs respectively (Table 6.5). The κ score agreement for 10% and +2/+3 cut-offs is in the 'substantial' agreement category and for 20% cut-off it is the 'moderate' agreement category. However, the κ score agreements between all seven pathologists for the three cut-offs were not statistically different (Figure 6.3).

In summary, the intra-observer agreements for the three cut-offs were reasonably good. In addition, the agreements for the three cut-offs were not statistically different from each other. Thus a good intra-observer agreement confirms the reproducibility of these cut-offs by pathologists and again this supports their use for IHC biomarkers.

The inter- and intra-observer agreements follow the same pattern i.e. 'substantial' agreement for 10% and +2/+3 and 'moderate' agreement for 20% cut-off.

### 6.3.4 Scoring of different cellular compartments

The aim was to determine if the pathologists tend to have more agreement for a particular staining pattern (cytoplasmic, nuclear and membranous). Depending on the staining pattern in the images studied, three patterns emerged: cytoplasmic only staining, cytoplasmic and/or nuclear staining (CN), and cytoplasmic and/or membranous staining (CM). Therefore, for each cut-off, three scenarios were identified i.e. cytoplasmic only staining compared to CN, cytoplasmic only staining compared to CM and CN compared to CM.

#### 6.3.4.1 Cytoplasmic only staining achieved higher agreement using all three cut-offs than the other staining patterns

The level of agreement on different staining patterns (cytoplasmic only, CN and CM) was analysed for the three cut-offs. Overall, the mean κ scores for cytoplasmic only staining were higher than the other staining patterns. More specifically, a statistically higher agreement for cytoplasmic only staining was observed in the following scenarios: cytoplasmic compared to CN category using +2/+3 cut-off; and cytoplasmic compared to CM category using 20% and +2/+3 cut-offs.

Moreover, a statistically higher agreement for CN staining was observed in the following scenarios: CN compared to CM category using 20% and +2/+3 cut-offs. No statistically significant difference between different staining patterns was observed for 10% cut-off (Table 6.6).

In summary, there is more agreement for cytoplasmic only staining followed by CN and CM. Clearly, this suggests that cytoplasmic only staining is easier to interpret than cytoplasmic staining mixed with nuclear or membranous staining. Finally 10% cut-off is easy to interpret regardless of the staining compartment of cell.

## 6.4 Discussion

**Three IHC cut-offs, namely 10%, 20% and +2/+3 were assessed for observer agreement between pathologists.** All cut-offs showed good inter- and intra-observer agreement between pathologists. Similarly, all three cut-offs showed high correlation with perceived ease of scoring. Finally, the observer agreement for cytoplasmic only staining is higher than cytoplasmic/nuclear staining and cytoplasmic/membranous staining.

**Establishing a cut-off for biomarker assessment is an essential pre-requisite for clinical translation.** A wide range of cut-offs have been used for diagnostic, prognostic and predictive IHC biomarkers in research and clinical settings. The purpose of a cut-off for a diagnostic biomarker is to assign patients into positive or negative categories with reasonable sensitivity without compromising specificity (336). Based on the expression level for a candidate biomarker in cancer and normal tissue, a cut-off is established. A good diagnostic cut-off has a low probability of false positivity. This removes the possibility of assigning normal population to the diseased group. Similarly, it has a low probability of false negativity to avoid missing the diseased group (336). The purpose of a cut-off for a prognostic biomarker is to divide the population into categories of longer and shorter survival for the outcome. In research settings a cut-off based on percentage of positive tumour cells is mostly used (337, 338). Similarly, the aim of a cut-off for predictive biomarkers is to stratify patients into likely responders and non-responders to treatment and intervention (339).

**IHC cut-offs used for prognostic and predictive biomarkers have been investigated for observer agreement but such studies are limited for diagnostic biomarkers.** The cut-offs of 10% and 30% positive cells with strong membranous staining for HER2 have been investigated for reproducibility amongst pathologists (322). In addition, for estrogen receptor (ER) and progesterone receptor (PR), the continuous H-score (range 0-300) and categorical scores (negative: H-score<1, positive: H-score ≥1) have been investigated for inter-observer agreement (321). These cut-offs for HER2, PR and ER are clinically important and are used in day-to-day practice by pathologists.

Clinically relevant cut-offs are important for biomarker evaluation but the intra- and inter-observer agreements for cut-offs between pathologists may be required to fully explore their clinical utility. We sought to investigate three cut-offs i.e. 10%, 20% and +2/+3 with the hope that if evidence of their scoring reproducibility is provided, they could potentially help the clinical translation of IHC biomarkers. Interestingly, the purpose of cut-offs differ for different biomarkers but these three cut-offs have been used for diagnostic (S100P, pVHL, KIT, HMGI(Y), CK20, P53, Ki-67) (156, 166, 265, 329, 335), prognostic (Ki-67, p53) and predictive (APAF-1, EGFR) biomarkers (340-343). Therefore, providing the strength of agreement between pathologists for these three cut-offs has significant clinical potential.

**Inter-observer agreement between pathologists was used to elucidate the reliability of cut-offs.** A 'substantial' agreement was observed with overall mean κ scores of 0.64 and 0.62 for 10% and +2/+3 cut-offs respectively, whereas 'moderate' agreement with a κ score of 0.59 was observed for 20% cut-off. In a study comparing the 10% positivity with 30% positivity for HER2, the mean κ scores for inter-observer agreement were 0.49 for 10% positive cells and 0.54 for 30% positive cells (322). Clearly, the κ scores generated for the three cut-offs under investigation in our project are comparable to the k scores for HER2 which is already in clinical practice as a predictive biomarker. Moreover, studies looking at the inter-observer reproducibility in histopathology and the IHC literature have shown that κ scores more than 0.60 (substantial agreement) are regarded as a good level of agreement. In comparison, κ scores less than 0.40 (fair agreement) are regarded as an unacceptably low level of agreements for diagnostic purposes (344-348).

**Intra-observer agreement of the scoring and then re-scoring of the same image was used to assess reproducibility of the cut-offs.** Again a pattern similar to inter-observer agreement emerged with 'substantial' agreements for the 10% and +2/+3 cut-offs and 'moderate' agreement for 20% cut-off. However, the intra-observer agreements (0.71, 0.60 and 0.73) in the present study are higher than inter-observer agreements (0.64, 0.59 and 0.62) for the three cut-offs. This finding agrees with the previous literature that the intra-observer agreement is more than the inter-observer agreement. For example the intra-

observer agreement (k=0.85) is better than the inter-observer agreement (k=0.80) for PDX-1 IHC staining intensity in prostate cancer (349). In addition, the intra-observer agreement (k=0.78) is better than the inter-observer agreement (k=0.65) for evaluation of focal cortical dysplasia categories (350).

**Taking 10% positive cells as a cut-off has been used for a variety of IHC biomarkers in different cancer types.** These include S100P and XIAP in the differentiation of pancreatic cancer from non-neoplastic pancreatic tissue, and for a panel of napsin-A, Thyroid transcription factor 1, Cytokeratin 5, and P63 in differentiating adenocarcinoma from squamous cell carcinoma of the lung (157, 158). Moreover, 10% cut-off is prognostic in breast cancer for a panel of Ki67 and p53, predictive of event-free survival in stage II colon cancer for VEGF and is predictive in rectal tumours treated with preoperative, high-dose-rate brachytherapy for APAF-1 (341, 343, 351). The use of a 10% cut-off in other areas of pathology means that the more experienced pathologists in the present study will have already had experience in applying this cut-off, which is a possible explanation for why they have a higher agreement than junior pathologists. Studies have attempted to show the reproducibility of the 10% cut-off and the κ scores achieved in the current study (0.64, substantial agreement) is similar to the k scores (0.57-0.77, moderate to substantial) in the reported literature (341, 352-354).

**The 20% positive staining cut-off has also been used for a variety of IHC biomarkers.** These include EGFR as a predictive biomarker in non-small cell lung cancer (390), Ki-67 as a prognostic biomarker in breast carcinoma (355) and NF-E2 in the differentiation of essential thrombocythemia from primary myelofibrosis (342, 356). However, studies investigating the variation in interpretation of this cut-off between pathologists are very limited. The current study investigated the 20% cut-off for observer agreement and our results suggested a good level of agreement.

**Moderate to strong staining intensity and any percentage of positive cells (+2/+3) as a cut-off has also been for IHC biomarkers.** These include CK20, P53, CK5/6, CD138, and Her2/Neu in the diagnosis of urothelial carcinoma in situ (382) and the use of claudin-4 to distinguish adenocarcinoma from malignant mesothelioma in effusion cytology (156, 357, 358). However, once again studies

observing the variation in interpretation of this cut-off between pathologists are very limited. Our results demonstrate that this cut-off is also reliable, reproducible and easy to score and it can be ranked second to 10% cut-off from the current study. Separating moderate staining from weak staining can sometime be difficult. Using only strong staining gives a high level of agreement but for diagnostic purposes if moderate staining is ignored the sensitivity will substantially drop for markers stained less intensely compared to biomarkers with high expression.

**The observer agreement was also assessed using staining in different cellular compartments.** Staining in only cytoplasmic compartment achieved higher agreements than other staining patterns. The k score (0.62) from membranous and/or cytoplasmic staining for 10% cut-off in the current study is comparable to the k score (0.49) for 10% cells with membranous staining in HER2 (322).

**The sample size was good and seven pathologists participated in the present study.** This number is comparable to the IHC biomarker and histopathology literature (4 to 7 participants) where observer agreement was investigated (283, 349, 359, 360). In addition, the participants in the current study were practising pathologists with variable levels of experience as compared to studies where either physicians (with no formal pathology experience) (349) or researchers with experience in IHC were recruited (283). Thus the results of this study provide good evidence on the use of cut-offs for IHC biomarkers.

**The strengths and limitations are:** The strengths of this project include: the participation by seven practising pathologists in two sessions; and investigation of three different cut-offs for the same image making this study time-effective and a broad study in the PDAC IHC literature for observer agreement. The limitations include: the relatively fewer number of images due to the time constraints imposed by the clinical work of the pathologists; and all pathologists were from the same institution; the aim was to carry out the study with all of the pathologists present at one session and this was achieved for the inter-observer part but for the intra-observer part we had to arrange an extra session. The other potential limitation might be that the images were shown as a PowerPoint presentation on screen rather than using a standard microscope.

**Conclusion:** In a day-to-day clinical practice pathologists need a scoring system and a cut-off that is reproducible and easy to use (189, 322). A wide range of cut-offs have been used for IHC biomarkers and they are becoming more important. We selected 10%, 20% and +2/+3 cut-offs that have good clinical potential from the already existing cut-offs. These three cut-offs are reliable and reproducible achieving a reasonably good agreement level between pathologists (when compared with the literature). They could facilitate translational biomarker studies and could potentially be used by scientists who are not trained pathologist but are involved in investigating potential IHC biomarkers. Further studies are required to assess these cut-offs with pathologists from different institutions and using a larger sample of images.

**What this study adds:** This is a wide study in PDAC biomarker research investigating cut-offs with implications for clinical practice. The translational significance of this work is providing a reasonable base for studies investigating biomarkers. A biomarker achieving diagnostic, prognostic and/or predictive significance with any of the three cut-offs might have a good translational potential as it might be easily interpreted by pathologists.

**What next:** We have validated biomarkers in surgical specimens, determined cut-offs achieving better diagnostic sensitivity and specificity and we have validated cut-offs for observer agreement. The next step is to investigate the previously investigated biomarkers (KOC, maspin, mesothelin and S100P) as a panel in archival cytology samples using the identified and validated cut-offs.

**Table 6.1: Explanation of the categories of k scores with colour codes**

| | | |
|---|---|---|
| 0.01-0.20 | Slight | |
| 0.21-0.40 | Fair | |
| 0.41-0.60 | Moderate | |
| 0.61-0.80 | Substantial | |
| 0.81-1 | Almost Perfect | |

**Note:** These colour codes will be used for all subsequent figures and tables.

**Table 6.2: Pairwise k scores of inter-observer agreements between pathologists for the three cut-offs.**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **10% Cut-Off** | | | | | | | |
| **Observers** | | | | | | | |
| **Observers** | **A** | **B** | **C** | **D** | **E** | **F** | **G** |
| **A** | | **0.8** | **0.82** | 0.62 | 0.47 | 0.72 | **0.89** |
| **B** | | | **0.82** | 0.62 | 0.47 | 0.53 | **0.89** |
| **C** | | | | 0.48 | 0.36 | 0.58 | 0.72 |
| **D** | | | | | 0.48 | 0.55 | 0.72 |
| **E** | | | | | | 0.74 | 0.54 |
| **F** | | | | | | | 0.6 |
| **G** | | | | | | | |
| **Mean k score** | **0.64 (95% CI, 0.57-0.70)** | | | | | | |
| **20% Cut-Off** | | | | | | | |
| **Observers** | | | | | | | |
| **Observers** | **A** | **B** | **C** | **D** | **E** | **F** | **G** |
| **A** | | **0.85** | 0.57 | 0.64 | 0.54 | 0.49 | **0.92** |
| **B** | | | 0.7 | 0.53 | 0.43 | 0.38 | 0.77 |
| **C** | | | | 0.64 | 0.42 | 0.48 | 0.51 |
| **D** | | | | | 0.62 | **0.82** | 0.56 |
| **E** | | | | | | 0.71 | 0.46 |
| **F** | | | | | | | 0.42 |
| **G** | | | | | | | |
| **Mean k score** | **0.59 (95% CI, 0.52-0.66)** | | | | | | |
| **+2/+3 Cut-Off** | | | | | | | |
| **Observers** | | | | | | | |
| **Observers** | **A** | **B** | **C** | **D** | **E** | **F** | **G** |
| **A** | | 0.75 | 0.42 | 0.6 | 0.55 | 0.46 | 0.65 |
| **B** | | | 0.48 | 0.54 | 0.49 | 0.52 | 0.59 |
| **C** | | | | 0.55 | 0.61 | 0.72 | 0.61 |
| **D** | | | | | 0.58 | 0.6 | 0.7 |
| **E** | | | | | | **0.88** | **0.88** |
| **F** | | | | | | | 0.77 |
| **G** | | | | | | | |
| **Mean k score** | **0.62 (95% CI, 0.56-0.67)** | | | | | | |

**Note:** Comparison of pairwise k scores with colour codes between pathologists (A-G) in the evaluation of immunohistochemistry using 10%, 20% and +2/+3 cut-offs. 21 inter-observer pairwise k scores (AB, AC, AD and so on……) were generated for all three cut-offs and are shown with mean k score and 95% CI separately for each cut-off**.**

**Table 6.3: Pairwise k scores of inter-observer agreements between experienced and junior pathologists for the three cut-offs.**

**A**

**Experienced pathologists**

**10% Cut-Off**

| | | Observers | | |
|---|---|---|---|---|
| | | **A** | **B** | **C** |
| **Observers** | **A** | | **0.80** | **0.82** |
| | **B** | | | **0.82** |
| | | | | |
| | | | | |
| | Mean K score | | 0.81 (Range: 0.80-0.82) | |

**20% Cut-Off**

| | | Observers | | |
|---|---|---|---|---|
| | | **A** | **B** | **C** |
| **Observers** | **A** | | **0.85** | 0.57 |
| | **B** | | | 0.70 |
| | | | | |
| | | | | |
| | Mean K score | | 0.71 (Range: 0.57-0.85) | |

**+2/+3 Cut-Off**

| | | Observers | | |
|---|---|---|---|---|
| | | **A** | **B** | **C** |
| **Observers** | **A** | | 0.75 | 0.42 |
| | **B** | | | 0.48 |
| | | | | |
| | | | | |
| | Mean K score | | 0.55 (Range: 0.42-0.75) | |

**B**

**Junior pathologists**

**10% Cut-Off**

| | | Observers | | | |
|---|---|---|---|---|---|
| | | **D** | **E** | **F** | **G** |
| **Observers** | **D** | | 0.48 | 0.55 | 0.72 |
| | **E** | | | 0.74 | 0.54 |
| | **F** | | | | 0.60 |
| | **G** | | | | |
| | Mean k score | | 0.61 (Range: 0.48-0.74) | | |

**20% Cut-Off**

| | | Observers | | | |
|---|---|---|---|---|---|
| | | **D** | **E** | **F** | **G** |
| **Observers** | **D** | | 0.62 | **0.82** | 0.56 |
| | **E** | | | 0.71 | 0.46 |
| | **F** | | | | 0.42 |
| | **G** | | | | |
| | Mean k score | | 0.60 (Range: 0.42-0.82) | | |

**+2/+3 Cut-Off**

| | | Observers | | | |
|---|---|---|---|---|---|
| | | **D** | **E** | **F** | **G** |
| **Observers** | **D** | | 0.58 | 0.60 | 0.7 |
| | **E** | | | **0.88** | **0.88** |
| | **F** | | | | 0.77 |
| | **G** | | | | |
| | Mean k score | | 0.73 (Range: 0.58-0.88) | | |

**Table 6.3: A,** Comparison of pairwise K scores with colour codes between experienced pathologists (A-C) in the evaluation of immunohistochemistry using 10%, 20% and +2/+3 cut-offs. Three inter-observer pairwise k scores (AB, AC and BC) were generated and are shown with mean k score and 95% CI separately for each cut-off.

**B,** Comparison of pairwise K scores with colour codes between junior pathologists (D-G) in the evaluation of immunohistochemistry using 10%, 20% and +2/+3 cut-offs. Six inter-observer pairwise k scores (DE, DF, DG, EF, EG and FG) were generated and are shown with mean k score and 95% CI separately for each cut-off**.**

**Table 6.4: Multivariable linear regression for 10%, 20% and +2/+3 cut-offs as predictor variables and perceived ease of interpretation as dependent variable.**

| Variables | Unstandardized Coefficients | | Standardized Coefficients | P value |
|---|---|---|---|---|
| | **B** | **Std. Error** | **Beta** | |
| (Constant) | -5.67 | 1.07 | | **<0.001** |
| 10% Cut-off | 0.71 | 0.16 | 0.41 | **<0.001** |
| +2/+3 Cut-off | .46 | 0.12 | 0.38 | **0.001** |
| 20% Cut-off | 0.51 | 0.16 | 0.34 | **0.004** |

**Note:** The 10%, 20% and +2/+3 are predictor variables i.e. they are variables that are predicting an outcome (the ease of interpretation). In this regression model ease of interpretation is a dependent variable i.e. a variable which "depends" on the predictor variable. The standardised beta coefficients were used as an estimate of association between predictor and dependent variable. The higher the beta coefficient the higher is the p-value significance and the stronger is the association between predictor and dependent variables. Beta coefficient in this model is highest (0.41) for 10% cut-off, followed by +2/+3 (0.38) and 20% (0.34). However, the p-value for all three cut-offs is significant showing a positive association with ease of interpretation.

**Table 6.5: Pairwise k scores of intra-observer agreements for pathologists for the three cut-offs.**

| K Scores | | | | | | |
|---|---|---|---|---|---|---|
| **Codes** | **10%** | **P value** | **20%** | **P value** | **+2/+3** | **P value** |
| A-A | 0.76 | 0.000 | 0.74 | 0.000 | 0.80 | 0.000 |
| B-B | **0.89** | 0.000 | 0.68 | 0.000 | 0.72 | 0.000 |
| C-C | **0.84** | 0.000 | 0.59 | 0.000 | 0.50 | 0.003 |
| D-D | 0.43 | 0.002 | 0.55 | 0.001 | 0.59 | 0.000 |
| E-E | 0.68 | 0.000 | 0.59 | 0.000 | **0.88** | 0.000 |
| F-F | 0.47 | 0.003 | 0.51 | 0.002 | **0.94** | 0.000 |
| G-G | **0.87** | 0.000 | 0.53 | 0.001 | 0.68 | 0.000 |
| **Mean (95% CI)** | **0.71 (0.53-0.88)** | | **0.60 (0.52-0.68)** | | **0.73 (0.59-0.87)** | |

**Note:** Pairwise k scores showing intra-observer reproducibility from scoring re-scoring (for example A-A) of all seven pathologists (A-G) in the evaluation of Immunohistochemistry using 10%, 20% and +2/+3 cut-offs.

**Table 6.6: Mean k scores with p values for staining of different cellular compartments.**

| | C vs. CN | | C vs. CM | | CN vs. CM | |
|---|---|---|---|---|---|---|
| **Cut-Offs** | **mean** | **p value*** | **mean** | **p value** | **mean** | **p value** |
| **10%** | 0.77 vs. o.71 | 0.380 | 0.77 vs. 0.64 | 0.150 | 0.71 vs. 0.64 | 0.500 |
| **20%** | 0.75 vs. 0.63 | 0.100 | 0.75 vs. 0.40 | <0.001 | 0.63 vs. 0.40 | 0.009 |
| **+2/+3** | 0.81 vs. 0.58 | 0.001 | 0.81 vs. 0.40 | <0.001 | 0.58 vs. 0.40 | 0.010 |

**Note:** *Paired sample t test

Abbreviations: C (cytoplasmic staining), CM (cytoplasmic and membranous staining) and CN (cytoplasmic and nuclear staining).

**Figure 6.1: Boxplots showing the distribution of inter-observer k scores for 10%, 20% and +2/+3 cut-offs.**

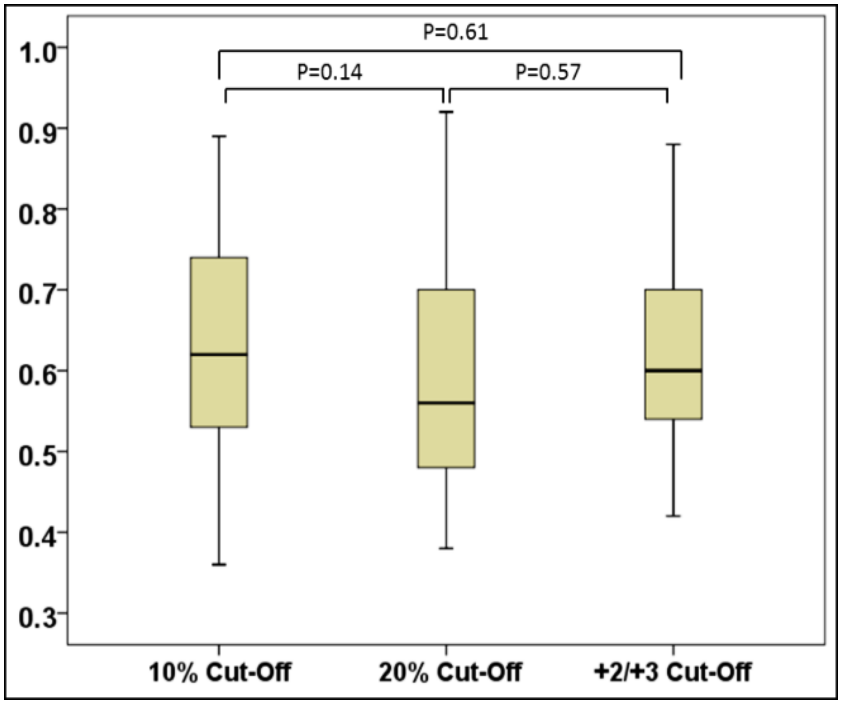



**Figure Legend:** For each cut-off 21 pairwise inter-observer k scores are shown in boxplots. As evident from boxplots the k scores for 10% cut-off appear higher than 20% and +2/+3 cut-0ffs. To investigate whether the k scores of one cut-off is statistically higher than the others, paired sample t test was used to compare the three cut-offs. Three possible comparisons can be made: 10% vs. 20%; 10% vs. +2/+3 and 20% vs. +2/+3. However, no statistically significant difference in the k scores was observed between the three cut-offs as shown by the p values.

**Figure 6.2: Representative images of high and low inter-observer agreement between all pathologists for 10%, 20% and +2/+3 cut-offs.**

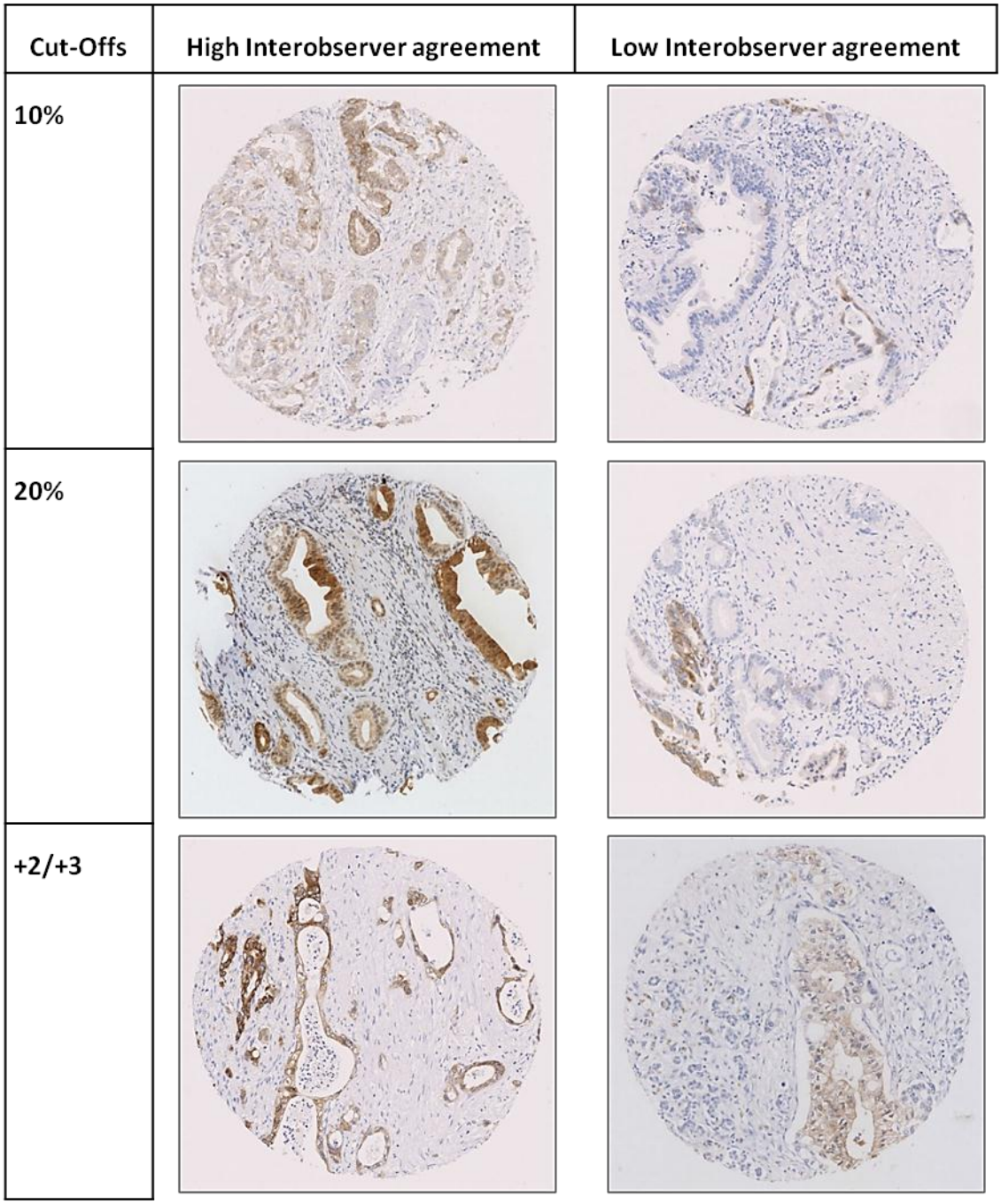



**Figure legend:** The high (left column grid) and low (right column grid) inter-observer agreement of IHC interpretation is shown for the three cut-offs. The staining for the three cut-offs was recorded only in the tumour epithelium. The high level agreement is attributed to the strong staining intensity and higher proportion of positive cells as illustrated in left column grid. All pathologists agreed on the images in the left column grid for all three cut-offs. However, there were differences in the number of pathologists agreeing on the images in the right column grid. For 10% cut-off (right upper image) 4/7 pathologists agreed, for 20% cut-off (right middle image) 5/7 pathologists agreed and for the +2/+3 cut-off (right lower image) 3/7 pathologists agreed.

**Figure 6.3: Boxplots showing the distribution of intra-observer k scores for 10%, 20% and +2/+3 cut-offs.**

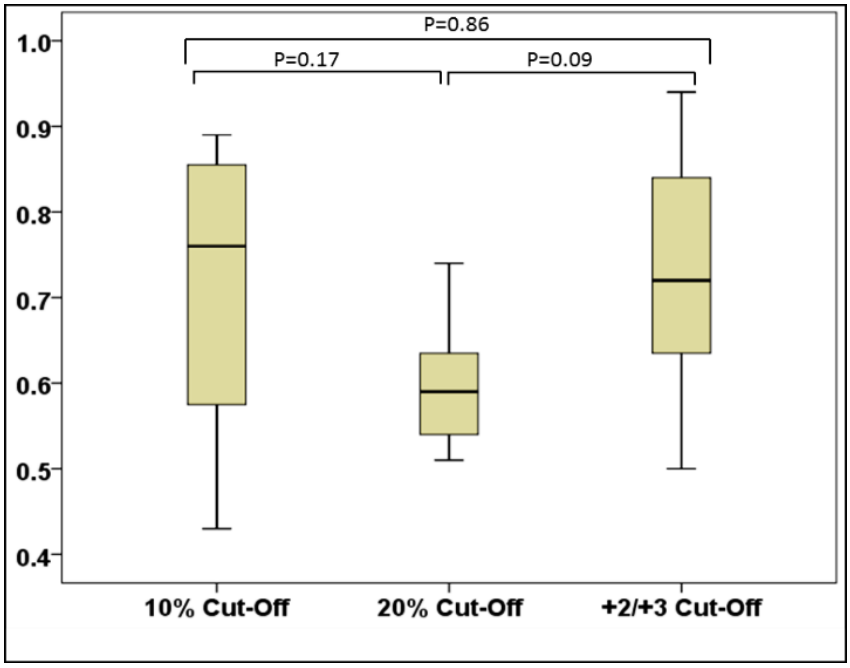



**Figure Legend:** For each cut-off seven pairwise intra-observer k scores are shown in boxplots. As evident from boxplots the k scores for 10% cut-off appear higher than 20% and +2/+3 cut-0ffs. To investigate whether the k scores of one cut-off is statistically higher than the others, Wilcoxon signed ranked test was used to compare the three cut-offs. Three possible comparisons can be made: 10% vs. 20%; 10% vs. +2/+3 and 20% vs. +2/+3. However, no statistically significant difference in the k scores was observed between the three cut-offs as shown by the p values.

# 7 A panel of KOC, maspin and mesothelin is a good working immunohistochemistry panel for PDAC diagnosis from cytology samples

## 7.1 Chapter Summary

Pre-treatment diagnosis of PDAC involves imaging followed by endoscopy with cytology. Diagnostic confirmation is important for patient management and is achieved from interpretation of cytology samples. However, cytology is not perfect and can be difficult, especially in chronic pancreatitis. Immunohistochemical (IHC) biomarkers could help but none is yet routinely used. IHC biomarkers validated in pancreatico-biliary (PB) TMAs (discussed in chapters 4 and 5) were used to stain the archival cytology samples in a pilot study. The aim was to generate a diagnostic panel that could potentially help as an adjunct to cytology in difficult to diagnose cases.

A cytology resource was generated from fresh Whipple's resections for the optimisation of IHC antibodies. Optimised antibodies for IHC were then used on archival cytology samples from 11 patients with PDAC and 10 patients with benign diagnoses. IHC was performed on an automated platform using antibodies against KOC, maspin and mesothelin. Tissue cores were scored for staining intensity and the proportion of tissue stained using the Histoscore method (range, 0-300). Sensitivity and specificity for individual biomarkers, as well as biomarker panels, were determined with different cut-offs for positivity and compared using summary receiver operating characteristic (ROC) curves.

The expression of all three biomarkers was high in PDAC cytology as compared to benign cytology with a mean Histoscore of 168 vs. 0 for KOC, 50 vs. 1.5 for maspin and 73 vs. 0.5 for mesothelin ($p<.001$). Five cut-offs, namely 5%, 10% or 20% positive cells, moderate-strong staining intensity and Histoscore 20 were carefully chosen for sensitivity/specificity analysis. Using 10% positive cells as a cut-off achieved higher sensitivity/specificity values: KOC 92%/100%; maspin 54%/100% and mesothelin 72%/100%. Analysis of a panel of KOC, S100P and mesothelin achieved 82% sensitivity and 100% specificity if at least 2 biomarkers were positive for the 10% cut-off; and 72% sensitivity and 100% specificity for the 20% cut-off.

A biomarker panel of KOC, maspin and mesothelin was found to be good panel with the 10% cut-off achieving high sensitivity and specificity in PDAC compared to benign diagnoses from cytology samples.

## 7.2 Introduction

Pre-treatment diagnosis is important for the management of patients with PDAC and is achieved from cytology samples. The difficulty of morphologic assessment of cytology samples to distinguish PDAC from non-neoplastic pancreas and the potential role of IHC biomarkers to improve the cytology diagnosis has been discussed in detail in chapter 1.

A systematic approach was adopted to address the unmet clinical need in the diagnosis of PDAC from cytology samples. A systematic review and meta-analysis was performed and five candidate diagnostic IHC biomarkers were selected for validation in our local cohort of resection specimens from patients with PBA. Of the five biomarkers tested, four appear to be specific and sensitive in tissue samples: KOC, maspin, S100P and mesothelin. These four candidate biomarkers will now be investigated in archival cytology samples to elucidate their diagnostic potential for clinical translation.

### 7.2.1 Aims

The aims of this chapter were as follows: to assess the sensitivity and specificity of biomarkers singly and in panels in archival cytology samples (PDAC and benign samples); and to assess the suitability of 'needling' cytology samples for IHC optimisation. This should enable us in determining a clinically useful diagnostic biomarker or panel of biomarkers that could potentially be applied in difficult to diagnose cytology cases.

## 7.3 Results

IHC on cytology samples was performed utilising an automated IHC platform for KOC, mesothelin, maspin and S100P. After immunostaining the scoring of cytology samples was performed using a standard light microscope. The scoring of cytology samples was performed by a pathologist (SB) and all subsequent sensitivity and specificity analyses are based on this scoring. A categorical scoring using 10% positive cells as a cut-off was performed by the same pathologist (SB) and by the author (AA) for inter-observer agreement. The scoring of cytology samples was performed blind to the final diagnosis.

The results of this chapter are divided into two main parts: the use of 'needling' cytology samples for antibody optimisation; and assessing the diagnostic sensitivity and specificity of biomarkers in the archival cytology samples from patients with PDAC and benign cytology.

## 7.3.1 The use of 'needling' cytology samples for antibody optimisation

A resource of normal (pancreatic, duodenal) and malignant (PDAC) cytology, in the form of "cell blocks", from Whipple's resections was developed. These samples were called 'needling' cytology samples. The purpose of this resource was four-fold: to stain these samples with same IHC parameters used for staining TMAs; to further optimise IHC for cytology samples to achieve potentially better staining; to observe biomarker expression in malignant and normal (pancreatic, duodenal) cytology samples; and to use these samples as positive controls for staining the archival cytology samples.

The 'needling' samples were first stained with 'IHC parameters optimised for TMAs' and then with the 'IHC parameters optimised for cytology'.

### 7.3.1.1 Staining the 'needling' samples with 'IHC parameters optimised for TMAs'

Ten 'needling' samples obtained from five Whipple specimens (5 samples from tumour and 5 from adjacent normal pancreas) were used for IHC staining for each biomarker (KOC, S100P, maspin and mesothelin). The corresponding tissue sections of these cytology samples were also used to compare IHC staining in cell and tissue specimens. The 'IHC parameters optimised for TMAs' were used for this IHC run (shown in Table 2.4, Chapter 2). The staining achieved is presented for the four biomarkers.

***PDAC 'needling' and tumour staining in corresponding tissue sections:*** The staining for **KOC** was cytoplasmic and present in 2 out of 5 PDAC 'needling' samples with Histoscore of 60 and 10 for the two samples. Overall the staining was weak in intensity and patchy. The staining in all five corresponding PDAC tissue sections was present as expected and was weak to strong in intensity and diffuse in PDAC (Figure 7.1).

The staining for **mesothelin** was cytoplasmic and membranous and for diagnostic purposes the presence of either cytoplasmic or membranous or both types of staining was considered. The staining was present in 2 out of 5 PDAC 'needling' samples with Histoscore of 80 and 70 for the two samples. Overall the staining was weak in intensity and diffuse. The staining in all five corresponding PDAC tissue sections was present as expected and was weak to strong in intensity and diffuse in PDAC (Figure 7.1).

The staining for **maspin** was cytoplasmic and nuclear but for diagnostic purposes only cytoplasmic staining was considered. The staining was present in 3 out of 5 PDAC 'needling' samples with Histoscores of 300, 60 and 90 for the three samples. Overall, the staining was weak to strong in intensity and diffuse but a significant amount of background staining was observed. The staining in all five corresponding PDAC tissue sections was present as expected and was moderate to strong in intensity and diffuse in PDAC (Figure 7.1).

No staining for **S100P** was observed in any of the PDAC 'needling' samples. However, the staining in all five corresponding PDAC tissue sections was present as expected and was moderate to strong in intensity and diffuse in PDAC (Figure 7.1).

***Normal pancreatic 'needling' and normal pancreas in corresponding resection specimens:*** For all four biomarkers, staining was absent from the normal pancreatic 'needling' samples and from normal tissues in tissue sections.

In summary, the purpose of this IHC phase was to verify the 'IHC parameters optimised for TMAs' using cytology samples. It appeared that the staining for maspin was good (but with a high background staining), staining for KOC and mesothelin was weak in intensity and less percentage of neoplastic cells showed positive staining and S100P showed no staining. Thus all four biomarkers might require further optimisation. The staining achieved in TMA sections has not fully reflected in cytology samples. The next section thus explains the optimisation of biomarkers for cytology specimens.

#### 7.3.1.2 Optimising the IHC parameters for four antibodies to achieve better staining using 'needling' cytology samples

IHC staining can sometimes be challenging to achieve due to technical issues. The available evidence on IHC parameters for these antibodies was reviewed in the PDAC cytology literature (shown in Table 2.2, Chapter 2). In assessing the PDAC IHC literature for these biomarkers, we identified other options used in two IHC parameters i.e. antigen retrieval and primary antibody dilutions. Hence these two parameters were reviewed.

The purpose of this IHC run on 'needling' samples was to optimise antibodies with the hope of achieving better staining. The aim for KOC and mesothelin was to achieve better staining in terms of both intensity and the proportion of positive cells. The aim for maspin was to reduce the background signal and the aim for S100P was to achieve any staining.

For each antibody two sets of five PDAC 'needling' samples were used to compare different IHC parameters. One set was used for staining the samples with 'IHC parameters optimised for TMAs' and the second set was used to vary antigen retrieval buffer for KOC and mesothelin, to vary antigen retrieval method for S100P and to vary antibody dilution for maspin. These were called 'IHC parameters optimised for cytology'. The details of these parameters are outlined in Tables 2.4, 2.5 and 2.6 in chapter 2.

***Comparing the staining achieved for various IHC parameters using 'needling' cytology samples:*** The issue of weak staining for KOC and mesothelin was addressed by changing the pH of antigen retrieval buffer solution.

For **KOC**, an improvement in IHC staining was observed with a mean Histoscore of 250 compared to 30 in malignant cells of all 'needling' samples for 'IHC parameters optimised for cytology' compared to 'IHC parameters optimised for TMAs'.

Similarly, for **mesothelin**, an improvement in IHC staining was observed with a mean Histoscore of 160 compared to 65 in malignant cells of all 'needling' samples. Thus by changing only one IHC parameter a significant improvement in staining was achieved for KOC and mesothelin (Figure 7.2).

For **maspin**, the issue of significant background staining was addressed by doubling the dilution of primary antibody. The mean Histoscore was 135 for 'IHC parameters optimised for cytology' compared to 150 for the 'IHC parameters optimised for TMAs'**.** Thus the sensitivity of maspin detection was not significantly affected but the background signal was decreased (Figure 7.2). For **S100P**, no staining was observed with varying the IHC parameters (both antibody dilution and antigen retrieval).

In summary, changing the antigen retrieval buffer for KOC and mesothelin significantly improved the staining in 'needling' samples. The background signal for maspin was decreased by doubling the antibody dilution without significantly affecting the antigen detection.

***Staining the normal pancreatic and normal duodenal 'needling' samples:*** The changes in IHC parameters might potentially affect the staining in normal tissue and hence the specificity of biomarkers. To address this issue normal pancreatic 'needling' samples were stained with the 'IHC parameters optimised for cytology'. Cytology samples may contain gastro-intestinal contaminants including normal duodenal epithelial cells and it would be of value to assess the biomarker expression in these cells. The normal duodenal 'needling' samples were therefore stained. Five normal pancreatic and five normal duodenal 'needling' samples were used for each of the four antibodies (KOC, S100P, maspin and mesothelin). No or focal (<5%) staining in all normal 'needling' samples was observed for all biomarkers. Hence changing the IHC parameters did not affect the specificity of these biomarkers providing more confidence for using of changed IHC parameters for staining the archival cytology samples.

The analysis of PDAC and normal 'needling' samples using 'IHC parameters optimised for cytology' achieved better staining and correlated well with the ultimate malignant versus benign diagnoses. The handling and processing of the 'needling' samples was performed as similarly as possible to how actual EUS-FNA cytology samples are collected and processed. We therefore considered that the IHC conditions that worked well on the 'needling' samples could be taken forward to stain archival cytology samples. The final optimised IHC conditions are summarised in Table 2.6, chapter 2. The optimisation was successful for

KOC, maspin and mesothelin and therefore these biomarkers will be taken forward for staining the archival cytology samples.

### 7.3.2 Immunohistochemical diagnostic biomarkers for pancreatic ductal adenocarcinoma - a pilot study on cytology samples.

The archival cytology samples were used to investigate the panel of KOC, maspin and mesothelin. 11 malignant cytology cell blocks (final diagnosis PDAC) and 10 benign cytology cell blocks were retrieved from the cytology archive of Southern General Hospital, Glasgow, UK. The detail of the selection process of the samples for this pilot project is outlined chapter 2. The cellularity of these samples was checked using the H&E section to ensure that malignant and benign samples contained adequate malignant and benign cells for testing. IHC was performed using 'IHC parameters optimised for cytology'.

The objectives of this pilot study were as follows: to assess the diagnostic sensitivity and specificity of KOC, maspin and mesothelin singly and in panel using archival cytology samples; to assess the utility of five cut-offs identified in TMA work but especially to validate the diagnostic performance of 10%, 20% and +2/+3 cut-offs in cytology samples; and finally to assess the observer agreement for 10% cut-off by two independent observers i.e. a pathologist (SB) and the author (AA).

The results of this section are subdivided into three main categories: assessing the expression of biomarkers in PDAC and benign cytology; using the cut-offs to calculate sensitivity and specificity singly and as a panel of biomarkers; and using the 10% cut-off as a binary categorical variable (positive/negative) to investigate inter-observer agreement.

#### 7.3.2.1 Expression of all biomarkers was epithelial and was significantly higher in PDAC than benign cytology

For each marker assessed in the cytology cohort, IHC staining was seen only in epithelial cells. As expected, KOC expression was observed in the cytoplasm; maspin was expressed in the cytoplasm and nucleus (but for diagnostic purposes cytoplasmic staining was scored), while mesothelin expression was cytoplasmic and membranous (Figure 7.3). In general, moderate to strong intensity of

staining was observed for KOC, while weak to strong staining was observed for maspin and mesothelin in malignant cytology samples. Moreover, for all three biomarkers significantly higher expression in malignant versus benign cytology was observed (Table 7.1, $p<0.003$, Independent sample t-test).

#### 7.3.2.2 Expression of KOC was higher than maspin and mesothelin in PDAC cytology samples as assessed by the percentage positivity and Histoscores

When scored simply as the proportion of positive staining cells, the mean percentage of positive cells in malignant cytology samples for KOC was 66%, for maspin it was 33% and for mesothelin it was 36%. A more comprehensive and semi-quantitative analysis was performed by utilizing a Histoscore as a scoring system. The mean Histoscore in malignant cytology for KOC was 168, while for maspin and mesothelin, the mean Histoscores were 50 and 73 respectively (Table 7.1).

The expression level for all three biomarkers in PDAC cytology samples was high making the interpretation of immunostaining easy. Next I assessed the expression levels of biomarkers in benign cytology samples to evaluate the specificity of biomarkers.

#### 7.3.2.3 Expression of all biomarkers in normal tissue was very low for all three biomarkers

The expression of the three biomarkers was very low (mean Histoscore for KOC 0, for maspin 1.5 and for mesothelin 0.5) in benign cytology samples (Table 7.1). It is therefore hoped that in a mixed population of tumour and normal cells, only positive staining would need to be taken into account for diagnostic purposes, as positivity is essentially associated only with tumour cells. Thus, IHC staining using these markers could facilitate the interpretation of cytology samples. The sensitivity and specificity of biomarkers using various cut-offs was next analysed.

#### 7.3.2.4 KOC is a better candidate diagnostic biomarker than maspin and mesothelin based on sensitivity and specificity values

The following number of PDAC cytology cases stained positive using 10% cut-off for KOC, maspin and mesothelin: for KOC 10 out of 11 PDAC cases; for maspin 6

out of 11 PDAC cases; and for mesothelin 8 out of 11 PDAC cases (Table 7.2). All PDAC cases (11 out of 11) were positively stained taking all biomarkers together as a panel. In contrast, all benign cytology cases (10 out of 10) were negative for all three biomarkers (Table 7.2).

The sensitivities and specificities of all three biomarkers were calculated using five cut-offs, as shown in Figure 7.4. KOC expression appears to show reasonably high sensitivity and specificity for all cut-offs. The 5% and 10% cut-offs achieve marginally better sensitivity (92%) and specificity (100%) values compared with other cut-offs for KOC (Figure 7.4A). Maspin appears to have similar sensitivity and specificity values for all cut-offs except the 5% cut-off that has low specificity (82%) (Figure 7.4B). Applying the five cuts-offs to the analysis of mesothelin expression resulted in different sensitivity and specificity values, however, the best combination was again achieved using the 10% cut-off, with 72% sensitivity and 100% specificity (Figure 7.4C).

In summary, KOC is a highly sensitive and specific biomarker for PDAC diagnosis. Maspin and mesothelin are highly specific biomarkers but their sensitivity in detecting malignancy in cytology samples is low. Now we know that these biomarkers are highly specific with very low expression in the normal pancreas but their sensitivity as a single candidate biomarker might not be clinically useful. To improve the sensitivity, panels of biomarkers were used. KOC, maspin and mesothelin were used in different panels for PDAC diagnosis in cytology samples as shown in the following section.

#### 7.3.2.5 A biomarker panel of KOC, maspin and mesothelin was found to be a good diagnostic panel for PDAC diagnosis from cytology

The sensitivity and specificity of panel of biomarkers were next assessed to find a suitable panel for PDAC diagnosis. The 10% and 20% cut-offs were selected for this investigation, based on their diagnostic performance.

Two panel approaches were used for sensitivity and specificity analysis. One approach categorised the patients into malignant and benign cytology if at least one biomarker in the panel was positive. For example for a 10% cut-off, the presence of 10% positive cells for one or more biomarker in a panel in the same

cytology sample is required to categorise the patient into a tumour category. Using this approach four panels of biomarkers were tested as shown in Table 7.3. These panels were: a panel comprising all three biomarkers; and three panels of two biomarkers in combination (KOC and mesothelin, KOC and maspin, maspin and mesothelin). A panel of all three biomarkers achieved 100% sensitivity and specificity for a 10% cut-off; and further achieved 90% sensitivity and 100% specificity for a 20% cut-off. These panels were compared by combined SROC curve, using both the 10% cut-off (Figure 7.5A) and 20% cut-offs (Figure 7.5B). The combined SROC curve showed that a panel of all three biomarkers is better than the other panels for both 10% and 20% cut-offs.

The second panel approach categorised the patients into malignant and benign cytology if at least two biomarkers in the panel were positive. For example for a 10% cut-off, the presence of 10% positive cells for two or more biomarkers in a panel in the same cytology sample is required to categorise the patient into a tumour category. A biomarker panel comprising all three biomarkers (KOC, maspin and mesothelin) was tested for sensitivity and specificity analysis using this approach. This panel achieved 82% sensitivity and 100% specificity for a 10% cut-off and 72% sensitivity and 100% specificity for a 20% cut-off (Table 7.3).

Figure 7.6 shows immunostaining of a PDAC cytology case to demonstrate the importance of the panel of biomarkers. This shows that the immunostaining may help in confirming the diagnosis in cases where the cytologist is unsure. Taken together, our results show that a panel comprising all three biomarkers could be used to improve the diagnosis of PDAC in difficult to diagnose cases.

#### 7.3.2.6 The 10% cut-off achieved a high inter-observer agreement for IHC interpretation in cytology samples

The 10% cut-off achieved a good combination of sensitivity and specificity for all three candidate biomarkers and panels in cytology samples. Therefore, I next analysed the inter-observer agreement for 10% cut-off as a categorical (positive/negative) variable. The cytology sample is called positive for malignancy when immunostaining is present in 10% or more cells. Conversely, the cytology sample is called negative for malignancy when immunostaining is present in less than 10% cells.

All malignant and benign cytology samples were scored by two observers. The inter-observer agreement for KOC using a 10% cut-off is in 'almost perfect' category (κ score=1.00). Similarly, the k scores using 10% cut-off for maspin and mesothelin are 0.92 and 0.91 respectively. Finally, the k score taking all three biomarkers as a panel for the 10% cut-off is 0.94 (Table 7.4).

Although the sample size is small, there is a high level agreement on the 10% cut-off between observers. This suggests that if the 10% cut-off is used in PDAC diagnostic cytology the observer variation is minimised. The use of the 10% cut-off is thus not only good in differentiating PDAC from benign but also achieves a good agreement level.

In summary, a panel of KOC, maspin and mesothelin using 10% cut-off is a good panel for the diagnosis of PDAC from cytology samples. This panel could potentially be used in difficult to diagnose cases as an adjunct to morphologic interpretation of cytology.

## 7.4 Discussion

**Biomarkers could help improve the diagnostic management of patients with PDAC**. We systematically addressed the issues delaying the translation of biomarkers from bench to clinic. After identification and validation in our surgical cohort, the utility of good biomarkers was investigated using archival cytology samples in a pilot study. A panel of KOC, maspin and mesothelin achieved high diagnostic sensitivity and specificity for PDAC diagnosis in cytology samples.

Biomarker research in cytology samples can present some challenges (361) and the discussion is thus broadly divided into two sections. The first section will discuss the issues related to assay development and IHC optimisation. The second section will discuss the utility of diagnostic IHC biomarkers in potentially improving PDAC diagnosis from cytology samples.

### 7.4.1 The role of 'needling' resource for biomarker optimisation

**Antibody optimisation is important to ensure good staining.** The 'needling' samples were processed into cell blocks and sections from these cell blocks were used for antibody optimisation and assay development. Cell block sections provide a good platform for IHC (362). The sample preparation of cell blocks differs from surgical specimens and this might affect the transfer of IHC conditions from tissue to cytology for some biomarkers. Briefly, surgical specimens are fixed in formalin for 24-48 hours, specimens are then cut, processed, embedded in wax and ultimately result in paraffin blocks for sectioning (363). In comparison, the fine needle aspiration specimens are put in PreservCyt solution, spun in a centrifuge, cells/tissue fragments settle to the bottom of the tube, the supernatant fluid is removed and then plasma and thrombin are added to make clot of cells or cell pellet. The cell pellet is then fixed in formalin overnight and with the help of wax is processed to paraffin cell blocks (167). Theoretically, the target antigen in the cell block sections might therefore need different antigen retrieval conditions than tissue sections to expose some antigens for immunostaining.

**A cytology resource from fresh Whipple resections was generated to allow antibody optimisation and to use as control sections.** Archival cytology samples are precious and given the limited availability, optimisation of IHC is challenging. In addition, the clinicopathologic information of all patients undergoing cell sample collection is maintained that requires a significant amount of time and cost. It is therefore vital to use the archival cytology samples for properly worked up biomarkers. A 'needling' cytology resource in the form of FFPE cell blocks was thus generated for optimising antibodies. The procurement of cell samples from Whipple resections and then the processing of these cell specimens into cell blocks were performed in a similar manner to the fine needle aspiration samples (167, 364). This comparable processing will help in transferring the optimised IHC conditions from 'needling' to archival samples.

An ideal diagnostic biomarker should have higher expression levels in the tissue of interest than its comparator (365, 366). In this study the tissue of interest was malignant cells in the PDAC cytology samples and the comparator was non-neoplastic pancreatic epithelial cells. To achieve good staining in PDAC cytology and low staining in non-neoplastic cells, the IHC conditions were optimised.

**IHC optimisation is the cornerstone of assay development in IHC biomarker research and optimisation is sometimes necessary to achieve better staining** (120, 122, 147). Usually the IHC conditions used for surgical specimens are used for cytology specimens (164, 167, 176). However, the IHC conditions of surgical specimens did not work well on 'needling' samples in the current project. Therefore, the IHC conditions were optimised using 'needling' samples to ensure good staining for archival cytology samples. A variety of parameters could be changed for optimisation of IHC and include, among others, antigen retrieval, antibody dilutions, incubation timing and temperature conditions (122, 146, 194). The antigen retrieval greatly improves the immunostaining in formalin fixed cytology specimens (367) and the method can be varied by time, temperature, type of enzyme, and buffer solutions. The heat induced epitope retrieval (HIER) is by far the most common method employed in clinical and research laboratories (146). In the present study, better staining for KOC and mesothelin was achieved after varying the antigen retrieval buffer for HIER and better staining for maspin was achieved after varying the antibody dilution. The

utilisation of HIER is subjective and there is no consensus on the type of HIER in IHC laboratories. The application of various buffers, timing and temperature conditions are used randomly. The level of immunostaining is directly related to the method adopted for HIER (195, 368, 369). We have also illustrated here that the change in HIER buffer solution significantly enhances the staining without compromising specificity. The optimisation for S100P was not successful despite varying both the antigen retrieval and antibody dilutions in the current study. This might probably be related to the processing of cytology samples. Liu et al (167) fixed the cells after EUS-FNA in formalin for 10 minutes compared to 12 hour formalin fixation in our cytology laboratory.

**The use of control specimens is important in IHC and they must be of similarly prepared material for quality assurance** (361). No cytology control is mentioned in the literature for diagnostic biomarkers investigated in PDAC (166, 167, 174). The PDAC 'needling' sample with positive staining for a biomarker was used as positive control and a normal 'needling' sample with negative staining for biomarker was used as negative control. As the 'needling' samples were prepared in a similar manner to the archival samples, they could serve as good control specimens for biomarkers.

The diagnostic utility of worked-up biomarkers (KOC, maspin and mesothelin) on archival cytology samples was then evaluated.

### 7.4.2 The staining of archival cytology samples and development of diagnostic IHC panel

Clinical translation of diagnostic biomarkers requires careful investigation from assay development to microscopic assessment of biomarkers and evaluation of their diagnostic accuracy (138, 147). A significant number of potential biomarkers have been investigated in surgical pathology specimens and a variety of cytology preparations (cell blocks, direct smear etc) to improve the diagnosis of PDAC. The usefulness of KOC, maspin and mesothelin as diagnostic biomarkers has been investigated singly and in a panel; however, none has entered routine clinical practice (166, 167, 217, 235). Lack of specific and sensitive biomarkers make the diagnosis of PDAC challenging in atypical or indeterminate to diagnose cases.

**Archival cytology samples were stained for the three biomarkers KOC, maspin and mesothelin.** The IHC staining in cytology samples was interpreted based on the overall staining irrespective of the benign or malignant mixture of cells. A variety of scoring systems have been adopted for biomarker research in PDAC cytology samples (141, 165, 173, 186, 255, 271). The Histoscore is a semi-quantitative IHC scoring system covering both staining intensity and the percentage of positive cells and gives a better indication of the expression level of the biomarker (321). Biomarkers with high expression levels in malignant tissue are better candidates for diagnostic purposes. Hence the cytology samples were scored using a Histoscore for each sample. KOC is highly expressed in malignant cytology samples as illustrated by the high Histoscore values compared to maspin and mesothelin. The expression of KOC is comparatively more homogenous than maspin and mesothelin in malignant cytology samples. In comparison, all three biomarkers have either no or focal expression in benign samples.

The sensitivity and specificity of biomarkers was assessed using cut-offs identified in chapter 4. A special focus was on the three cut-offs i.e. 10%, 20% and +2/+3 which were tested for reliability and reproducibility in chapter 6. The three most common cut-offs used in the literature for PDAC cytology are positive or negative staining (255, 256, 271), 5% positive cells (175) and 10% positive cells (159, 186). The 10% cut-off was therefore investigated for reproducibility between two observers.

The sensitivity and specificity of biomarkers in PDAC cytology is usually tested using only one diagnostic cut-off for categorizing patients into diagnostic groups (166, 167, 217, 254). Here, we report the use of multiple cut-offs for the same biomarker investigated in the same samples. Based on the results reported, the translational scientists or pathologists can then select a cut-off that is more appropriate to their situation. A cut-off should deliver a good combination of sensitivity and specificity. We know (from chapter 6) that 10%, 20% and +2/+3 are good cut-offs, therefore, let us discuss the diagnostic ability of biomarkers individually utilizing these cut-offs.

**KOC,** is a highly sensitive and specific biomarker across all cut-offs and the values are comparable to the TMA work (presented in chapters 4). The

sensitivity and specificity achieved in the current study is 92% and 100%. A sensitivity range of 71%-95% and a specificity range of 91%-100% have been reported for KOC in PDAC cytology literature (174, 175, 217). The specificity of **maspin** is good for all cut-offs but the sensitivity is lower compared to our TMA work (presented in chapter 5) and literature. The sensitivity and specificity achieved in the current study is 54% and 100%. The sensitivity and specificity for maspin reported in a PDAC cytology study is 100% and 90% (167). Finally, **mesothelin** is moderately sensitive and specific and the values are comparable to our TMA work except for 20% cut-off where the sensitivity is less than the TMAs (presented in chapters 4). The sensitivity and specificity achieved in the current study is 72% and 100%. A sensitivity range of 62%-69% and a specificity range of 91%-100% have been reported for mesothelin in PDAC cytology literature (164, 173, 254, 255).

Taken together, these biomarkers could help diagnose PDAC from clinical cytology samples with good diagnostic accuracy for KOC followed by mesothelin and maspin. KOC appears to be a very good candidate for categorizing patients in malignant and benign groups. This also agrees with the PDAC cytology literature for KOC and is highly recommended for future diagnostic studies. Maspin and mesothelin have comparatively low sensitivity but they are highly specific which makes them suitable candidates for investigation as a panel. Lastly, these results should be interpreted with caution due to the small sample size of this pilot project. These biomarkers were next investigated as a panel.

**A panel of KOC, maspin and mesothelin is highly sensitive and specific for PDAC diagnosis from cytology.** The high specificity of KOC, maspin and mesothelin will ensure that the combination of biomarkers will not affect the specificity and will improve the sensitivity. The inter- and intra-tumour heterogeneity of biomarker expression makes it hard for a single candidate to be 100% sensitive (191, 192). Our results also show that none of the proposed biomarkers provide perfect sensitivity. Thus, a negative test result based on a single biomarker expression can lead to false negativity. Therefore, a panel of biomarkers seems more appropriate as it improves the sensitivity and the chance of missing malignancy is significantly reduced. Highly specific biomarkers are important to avoid false positivity and misclassification of disease (366, 370).

Clinically, this would mean that positivity is essentially limited to the condition of interest (malignant disease) and negative results essentially mean a benign etiology.

Two different approaches were adopted in this study to determine a panel delivering clinically useful sensitivity and specificity. One approach categorised patients into malignant and benign disease if one or more biomarker in a panel was positive. In contrast, the second approach assigned patient into malignant and benign groups if two or more biomarkers were positive in the panel. Clearly, the second approach provides more diagnostic confidence to the cytopathologist reporting the sample based on biomarker expression. However, the first approach provides better sensitivity compared to the second approach. The panel of KOC, maspin, mesothelin utilising 10% cut-off and two or more positive biomarkers achieve good sensitivity (82%) and specificity (100%). This panel is recommended for further investigation in prospective samples.

Diagnostic biomarkers have been previously investigated in cytology samples collected during EUS-FNA and pancreatico-biliary brushing. They have been investigated as single biomarkers e.g. KOC, MUC1, S100P, XIAP, mesothelin, P53 and SMAD4 (141, 170, 174, 177, 217, 254, 258) or as panel of biomarkers (139, 159, 165, 173, 175, 176, 188). The panel of biomarkers investigated in pancreatic cytology mostly compose of two biomarkers for example panels of B7-H4 and P53 (165), S100P and XIAP (159), MUC1 and MU2 (170), KOC and S100A4 (175) and mesothelin and 14-3-3σ (164). The sensitivity and specificity achieved in the current study for a panel of KOC, maspin and mesothelin is 100% and 100% (at least one positive marker in panel) and, 82% and 100% (at least two positive biomarkers in panel). This is comparable to the panel of KOC and S100A4 with a sensitivity and specificity of 100% and 95% (175) , panel of S100P and XIAP with sensitivity of 83% (specificity not given) (159). Furthermore, the specificity of our panel is better than the panel of mesothelin and 14-3-3σ with sensitivity and specificity of 96% and 75% (164).

The panel of KOC, maspin and mesothelin has thus good translational potential. Neuroendocrine differentiation markers such as synaptophysin, CD56 and chromogranin A are clinically used for confirming the diagnosis of neuroendocrine tumours in pancreatic cytology (153, 154, 371). The staining of

PDAC cytology achieved for the biomarkers investigated in the current project (Figure 7.2 and Figure 7.3) is comparable to the neuroendocrine markers (synaptophysin and CD56) staining in pancreatic cytology (Figure1c and 1d (153)). The clinical translation of the proposed biomarkers might thus be easy for improving PDAC diagnosis from cytology samples in difficult to diagnose cases.

**Finally, the 10% cut-off was tested for observer agreement in PDAC cytology samples.** From chapter 6, we know that the 10% cut-off achieves high observer agreement. Almost perfect (k score >0.80) agreement was observed for the 10% cut-off for all three biomarkers. The k score agreement generated from our work is comparable to (k score=0.71-0.89) (372, 373) and sometimes better (k score=0.44-0.54) (374, 375) than the k scores reported in the cytology literature. The PDAC cytology literature was searched (Pubmed, 12 August 2014) but no study was found assessing observer agreement for IHC markers in pancreatic cytology. The samples size is small and hence these results should be interpreted with caution, however, this study adds to our understanding of observer agreement in the PDAC cytology. This may open avenues for future studies investigating biomarkers and cut-offs in cytology samples for observer agreements.

**The strengths of this study are**: utilising 'needling' cytology samples for IHC optimisation and as controls for cytology IHC run before staining the precious and finite archival samples; using Histoscore for scoring purposes providing a semi-quantitative expression level of biomarkers; assessing of IHC staining by an independent pathologist (SB) to remove the subjective bias that the author (AA) might introduce to this important pilot study; evaluating sensitivity and specificity of biomarkers utilising consensus based cut-offs; investigating 10% cut-off for inter-observer agreement.

**The limitations are:** the sample size is relatively small (n=21 cytology samples), however, the pancreatic cytology literature for biomarker studies can be as small as 15 (165, 376), 21 (174), 22 (258) and 25 (217); the inability to stain cytology samples for S100P which will be part of our future work; lower than expected sensitivity achieved for maspin compared to our TMA work and literature; and only two observers for inter-observer agreement analysis.

**Conclusions:** A pilot project for improving PDAC diagnosis from cytology samples was evaluated utilising a panel of KOC, maspin and mesothelin. This panel appears to achieve good sensitivity and specificity in diagnosing PDAC and excluding benign cytology. A threshold of 10% positive cells appears to achieve better diagnostic accuracy for PDAC diagnosis for all biomarkers and panels. In addition, the inter-observer agreement for 10% cut-off is very good for all three biomarkers. To sum up, our data show that a panel of KOC, maspin and mesothelin using 10% cut-off could potentially be used as an adjunct to morphologic interpretation (like neuroendocrine markers) to confirm the diagnosis of PDAC from cytology.

**What this study adds:** This study proposes the utility of a panel for biomarkers for improving the diagnosis of PDAC from cytology samples as an adjunct to cytology in difficult-to-diagnose cases. It also provides evidence of the utility of 'needling' samples for IHC biomarker optimisation.

**What next:** Applying the suggested panel to prospectively collected cytology samples and assessing the role of the panel in PDAC diagnosis from cytology samples.

**Table 7.1: Summary statistics of KOC, maspin and mesothelin expression comparing PDAC and benign cytology samples**

| Biomarkers | | Malignant cytology | Benign cytology | P value |
|---|---|---|---|---|
| **KOC** | | | | |
| ***Positivity**** | Mean | 66% | 0% | <0.001 |
| | Median | 80% | 0% | |
| ***Histoscore*** | Mean | 168 | 0 | <0.001 |
| | Median | 175 | 0 | |
| **Maspin** | | | | |
| ***Positivity*** | Mean | 33% | 1.2% | <0.003 |
| | Median | 50% | 0% | |
| ***Histoscore*** | Mean | 50 | 1.5 | <0.003 |
| | Median | 60 | 0 | |
| **Mesothelin** | | | | |
| ***Positivity*** | Mean | 36% | 0.5% | <0.001 |
| | Median | 20% | 0% | |
| ***Histoscore*** | Mean | 73 | 0.5 | <0.001 |
| | Median | 40 | 0 | |

**Note:** *Positivity (percentage of positive cells with any staining intensity in tumour and normal tissue); P value (shows the statistical significance of the difference in expression of a biomarker in tumour vs. normal tissue); Positivity range (0-100), Histoscore range (0-300).

**Table 7.2: Distribution of staining results for biomarkers in PDAC and benign cytology samples**

| Biomarkers | PDAC Cytology | Benign Cytology |
|---|---|---|
| **KOC** | 10/11 (92%) | 0/10 (0%) |
| **Maspin** | 6/11 (54%) | 0/10 (0%) |
| **Mesothelin** | 8/11 (72%) | 0/10 (0%) |
| **KOC, maspin, mesothelin** | 11/11 (100%) | 0/10 (0%) |

**Note:** Data are given as number of positive cases for biomarker divided by total number of cases. The percentage of positive cases is shown in brackets. These results are based on 10% cut-off. A positive case is defined as the presence of 10% cells with positive staining in the cytology sample.

**Table 7.3: Panels of biomarkers used for sensitivity and specificity analyses, using 10% and 20% positive cells as cut-offs for positivity**

| **10% positive cells as cut-off** | | |
|---|---|---|
| **Panels** | **Sensitivity** | **Specificity** |
| **KOC, Maspin, Mesothelin** | 100% | 100% |
| **KOC, Mesothelin** | 100% | 100% |
| **KOC, Maspin** | 90% | 100% |
| **Maspin, Mesothelin** | 90% | 100% |
| **KOC, Maspin, Mesothelin*** | 82% | 100% |
| **20% positive cells as cut-off** | | |
| **Panels** | **Sensitivity** | **Specificity** |
| **KOC, Maspin, Mesothelin** | 90% | 100% |
| **KOC, Mesothelin** | 90% | 100% |
| **KOC, Maspin** | 82% | 100% |
| **Maspin, Mesothelin** | 82% | 100% |
| **KOC, Maspin, Mesothelin*** | 72% | 100% |

**Note: *** At least 2 biomarkers required to be positive in this panel. In the rest of the panels only one biomarker was required to be positive in a panel.

**Table 7.4: Inter-observer agreement (kappa score) for 10% cut-off between two observers (A and B) for KOC, maspin, mesothelin and all 3 biomarkers**

| K Scores | | |
|---|---|---|
| **Biomarker-Inter-observer** | **10% Cut-off** | **P-Value** |
| **KOC-AB** | 1.00 | 0.000 |
| **Maspin-AB** | 0.84 | 0.000 |
| **Mesothelin-AB** | 0.83 | 0.001 |
| **All 3 Biomarkers-AB** | 0.89 | 0.000 |

**Figure 7.1: Immunostaining of KOC, maspin and mesothelin comparing the biomarker expression in needling cytology samples and corresponding resection specimens**

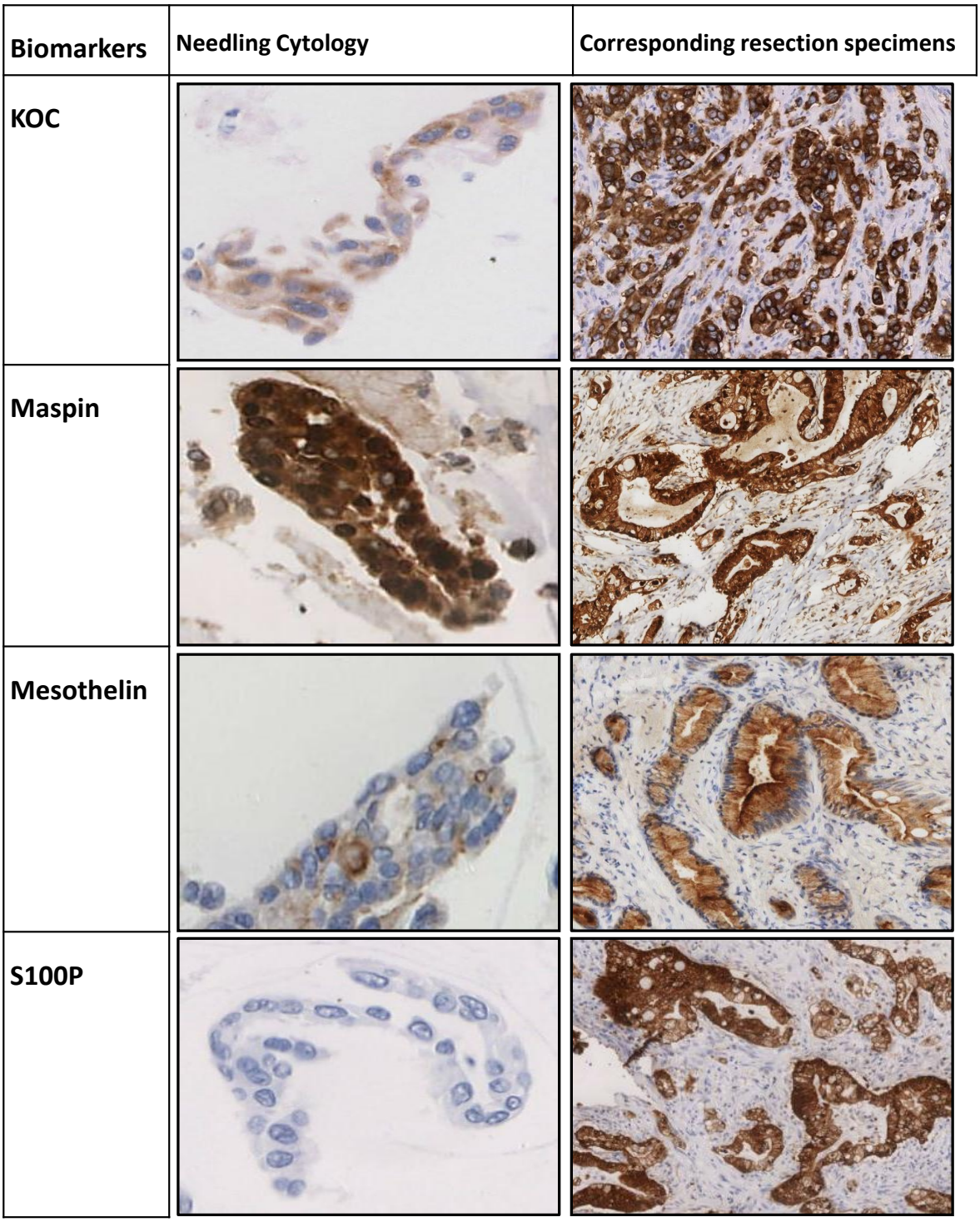



**Figure Legend:** The IHC staining of needling and corresponding resection specimens is shown for in a grid manner for KOC (cytoplasmic staining), maspin (cytoplasmic and nuclear staining), mesothelin (cytoplasmic and memebranous staining) and S100P (cytoplasmic and nuclear). This staining was achieved with 'IHC parameters optimised for TMAs'.

**Figure 7.2: Immunostaining of KOC, maspin and mesothelin comparing the different optimisation parameters using 'needling' cytology samples**

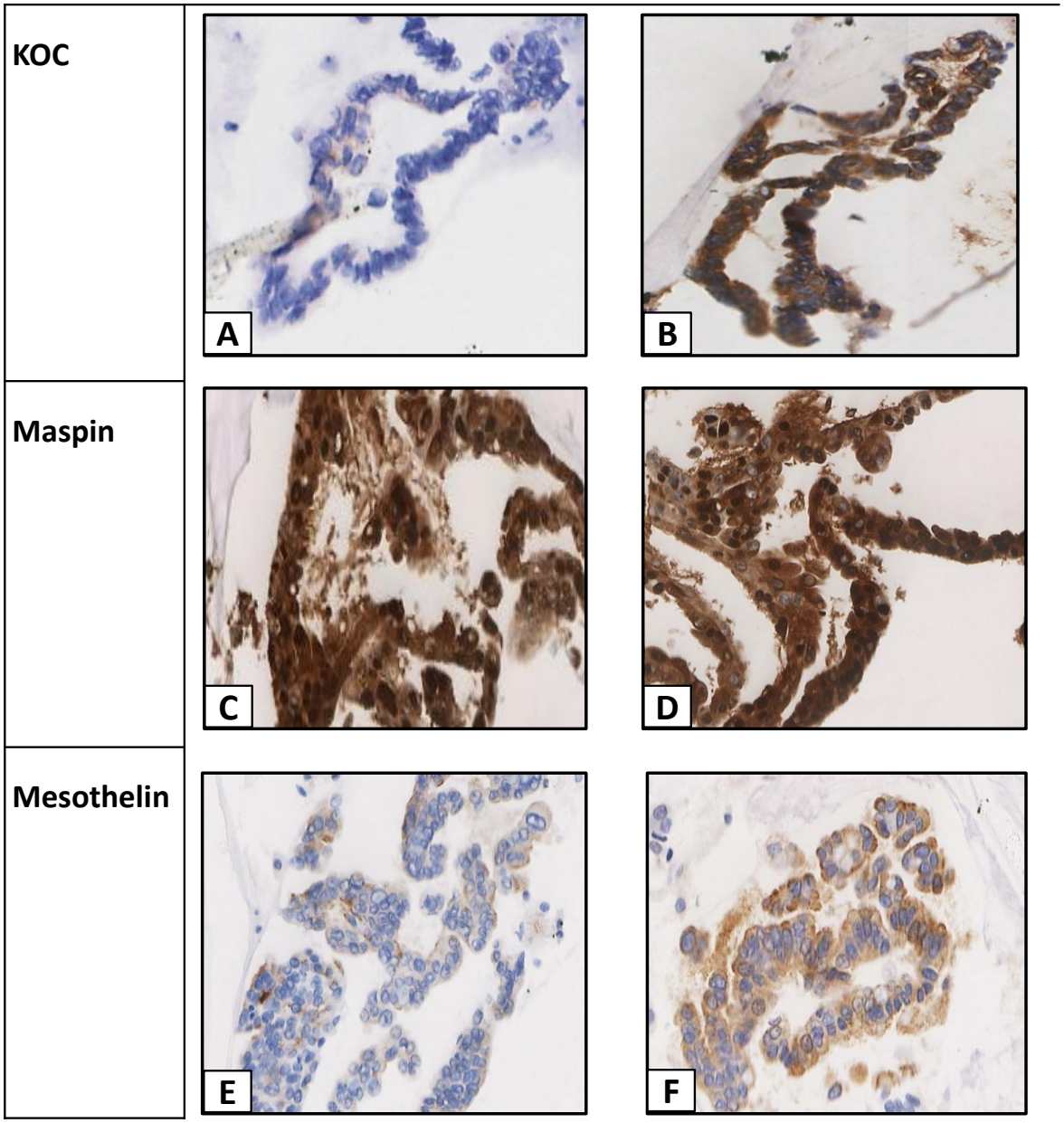



**Figure Legend:** Comparison of the staining patterns achieved with 'IHC parameters optimised for TMAs' **(A, C, E)** and 'IHC parameters optimised for cytology' **(B, D, F)** for KOC, maspin and mesothelin. For KOC comparing the staining achieved with **(A)** pH 6.0 as a heat induced antigen retrieval buffer with staining achieved using **(B)** pH 9.0. For maspin comparing the staining achieved with **(C)** 1/75 primary antibody dilution with **(D)** 1/150. For mesothelin comparing the staining achieved with **(E)** pH 6.0 as a heat induced antigen retrieval buffer with staining achieved using **(F)** pH 8.0.

**Figure 7.3: Immunostaining of KOC, Maspin and Mesothelin in benign and PDAC cytology samples.**

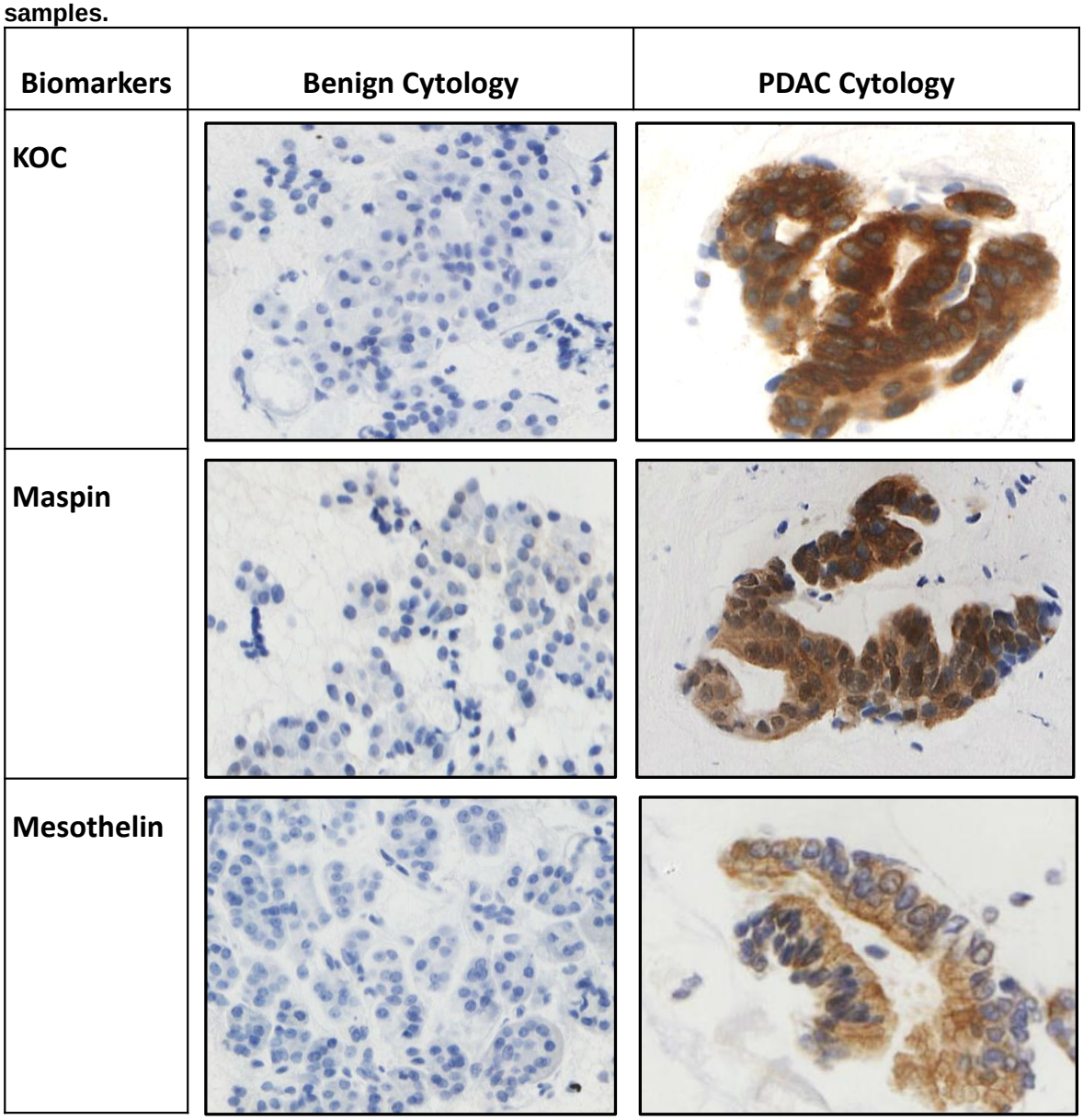



**Figure Legend:** Representative images of staining observed for KOC (cytoplasmic), maspin (cytoplasmic) and mesothelin (membranous and/or cytoplasmic) in benign compared to malignant PDAC cytology samples. This staining was achieved with 'IHC parameters optimised for cytology'.

**Figure 7.4: Sensitivity and specificity analysis based on five cut-offs for biomarkers**

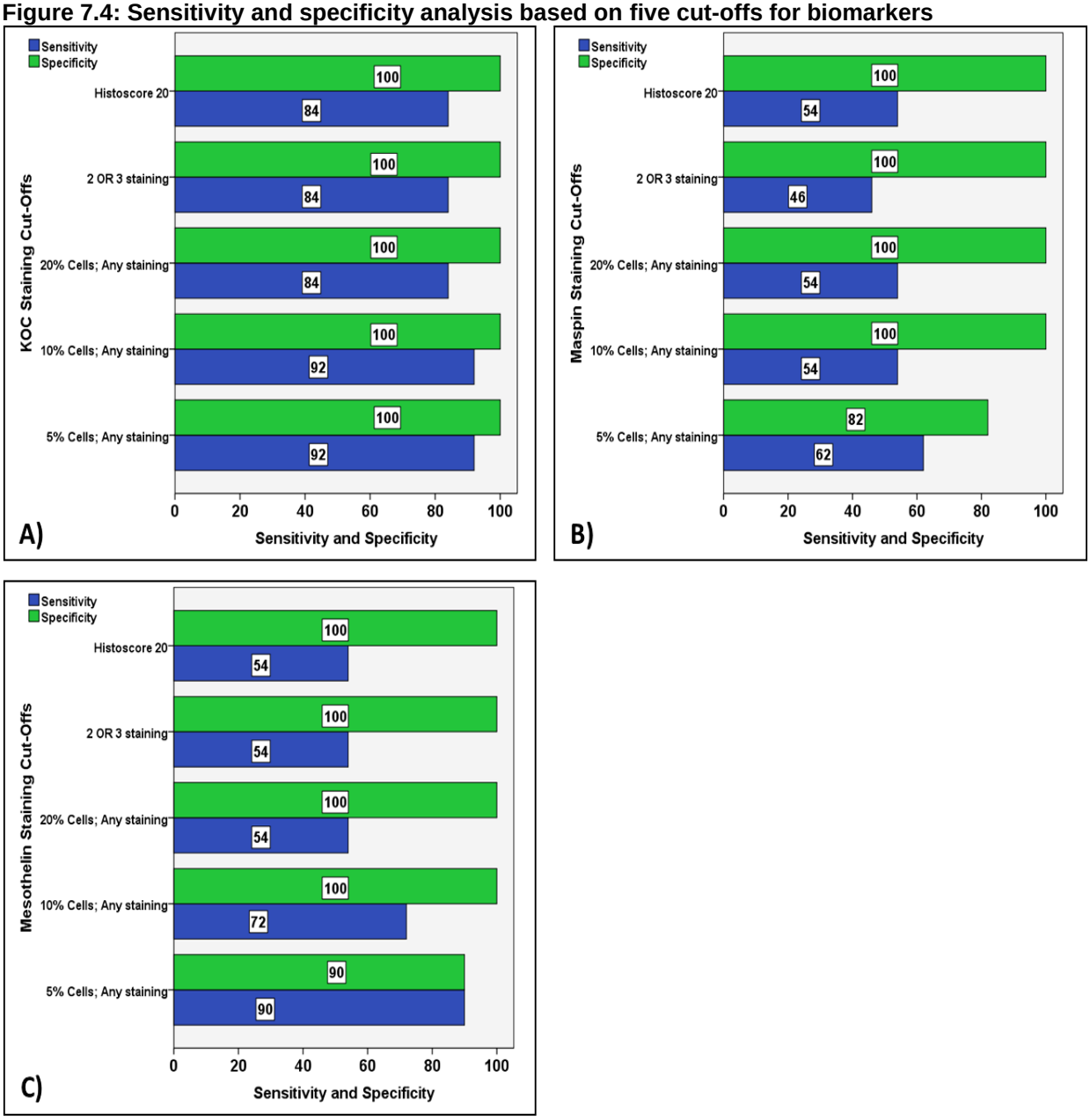



**Figure Legend:** Sensitivity and specificity analysis of biomarkers for the diagnosis of PDAC compared to benign disease, based on five cut-offs for positivity: 5% positive cells of any staining intensity; 10% positive cells of any staining intensity; 20% positive cells of any staining intensity; 2 OR 3 intensity i.e. moderate or strong staining of cells; and Histoscore 20. Analysis is presented for **A)** KOC, **B)** Maspin and **C)** Mesothelin.

**Figure 7.5: Combined Summary ROC curves for comparing panels of biomarkers.**

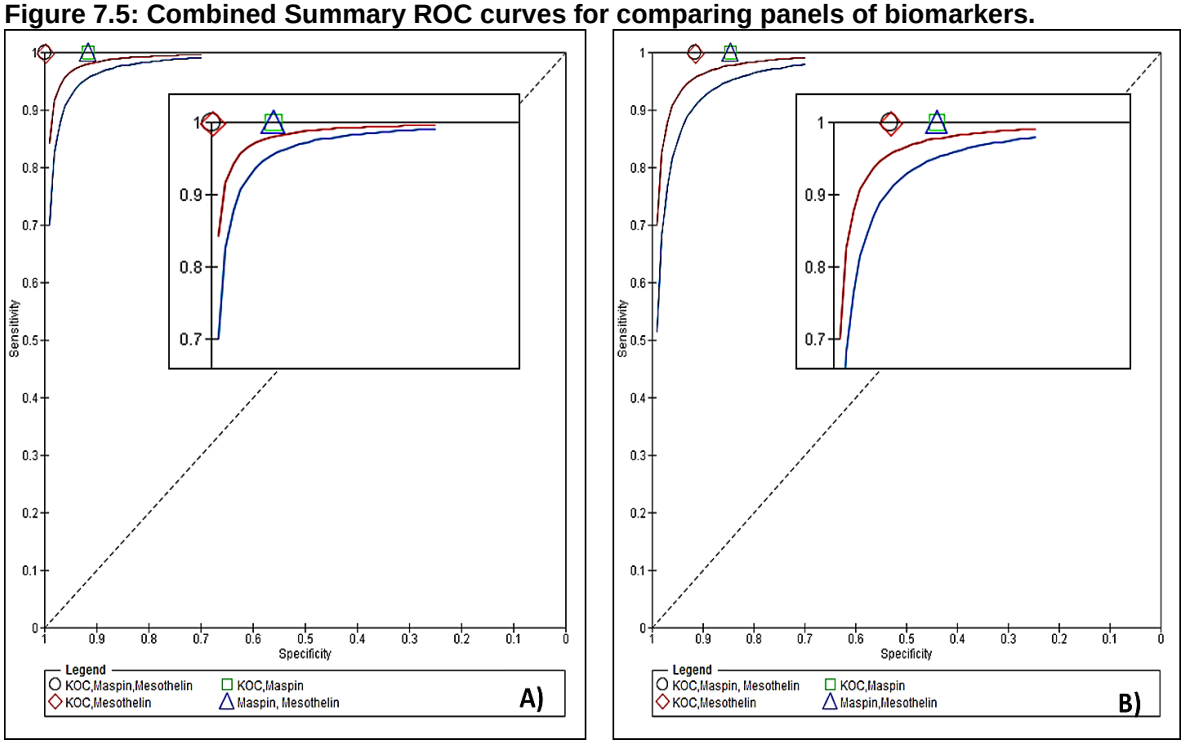



**Figure Legend:** *Combined Summary ROC curves for 10% **(A)** and 20% **(B)** cut-offs if only one biomarker was required to be positive in a panel. Four panels of biomarkers were compared. Panel 1 - KOC, Maspin, Mesothelin; Panel 2 - KOC, Mesothelin; Panel 3 - KOC, Maspin; Panel 4 - Maspin, Mesothelin.

*Summary ROC curves plot sensitivity against specificity and draw a summary line depicting combined sensitivity and specificity of a panel. Combined Summary ROC curves compare different panels to show the most “accurate” panel. The summary line at the top left corner shows the biomarker which is most accurate compared to others lying lower and further to the right. This enables the most accurate panel to be identified.

**Figure 7.6: Hematoxylin & eosin and IHC staining of a PDAC cytology case to demonstrate the utility of panel of biomarkers**

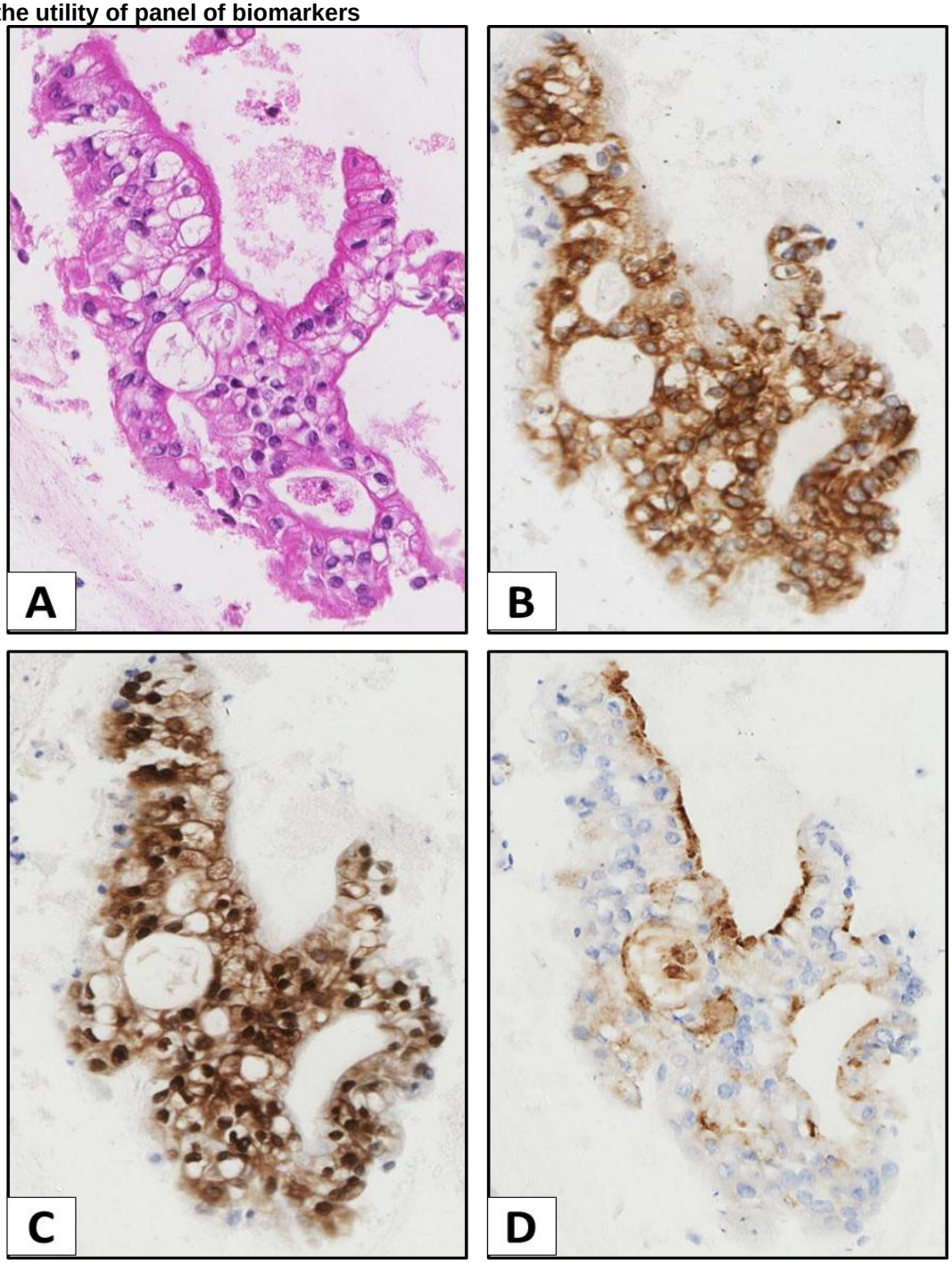



**Figure legend:** These images are showing the importance of panel of biomarkers. **A)** H&E staining of a PDAC cytology case, **B)** Cytoplasmic staining of KOC showing moderate-strong and diffuse expression **C)** Cytoplasmic and nuclear staining of maspin showing moderate-strong and diffuse expression **D)** Membranous staining of mesothelin showing moderate-strong and diffuse expression, some weak cytoplasmic staining and patchy distribution. This staining was achieved with 'IHC parameters optimised for cytology'.

# 8 Overall discussion

## 8.1 Thesis summary

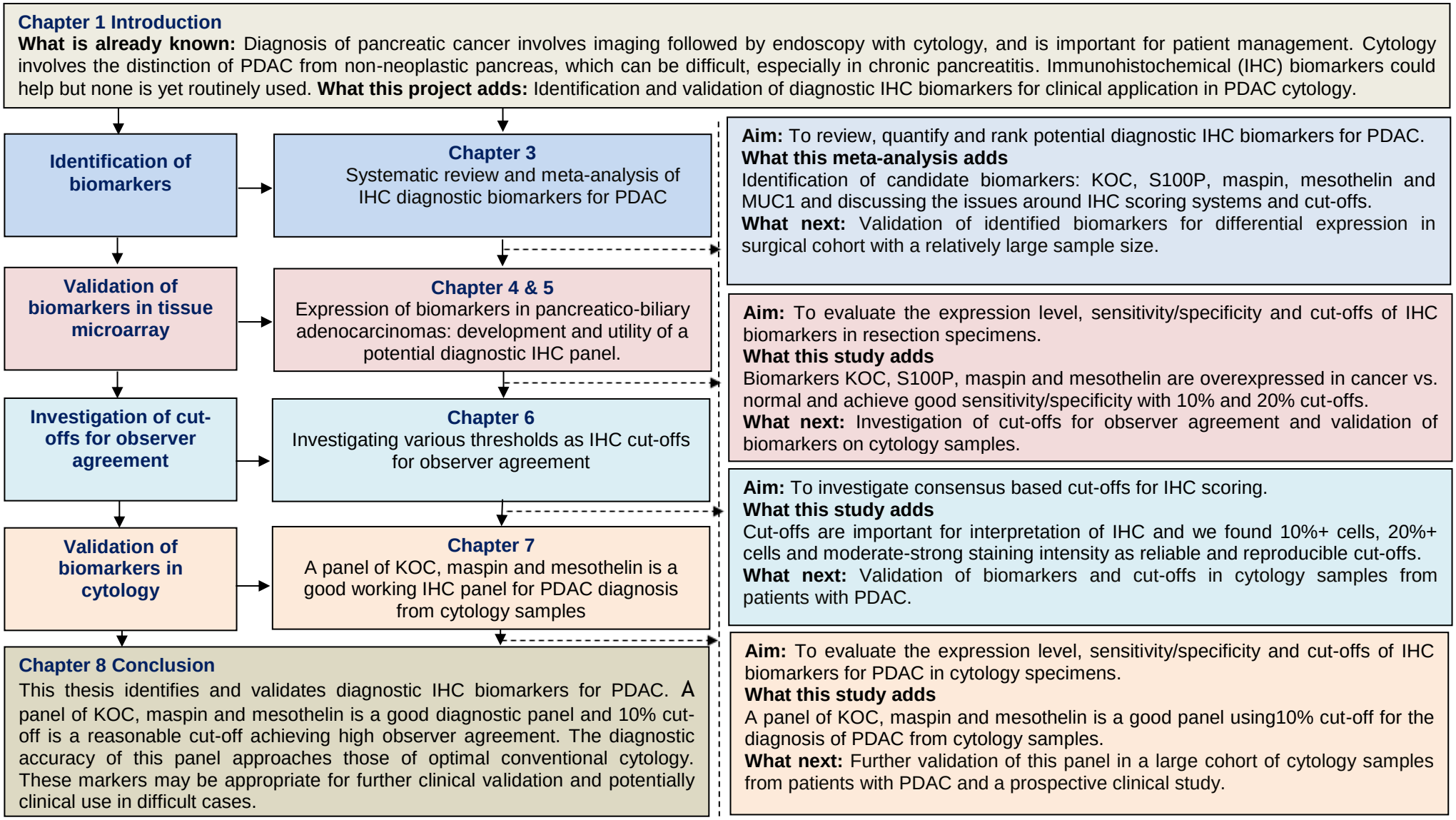
This thesis presents the role of immunohistochemistry (IHC) biomarkers in improving the diagnosis of pancreatic ductal adenocarcinoma (PDAC) from cytology samples. To achieve this goal, a systematic approach was adopted to identify and validate biomarkers for potential clinical application.

The identification phase involved reviewing the literature to identify already existing IHC biomarkers investigated in PDAC resection and cytology specimens. This was performed as discussed in chapter 3, with a focus on IHC biomarkers showing differential expression between PDAC and non-neoplastic pancreas. Rigorous search criteria were used to identify the most relevant articles for systematic review and meta-analysis. The meta-analysis helped to quantify and rank the performance of biomarkers.

High ranking candidate biomarkers from the meta-analysis were then validated in our cohort of resection specimens from patients with pancreatico-biliary adenocarcinoma (PBA) in an attempt to determine a suitable candidate biomarker or panel of biomarkers. This was presented in chapter 4 and 5, which also discuss various cut-offs for a positive diagnosis based on IHC staining. The project evolved and a need to investigate a consensus based cut-off with high inter- and intra-observer agreement emerged. In chapter 6, various cut-offs for observer agreement between practicing pathologists were investigated. It was found that cut-offs from our IHC validation work on TMAs and the literature achieve a good level of agreement between pathologists.

Finally, chapter 7 reports the utility of a panel of biomarkers (KOC, maspin and mesothelin) in archival cytology samples. This panel is a good working panel with a potential to improve the diagnosis of PDAC as an adjunct to cytology. Figure 8.1 is a summarising diagram of the project.

**Figure 8.1: Schematic diagram of thesis**



## 8.2 Implications for current diagnostic practice

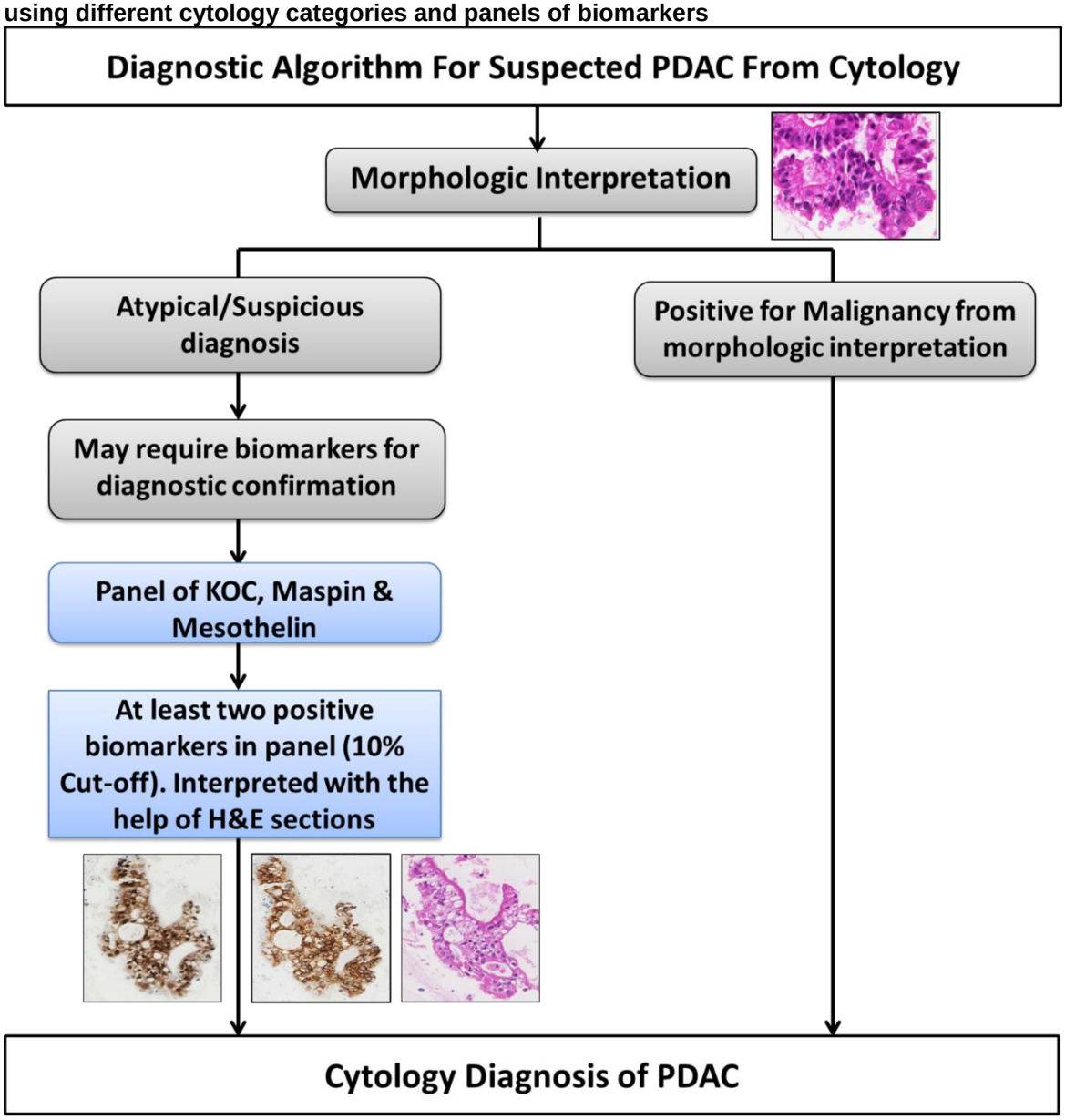
Biomarkers investigated in this project could help to improve the diagnosis of patients from cytology samples in patients suspected of PDAC. IHC is already in clinical practice in pancreatic cytology samples from patients with neuroendocrine tumours (NET) (153, 154, 371). IHC staining of CD56, chromogranin A and synaptophysin is used for confirming the diagnosis of NET from pancreatic cytology samples. Hence IHC biomarkers could potentially be translated to clinical cytology samples for PDAC.

Currently, only the morphological features of cells in the cytology samples are used for the diagnosis of PDAC. However, like NET biomarkers might be used to confirm the diagnosis in difficult cytology cases. Improvement in the diagnosis will further advance the diagnostic management of patients with PDAC and will remove the uncertainty associated with a PDAC diagnosis in indeterminate cases.

A diagnostic algorithm (Figure 8.2) can be developed for cost effectiveness of IHC biomarkers. One possible algorithm can be as follows: If routine haematoxylin and eosin staining is sufficient for confirmation of malignant or benign diagnosis, no biomarker panel will be applied. However, in difficult cases when the cytologist reports an ‘atypical’ or ‘suspicious’ diagnosis; the panel of KOC, maspin and mesothelin could be applied. The presence of at least two positive biomarkers in the panel using a 10% cut-off might be considered as a cytology adjunct for PDAC diagnosis. It is important to note here that the final diagnosis is made by a multi-disciplinary team of oncologists, surgeons, radiologists, pathologists and other supporting staff. Therefore, this proposed algorithm should be used in combination with all the clinical, radiological and pathological findings.

Some centres are now employing neoadjuvant treatment to all patients eligible for surgery and these treatment strategies require confirmation of tissue diagnosis. Hence a panel of biomarkers providing increased diagnostic confidence from cytology samples will advance the management of patients suspected of PDAC.

**Figure 8.2: Diagnostic suggested algorithm for the diagnosis of PDAC cytology samples using different cytology categories and panels of biomarkers**




## 8.3 Implications for translational biomarkers research

A systematic strategy was adopted in this thesis to address a clinical issue in PDAC diagnosis. IHC biomarkers were identified and sequentially validated in pancreatic surgical and cytology specimens. We believe that this strategy might potentially be useful in addressing clinical problems in other cancer types. This is particularly helpful for biomarkers that have been discovered and validated in clinical samples addressing a specific clinical issue. In addition, this strategy is not limited to diagnostic biomarkers but could provide a platform for prognostic and predictive biomarkers. The researcher can start with a clinical problem, identify high ranking biomarkers through a meta-analysis and then purposefully validate in tissue and cell samples as required. This type of work would require

close collaboration between pathologists, surgeons, oncologists, scientists and public health researchers.

In addition, the cut-offs validated for observer agreements in this thesis have potential implications for translational research. These cut-offs could potentially be used for evaluating the diagnostic or prognostic utility of biomarkers as the current study provides evidence that they are reliable and reproducible. Biomarkers already using these cut-offs in translational studies could possibly be easily translated to clinical practice as the observer agreement for these cut-offs is reasonably good.

## 8.4 Strengths and Limitations of this thesis

First I will cover the more specific strengths and limitations of this thesis and then I will provide a more general discussion on the strengths and limitations of this thesis.

The strengths of this thesis are as follows: the meta-analysis designed specifically for potential diagnostic IHC biomarkers in PDAC. Although good reviews (229, 230) on biomarkers for pancreatic cancer are present but such a meta-analysis is lacking for diagnostic IHC biomarkers; the validation of biomarkers identified in meta-analysis in a single experimental setting and a relatively good sample size (n=99 resection specimens cases). Our sample size is comparable to studies investigating biomarkers in PDAC resection specimens (n=93-96 cases) (171, 235) and sometimes better than other studies (n=27-53 cases) (186, 234, 237); we have evaluated the diagnostic performance of biomarkers across multiple cut-offs which is very limited in PDAC literature. Biomarkers investigated in PDAC resection and cytology specimens use only one cut-off for biomarker assessment (139, 164, 166, 174, 276); the investigation of cut-offs for observer agreement involving seven practicing pathologists. This type of study for observer agreement is very limited in PDAC literature; the IHC conditions from surgical specimens are usually used for staining the cytology samples (159, 165, 176, 186). We optimised antibodies on ‘needling’ cytology samples allowing the comparison of different IHC parameters and selecting good IHC conditions for final staining. The development of such a resource for IHC optimization is lacking in PDAC literature.

The limitations of the thesis are as follows: The use of the same tissue microarray (TMA) cohort for KOC, mesothelin, S100P and MUC1 but the use of another TMA set for maspin; the sample size (n=21 cases) of our cytology project was small but was comparable to the published literature reporting IHC biomarker in cytology samples (n=15-25 cases) (139, 165, 174, 217, 258, 376); S100P though a good marker in resection specimens could not be stained successfully in cytology samples; and lastly the sample size of cholangiocarcinoma (CCC) was low (n=14).

New research on diagnostic IHC biomarkers for PDAC has been published since our meta-analysis. Of special interest is the paper from Liu et al (167) where the authors investigated the utility of 26 candidate biomarkers in PDAC. IHC was performed for these 26 biomarkers using TMAs from 60 patients with resection specimens. From this initial investigation, four biomarkers maspin, S100P, IMP-3 (KOC), and pVHL were selected based on their expression levels in tissue sections for further investigation in cytology sections. Three biomarkers maspin, S100P, IMP-3 (KOC) were positive biomarkers and pVHL was a negative biomarker for PDAC. Finally, a panel of maspin, S100P, IMP-3 (KOC), and pVHL was suggested for potential clinical use (167). By comparison, we identified 49 potential IHC biomarkers in our systematic review including 24 of the 26 IHC biomarkers investigated in Liu et al's study. Afterwards, we performed a meta-analysis and through quantification identified IMP-3 (KOC), maspin, meosthelin and S100P as high ranking candidates which are similar to the final proposed biomarkers by Liu et al (167). In addition, pVHL as negative biomarker was identified in our systematic review but we did not investigate this marker in our project. To sum up, a significant amount of time, resource and money might be required to conduct a laboratory study when investigating a large number of biomarkers. In comparison, a systematic review and meta-analysis can review the available evidence, quantify the performance of existing biomarkers and help in selecting better candidates with significantly less time and resource. It may be of value first to perform a focused meta-analysis in biomarker research to address a specific clinical problem.

Other candidate IHC diagnostic biomarkers reported since our meta-analysis as potentially differentiating PDAC from non-neoplastic pancreas include Ki-67,

Smac (377), PAM4 (378), ApoE (137), CRABP-II (379), MUC4 and MUC16 (380). Three biomarkers CRABP-II, MUC4 and MUC16 in particular have been investigated in fine needle aspiration cytology and are worth validating in other cohorts.

First let us discuss the markers investigated in resection specimens and compare their specificities and sensitivities with the markers investigated in resection specimens in the current project. The specificity/sensitivity of differentiating PBA from normal pancreas for KOC is 100%/84%, maspin 99%/96%, S100P 100%/83% and mesothelin is 92%/88%. In comparison, the specificity and sensitivity of Ki-67 in differentiating PDAC from normal pancreas was 84% and 94%. Similarly, the specificity and sensitivity of Smac (a pro-apoptotic mitochondrial protein) was 92% and 100% (377). The sample size of this study was relatively small (n=47) but Smac appears to be a good candidate for further investigation (377). PAM4 was investigated in PDAC and chronic pancreatitis cases. The sensitivity of PAM4 was 79%, while the specificity was 81% if staining in the pancreatic intraepithelial neoplasias associated with CP were included. However, the specificity was 100% if staining in the PanINs is excluded from the specificity analysis. PAM4 could thus be a candidate for further investigation in cytology samples (378). The specificity and sensitivity of ApoE in differentiating PDAC from normal pancreas was 85% and 78% (137). These markers in particular Smac and PAM4 may be worth validating in cytology samples. Their diagnostic accuracies are good and the specificity and sensitivity of Smac is comparable to markers investigated in the current PhD project. However, KOC, maspin and mesothelin have been investigated in various studies from independent research groups. The new markers might thus require independent validation studies on FFPE tissue sections from patients with PDAC to further explore their diagnostic potential in PDAC.

Three markers CRABP-II, MUC4 and MUC16 appear promising (379, 380). CRABP-II was investigated in both resection specimens and cytology specimens, while MUC4 and MUC16 were investigated in cytology specimens. The sensitivity and specificity of CRABP-II in differentiating PDAC from normal pancreas and chronic pancreatitis in resection specimens was 100% and 100%. However, the sensitivity dropped to 61% in cytology samples but the specificity was still 100% (379).

MUC4 and MUC16 are highly specific biomarkers (specificity 100% for both) with 74% and 63% sensitivities respectively (380). Due to their high specificity they might potentially be used as a panel and appear to be good potential diagnostic candidates. However, their sensitivities are less than KOC (sensitivity 92%) and comparable to mesothelin (sensitivity 72%) and maspin (sensitivity 54%) as shown in our archival cytology project. Independent validation is required to further explore the clinical utility of CRABP-II, MUC4 and MUC16.

## 8.5 Future research

Referring back to the phases of biomarker development and factors delaying the clinical translation of biomarkers discussed in chapter 1 (specifically Section 1.9.2 and 1.10); we have now identified potential biomarkers and validated them in our surgical and cytology cohort. However, independent validation is very important for translation of biomarkers into clinically useful tools (142, 181, 199). Future research should therefore focus on validation studies enabling the clinical translation of biomarkers for pancreatic cytology samples.

The panel of KOC, maspin and mesothelin would require validation in retrospective and prospective studies. The analytic phase of these biomarkers is well analysed and reproducible in our laboratory. However, the pre-analytic handling and processing of tissues should be standardized to allow for the reproducibility of staining and diagnostic impact in other laboratories. A larger retrospective study on archival cytology samples would further determine the clinical utility of this panel. The retrospective study allows comparison of multiple biomarkers and helps in the development of algorithms for clinical use (381, 382). For example, a trend can be established whereby two or more biomarkers can be found positive in one clinical scenario for example in a ‘suspicious’ diagnostic category. Furthermore, retrospective longitudinal studies are relatively quicker than prospective studies. This holds particularly true for pancreatic cancer where very few cases are seen in most centers compared to cancers from other organs. A retrospective study could thus provide a solid and quick platform for the usefulness of candidate biomarkers investigated in PDAC. Sample selection in retrospective studies is very important in evaluating the diagnostic utility of biomarkers. The application of the proposed panel for PDAC cytology samples should thus be from a range of atypical, suspicious and positive

for malignancy diagnosis. This will provide more information on the diagnostic potential of biomarkers in difficult to diagnose cases.

After evaluating the clinical utility of biomarkers in retrospective samples the next validation step is the application of panels to prospective cytology samples. The purpose of the prospective study is to determine if the biomarkers could deliver the proposed diagnostic utility in routine clinical practice (382). These prospective clinical studies are therefore important to show the true clinical utility of biomarkers for improving the diagnosis of PDAC from cytology samples. Finally, the utility of our proposed panel in independent pathology laboratories will further enhance their clinical utility for routine diagnostic use.

## 8.6 Final conclusion (an amalgamation of all results chapters)

IHC biomarkers have been extensively investigated to improve the diagnosis, prognosis and prediction of response to treatment for a variety of cancers including pancreatic cancer. The aim of this project was to improve the diagnosis of PDAC from cytology samples using IHC biomarkers as an adjunct. To achieve this goal, the project involved the identification of IHC biomarkers from literature through meta-analysis and validation in PDAC resection and cytology samples. This project also involved the investigation of scoring systems and consensus based cut-offs.

The fact that many potentially diagnostic IHC biomarkers for PDAC already exist in the literature was demonstrated by our systematic review. Current evidence shown by our systematic review is that biomarkers identified are validated singly with less utilisation of panel approach. Moreover, panels of biomarkers are randomly selected for validation studies with a no ‘evidence based’ approach. In this project biomarker identification was through meta-analysis and a panel of biomarkers was proposed. Interestingly, the sensitivity and specificity values of biomarkers in our validation cohort agree with the values reported in our meta-analysis. This demonstrates the utility of meta-analysis in biomarker research to prioritise candidates for potential validation from a pool of biomarkers

The interpretation of IHC staining for diagnostic purposes requires a cut-off for assigning patients into positive or negative categories. From the TMA work five diagnostic cut-offs were identified. But for clinical translation the cut-offs should have high observer agreement. Consensus based cut-offs were therefore investigated among practicing pathologists. It was found that all three cut-offs (10%, 20%, +2/+3) achieve reasonable strength of observer agreements. The work presented here is evidence that these cut-offs are reliable and reproducible and might be used in other areas of cancer pathology and could potentially be applied to cytology samples in PDAC.

The technique of developing a 'needling' cytology resource presented here shows the value of generating such a resource for IHC optimisation purposes. The rationale behind this optimisation was to achieve better staining in archival cytology samples. After addressing the issues hindering the clinical translation of biomarkers a pilot study on the precious but scarce archival cytology samples was performed to elucidate the clinical potential of biomarkers. A panel of KOC, maspin and mesothelin was applied to the archival cytology samples. Individual biomarkers have relatively low sensitivity but as a panel they achieve good sensitivity and specificity. In addition, the observer agreement for 10% cut-off was very good for all three biomarkers in cytology samples.

In brief, the approach of identification of biomarkers presented here demonstrates the value of utilising the published literature for addressing the clinical problems. The successful panel of biomarkers subsequently developed through validation in resection and cytology samples addresses an important clinical issue in the diagnosis of PDAC from cytology. The diagnostic accuracy of this panel approaches the optimal conventional cytology. These markers may be appropriate for further clinical validation and potentially routine use in difficult cases.

# Appendices

# Publications, Presentations and Awards

## Publications in peer reviewed journals

1. Ali A, Brown V, Denley S, Jamieson N, Morton J, Nixon C, Graham J, Sansom O, Carter CR, McKay C et al: Expression of KOC, S100P, mesothelin and MUC1 in pancreatico-biliary adenocarcinomas: development and utility of a potential diagnostic immunohistochemistry panel. BMC Clinical Pathology 2014, 14(1):35.

2. Jamieson NB, Morran DC, Morton JP, Ali A, Dickson EJ, Carter CR, Sansom OJ, Evans TR, McKay CJ, Oien KA: MicroRNA molecular profiles associated with diagnosis, clinicopathologic criteria, and overall survival in patients with resectable pancreatic ductal adenocarcinoma. Clin Cancer Res 2012, 18(2):534-545.

3. Ali A, Evans P: Multi-resource peer assisted learning in postgraduate setting: a pilot study. Journal of the College of Physicians and Surgeons--Pakistan: JCPSP 2013, 23(4):251-256.

## Presentations at national and international conferences:

1. Ali A, Brown V, Jamieson NB, Denley SM, Carter CR, McKay CJ, Oien KA: Investigation of KOC expression as a potential diagnostic biomarker for pancreatic ductal adenocarcinoma. Abstracts: Pathology Informatics 2011 Meeting. J Pathol Inform 2011;2:43.

2. Ali A, Ul-Haq Z, Mohamed M, MacKay DF, Duthie F, Oien K: Abstract 1142: Systematic review and meta-analysis of immunohistochemical diagnostic markers for pancreatic ductal adenocarcinoma. Cancer Research 2013, 73(8 Supplement):1142.

   Ali A, Ul-Haq Z, Mohamed M, MacKay DF, Duthie F, Oien K: Abstract 1142: Systematic review and meta-analysis of immunohistochemical diagnostic markers for pancreatic ductal adenocarcinoma. Edinburgh Pathology, June 2013, available from:

http://path.org.uk/edinburgh2013/files/mainprog.pdf

## Awards

1. Travel Bursary to attend Pathology Informatics Conference 2011 at the University of Pittsburgh from University of Pittsburgh, USA to present my work titled "Investigation of KOC expression as a potential diagnostic biomarker for pancreatic ductal adenocarcinoma."

2. Travel Bursary to attend American Association for Cancer Research Annual Conference 2013 at Washington D.C, USA from University of Glasgow, UK and Pathosoc to present my work titled "Systematic review and meta-analysis of immunohistochemical diagnostic markers for pancreatic ductal adenocarcinoma".

3. Travel Bursary to attend Cancer Genomics Conference 2013 at University of Cambridge, UK from British Association for Cancer Research to present my work titled "Systematic review and meta-analysis of immunohistochemical diagnostic markers for pancreatic ductal adenocarcinoma".

4. Travel Bursary to attend Molecular Approach to Pathology course at Amsterdam, The Netherlands arranged by the European Association for Cancer Research.

**List of References**


1. Encyclopaedia Britannica: *Encyclopaedia Britannica Online Academic Edition* [cited 2014July25]. Pancreas. Available from: http://www.britannica.com/EBchecked/topic/440971/pancreas.
2. Campbell F, Verbeke C. Embryology, Anatomy, and Histology. Pathology of the Pancreas: Springer London; 2013. p. 3-20.
3. Rovasio R. Development and Structure of the Pancreas. Pancreatic Cancer: Springer New York; 2010. p. 27-38.
4. Adsay NV, Basturk O, Cheng JD, Andea AA. Ductal neoplasia of the pancreas: nosologic, clinicopathologic, and biologic aspects. Semin Radiat Oncol. 2005;15(4):254-64.
5. Cancer Research UK. London: [cited 2014July23]. Pancreatic cancer: UK incidence statistics. Available from: http://info.cancerresearchuk.org/cancerstats/types/pancreas/incidence.
6. Cancer Research UK. London: News and Resources [cited 2014July23]. Pancreatic cancer statistics-Key facts. Available from: http://info.cancerresearchuk.org/cancerstats/keyfacts/pancreatic-cancer/.
7. Hidalgo M. Pancreatic Cancer. New England Journal of Medicine. 2010;362(17):1605-17.
8. Olson SH, Kurtz RC. Epidemiology of pancreatic cancer and the role of family history. J Surg Oncol. 2013;107(1):1-7.
9. Zavoral M, Minarikova P, Zavada F, Salek C, Minarik M. Molecular biology of pancreatic cancer. World J Gastroenterol. 2011;17(24):2897-908.
10. Jacobs EJ, Chanock SJ, Fuchs CS, Lacroix A, McWilliams RR, Steplowski E, et al. Family history of cancer and risk of pancreatic cancer: a pooled analysis from the Pancreatic Cancer Cohort Consortium (PanScan). Int J Cancer. 2010;127(6):1421-8.
11. Permuth-Wey J, Egan KM. Family history is a significant risk factor for pancreatic cancer: results from a systematic review and meta-analysis. Fam Cancer. 2009;8(2):109-17.
12. Koorstra JB, Feldmann G, Habbe N, Maitra A. Morphogenesis of pancreatic cancer: role of pancreatic intraepithelial neoplasia (PanINs). Langenbecks Arch Surg. 2008;393(4):561-70.
13. Maitra A, Fukushima N, Takaori K, Hruban RH. Precursors to invasive pancreatic cancer. Adv Anat Pathol. 2005;12(2):81-91.
14. Hruban RH, Adsay NV, Albores-Saavedra J, Compton C, Garrett ES, Goodman SN, et al. Pancreatic intraepithelial neoplasia: a new nomenclature and classification system for pancreatic duct lesions. Am J Surg Pathol. 2001;25(5):579-86.
15. Hruban RH, Takaori K, Klimstra DS, Adsay NV, Albores-Saavedra J, Biankin AV, et al. An illustrated consensus on the classification of pancreatic intraepithelial neoplasia and intraductal papillary mucinous neoplasms. Am J Surg Pathol. 2004;28(8):977-87.
16. Hruban RH, Fukushima N. Pancreatic adenocarcinoma: update on the surgical pathology of carcinomas of ductal origin and PanINs. Mod Pathol. 2007;20 Suppl 1:S61-70.
17. Singh M, Maitra A. Precursor lesions of pancreatic cancer: molecular pathology and clinical implications. Pancreatology. 2007;7(1):9-19.
18. Basturk O, Coban I, Adsay NV. Pathologic Classification and Biological Behavior of Pancreatic Neoplasia. Pancreatic Cancer: Springer New York; 2010. p. 39-70.
19. Campbell F, Verbeke C. Ductal Adenocarcinoma. Pathology of the Pancreas: Springer London; 2013. p. 111-51.
20. Adsay NV, Basturk O, Klimstra DS, Kloppel G. Pancreatic pseudotumors: non-neoplastic solid lesions of the pancreas that clinically mimic pancreas cancer. Semin Diagn Pathol. 2004;21(4):260-7.
21. Barreto SG, Shukla PJ, Shrikhande SV. Tumors of the pancreatic body and tail. World Journal of Oncology. 2010;1(2):52-65.
22. Bond-Smith G, Banga N, Hammond TM, Imber CJ. Pancreatic adenocarcinoma. Bmj. 2012;344:e2476.
23. Wolfgang CL, Herman JM, Laheru DA, Klein AP, Erdek MA, Fishman EK, et al. Recent progress in pancreatic cancer. CA Cancer J Clin. 2013;63(5):318-48.
24. Tummala P, Junaidi O, Agarwal B. Imaging of pancreatic cancer: An overview. J Gastrointest Oncol. 2011;2(3):168-74.
25. Callery MP, Chang KJ, Fishman EK, Talamonti MS, William Traverso L, Linehan DC. Pretreatment assessment of resectable and borderline resectable pancreatic cancer: expert consensus statement. Ann Surg Oncol. 2009;16(7):1727-33.
26. Al-Haddad M, Eloubeidi MA. Interventional EUS for the diagnosis and treatment of locally advanced pancreatic cancer. Jop. 2010;11(1):1-7.
27. Burnett AS, Calvert TJ, Chokshi RJ. Sensitivity of endoscopic retrograde cholangiopancreatography standard cytology: 10-y review of the literature. J Surg Res. 2013;184(1):304-11.

28. Chang KJ. State of the art lecture: endoscopic ultrasound (EUS) and FNA in pancreatico-biliary tumors. Endoscopy. 2006;38 Suppl 1:S56-60.
29. Fusaroli P, Kypraios D, Caletti G, Eloubeidi MA. Pancreatico-biliary endoscopic ultrasound: a systematic review of the levels of evidence, performance and outcomes. World J Gastroenterol. 2012;18(32):4243-56.
30. Edge S, Byrd, D.R., Compton, C.C., Fritz, A.G., Greene, F.L., Trotti, A. . American Joint Committee on Cancer (AJCC) Cancer Staging Manual. 7th ed. New York: Springer; 2009.
31. Assifi MM, Lu X, Eibl G, Reber HA, Li G, Hines OJ. Neoadjuvant therapy in pancreatic adenocarcinoma: a meta-analysis of phase II trials. Surgery. 2011;150(3):466-73.
32. Evans DB, Varadhachary GR, Crane CH, Sun CC, Lee JE, Pisters PW, et al. Preoperative gemcitabine-based chemoradiation for patients with resectable adenocarcinoma of the pancreatic head. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2008;26(21):3496-502.
33. Tempero MA, Arnoletti JP, Behrman SW, Ben-Josef E, Benson AB, 3rd, Casper ES, et al. Pancreatic Adenocarcinoma, version 2.2012: featured updates to the NCCN Guidelines. J Natl Compr Canc Netw. 2012;10(6):703-13.
34. Iqbal N, Lovegrove RE, Tilney HS, Abraham AT, Bhattacharya S, Tekkis PP, et al. A comparison of pancreaticoduodenectomy with pylorus preserving pancreaticoduodenectomy: a meta-analysis of 2822 patients. Eur J Surg Oncol. 2008;34(11):1237-45.
35. Tran KT, Smeenk HG, van Eijck CH, Kazemier G, Hop WC, Greve JW, et al. Pylorus preserving pancreaticoduodenectomy versus standard Whipple procedure: a prospective, randomized, multicenter analysis of 170 patients with pancreatic and periampullary tumors. Ann Surg. 2004;240(5):738-45.
36. Ferrone CR, Pieretti-Vanmarcke R, Bloom JP, Zheng H, Szymonifka J, Wargo JA, et al. Pancreatic ductal adenocarcinoma: long-term survival does not equal cure. Surgery. 2012;152(3 Suppl 1):S43-9.
37. Richter A, Niedergethmann M, Sturm JW, Lorenz D, Post S, Trede M. Long-term results of partial pancreaticoduodenectomy for ductal adenocarcinoma of the pancreatic head: 25-year experience. World J Surg. 2003;27(3):324-9.
38. Diener MK, Knaebel HP, Heukaufer C, Antes G, Buchler MW, Seiler CM. A systematic review and meta-analysis of pylorus-preserving versus classical pancreaticoduodenectomy for surgical treatment of periampullary and pancreatic carcinoma. Ann Surg. 2007;245(2):187-200.
39. Sohn TA, Yeo CJ, Cameron JL, Koniaris L, Kaushal S, Abrams RA, et al. Resected adenocarcinoma of the pancreas-616 patients: results, outcomes, and prognostic indicators. J Gastrointest Surg. 2000;4(6):567-79.
40. Winter JM, Cameron JL, Campbell KA, Arnold MA, Chang DC, Coleman J, et al. 1423 pancreaticoduodenectomies for pancreatic cancer: A single-institution experience. J Gastrointest Surg. 2006;10(9):1199-210; discussion 210-1.
41. Yeo CJ, Abrams RA, Grochow LB, Sohn TA, Ord SE, Hruban RH, et al. Pancreaticoduodenectomy for pancreatic adenocarcinoma: postoperative adjuvant chemoradiation improves survival. A prospective, single-institution experience. Ann Surg. 1997;225(5):621-33; discussion 33-6.
42. Neoptolemos JP, Stocken DD, Friess H, Bassi C, Dunn JA, Hickey H, et al. A randomized trial of chemoradiotherapy and chemotherapy after resection of pancreatic cancer. N Engl J Med. 2004;350(12):1200-10.
43. Neoptolemos JP, Stocken DD, Bassi C, Ghaneh P, Cunningham D, Goldstein D, et al. Adjuvant chemotherapy with fluorouracil plus folinic acid vs gemcitabine following pancreatic cancer resection: a randomized controlled trial. Jama. 2010;304(10):1073-81.
44. Conroy T, Desseigne F, Ychou M, Bouche O, Guimbaud R, Becouarn Y, et al. FOLFIRINOX versus gemcitabine for metastatic pancreatic cancer. N Engl J Med. 2011;364(19):1817-25.
45. Von Hoff DD, Ervin T, Arena FP, Chiorean EG, Infante J, Moore M, et al. Increased Survival in Pancreatic Cancer with nab-Paclitaxel plus Gemcitabine. New England Journal of Medicine. 2013;369(18):1691-703.
46. Gillen S, Schuster T, Meyer Zum Buschenfelde C, Friess H, Kleeff J. Preoperative/neoadjuvant therapy in pancreatic cancer: a systematic review and meta-analysis of response and resection percentages. PLoS Med. 2010;7(4):e1000267.
47. Heinemann V, Haas M, Boeck S. Neoadjuvant treatment of borderline resectable and non-resectable pancreatic cancer. Annals of Oncology. 2013.
48. Katz MH, Pisters PW, Evans DB, Sun CC, Lee JE, Fleming JB, et al. Borderline resectable pancreatic cancer: the importance of this emerging stage of disease. J Am Coll Surg. 2008;206(5):833-46; discussion 46-8.
49. Maire F, Hammel P, Ponsot P, Aubert A, O'Toole D, Hentic O, et al. Long-term outcome of biliary and duodenal stents in palliative treatment of patients with unresectable adenocarcinoma of the head of pancreas. Am J Gastroenterol. 2006;101(4):735-42.

50. Van Heek NT, De Castro SM, van Eijck CH, van Geenen RC, Hesselink EJ, Breslau PJ, et al. The need for a prophylactic gastrojejunostomy for unresectable periampullary cancer: a prospective randomized multicenter trial with special focus on assessment of quality of life. Ann Surg. 2003;238(6):894-902; discussion -5.
51. Yang YM, Tian XD, Zhuang Y, Wang WM, Wan YL, Huang YT. Risk factors of pancreatic leakage after pancreaticoduodenectomy. World J Gastroenterol. 2005;11(16):2456-61.
52. Akobeng AK. Understanding diagnostic tests 3: Receiver operating characteristic curves. Acta Paediatr. 2007;96(5):644-7.
53. Altman DG, Bland JM. Diagnostic tests. 1: Sensitivity and specificity. Bmj. 1994;308(6943):1552.
54. Lalkhen AG, McCluskey A. Clinical tests: sensitivity and specificity. Continuing Education in Anaesthesia, Critical Care & Pain. 2008;8(6):221-3.
55. Brugge W, Dewitt J, Klapman JB, Ashfaq R, Shidham V, Chhieng D, et al. Techniques for cytologic sampling of pancreatic and bile duct lesions. Diagn Cytopathol. 2014;42(4):333-7.
56. Eloubeidi MA, Chen VK, Eltoum IA, Jhala D, Chhieng DC, Jhala N, et al. Endoscopic ultrasound-guided fine needle aspiration biopsy of patients with suspected pancreatic cancer: diagnostic accuracy and acute and 30-day complications. Am J Gastroenterol. 2003;98(12):2663-8.
57. Gress F, Gottlieb K, Sherman S, Lehman G. Endoscopic ultrasonography-guided fine-needle aspiration biopsy of suspected pancreatic cancer. Ann Intern Med. 2001;134(6):459-64.
58. Adler D, Max Schmidt C, Al-Haddad M, Barthel JS, Ljung BM, Merchant NB, et al. Clinical evaluation, imaging studies, indications for cytologic study, and preprocedural requirements for duct brushing studies and pancreatic FNA: the Papanicolaou Society of Cytopathology recommendations for pancreatic and biliary cytology. Diagn Cytopathol. 2014;42(4):325-32.
59. Iqbal S, Friedel D, Gupta M, Ogden L, Stavropoulos SN. Endoscopic-Ultrasound-Guided Fine-Needle Aspiration and the Role of the Cytopathologist in Solid Pancreatic Lesion Diagnosis. Pathology Research International. 2012;2012:17.
60. Pang JC, Minter RM, Kwon RS, Simeone DM, Roh MH. The role of cytology in the preoperative assessment and management of patients with pancreaticobiliary tract neoplasms. J Gastrointest Surg. 2013;17(3):501-10.
61. Kudo T, Kawakami H, Kuwatani M, Eto K, Kawahata S, Abe Y, et al. Influence of the safety and diagnostic accuracy of preoperative endoscopic ultrasound-guided fine-needle aspiration for resectable pancreatic cancer on clinical performance. World J Gastroenterol. 2014;20(13):3620-7.
62. Chhieng D, Stelow E. Introduction to Pancreatic Cytopathology. Pancreatic Cytopathology. Essentials in Cytopathology. 3: Springer US; 2007. p. 1-21.
63. Ferrari Junior AP, Lichtenstein DR, Slivka A, Chang C, Carr-Locke DL. Brush cytology during ERCP for the diagnosis of biliary and pancreatic malignancies. Gastrointest Endosc. 1994;40(2 Pt 1):140-5.
64. DeWitt J, Devereaux B, Chriswell M, McGreevy K, Howard T, Imperiale TF, et al. Comparison of endoscopic ultrasonography and multidetector computed tomography for detecting and staging pancreatic cancer. Ann Intern Med. 2004;141(10):753-63.
65. Iglesias Garcia J, Larino Noia J, Dominguez Munoz JE. Endoscopic ultrasound in the diagnosis and staging of pancreatic cancer. Rev Esp Enferm Dig. 2009;101(9):631-8.
66. Rosch T, Lorenz R, Braig C, Feuerbach S, Siewert JR, Schusdziarra V, et al. Endoscopic ultrasound in pancreatic tumor diagnosis. Gastrointest Endosc. 1991;37(3):347-52.
67. Eloubeidi MA, Tamhane A. EUS-guided FNA of solid pancreatic masses: a learning curve with 300 consecutive procedures. Gastrointest Endosc. 2005;61(6):700-8.
68. Hewitt MJ, McPhail MJ, Possamai L, Dhar A, Vlavianos P, Monahan KJ. EUS-guided FNA for diagnosis of solid pancreatic neoplasms: a meta-analysis. Gastrointest Endosc. 2012;75(2):319-31.
69. Affi A, Vazquez-Sequeiros E, Norton ID, Clain JE, Wiersema MJ. Acute extraluminal hemorrhage associated with EUS-guided fine needle aspiration: frequency and clinical significance. Gastrointest Endosc. 2001;53(2):221-5.
70. Yoshinaga S, Suzuki H, Oda I, Saito Y. Role of endoscopic ultrasound-guided fine needle aspiration (EUS-FNA) for diagnosis of solid pancreatic masses. Dig Endosc. 2011;23 Suppl 1:29-33.
71. Wang W, Shpaner A, Krishna SG, Ross WA, Bhutani MS, Tamm EP, et al. Use of EUS-FNA in diagnosing pancreatic neoplasm without a definitive mass on CT. Gastrointest Endosc. 2013;78(1):73-80.
72. O'Toole D, Palazzo L, Arotcarena R, Dancour A, Aubert A, Hammel P, et al. Assessment of complications of EUS-guided fine-needle aspiration. Gastrointest Endosc. 2001;53(4):470-4.
73. Micames C, Jowell PS, White R, Paulson E, Nelson R, Morse M, et al. Lower frequency of peritoneal carcinomatosis in patients with pancreatic cancer diagnosed by EUS-guided FNA vs. percutaneous FNA. Gastrointest Endosc. 2003;58(5):690-5.

74. Eloubeidi MA, Jhala D, Chhieng DC, Chen VK, Eltoum I, Vickers S, et al. Yield of endoscopic ultrasound-guided fine-needle aspiration biopsy in patients with suspected pancreatic carcinoma. Cancer. 2003;99(5):285-92.
75. Puli SR, Kalva N, Bechtold ML, Pamulaparthy SR, Cashman MD, Estes NC, et al. Diagnostic accuracy of endoscopic ultrasound in pancreatic neuroendocrine tumors: a systematic review and meta analysis. World J Gastroenterol. 2013;19(23):3678-84.
76. Ardengh JC, Lopes CV, Campos AD, Pereira de Lima LF, Venco F, Modena JL. Endoscopic ultrasound and fine needle aspiration in chronic pancreatitis: differential diagnosis between pseudotumoral masses and pancreatic cancer. Jop. 2007;8(4):413-21.
77. Takahashi K, Yamao K, Okubo K, Sawaki A, Mizuno N, Ashida R, et al. Differential diagnosis of pancreatic cancer and focal pancreatitis by using EUS-guided FNA. Gastrointest Endosc. 2005;61(1):76-9.
78. Varadarajulu S, Tamhane A, Eloubeidi MA. Yield of EUS-guided FNA of pancreatic masses in the presence or the absence of chronic pancreatitis. Gastrointest Endosc. 2005;62(5):728-36; quiz 51, 53.
79. Hwang CY, Lee SS, Song TJ, Moon SH, Lee D, Park do H, et al. Endoscopic ultrasound guided fine needle aspiration biopsy in diagnosis of pancreatic and peripancreatic lesions: a single center experience in Korea. Gut Liver. 2009;3(2):116-21.
80. Touchefeu Y, Le Rhun M, Coron E, Alamdari A, Heymann MF, Mosnier JF, et al. Endoscopic ultrasound-guided fine-needle aspiration for the diagnosis of solid pancreatic masses: the impact on patient-management strategy. Aliment Pharmacol Ther. 2009;30(10):1070-7.
81. Agarwal B, Abu-Hamda E, Molke KL, Correa AM, Ho L. Endoscopic ultrasound-guided fine needle aspiration and multidetector spiral CT in the diagnosis of pancreatic cancer. Am J Gastroenterol. 2004;99(5):844-50.
82. Stewart CJ, Mills PR, Carter R, O'Donohue J, Fullarton G, Imrie CW, et al. Brush cytology in the assessment of pancreatico-biliary strictures: a review of 406 cases. Journal of clinical pathology. 2001;54(6):449-55.
83. Wakatsuki T, Irisawa A, Bhutani MS, Hikichi T, Shibukawa G, Takagi T, et al. Comparative study of diagnostic value of cytologic sampling by endoscopic ultrasonography-guided fine-needle aspiration and that by endoscopic retrograde pancreatography for the management of pancreatic mass without biliary stricture. J Gastroenterol Hepatol. 2005;20(11):1707-11.
84. Valls C, Andía E, Sanchez A, Fabregat J, Pozuelo O, Quintero JC, et al. Dual-Phase Helical CT of Pancreatic Adenocarcinoma. American Journal of Roentgenology. 2002;178(4):821-6.
85. Glasbrenner B, Ardan M, Boeck W, Preclik G, Moller P, Adler G. Prospective evaluation of brush cytology of biliary strictures during endoscopic retrograde cholangiopancreatography. Endoscopy. 1999;31(9):712-7.
86. Rosch T, Hofrichter K, Frimberger E, Meining A, Born P, Weigert N, et al. ERCP or EUS for tissue diagnosis of biliary strictures? A prospective comparative study. Gastrointest Endosc. 2004;60(3):390-6.
87. Layfield LJ, Wax TD, Lee JG, Cotton PB. Accuracy and morphologic aspects of pancreatic and biliary duct brushings. Acta Cytol. 1995;39(1):11-8.
88. McGuire DE, Venu RP, Brown RD, Etzkorn KP, Glaws WR, Abu-Hammour A. Brush cytology for pancreatic carcinoma: an analysis of factors influencing results. Gastrointest Endosc. 1996;44(3):300-4.
89. Logrono R, Kurtycz DF, Molina CP, Trivedi VA, Wong JY, Block KP. Analysis of false-negative diagnoses on endoscopic brush cytology of biliary and pancreatic duct strictures: the experience at 2 university hospitals. Arch Pathol Lab Med. 2000;124(3):387-92.
90. Mitsuhashi T, Ghafari S, Chang CY, Gu M. Endoscopic ultrasound-guided fine needle aspiration of the pancreas: cytomorphological evaluation with emphasis on adequacy assessment, diagnostic criteria and contamination from the gastrointestinal tract. Cytopathology. 2006;17(1):34-41.
91. Layfield LJ, Jarboe EA. Cytopathology of the pancreas: neoplastic and nonneoplastic entities. Ann Diagn Pathol. 2010;14(2):140-51.
92. Chhieng D, Stelow E. Pancreatic Ductal Adenocarcinoma and Its Variants. Pancreatic Cytopathology. Essentials in Cytopathology. 3: Springer US; 2007. p. 35-66.
93. Schwartz MR. Endoscopic ultrasound-guided fine-needle aspiration. Cancer Cytopathology. 2004;102(4):203-6.
94. Carroll N PI. Carroll N, Penman I. London [cited 2014July23]. BSG: Recommendations for training in endosonography UK EUS Users: Available from. Available from: http://www.bsg.org.uk/clinical-guidance/endoscopy/guidance-for-endoscopic-ultrasonography-uk-eus-users-groups-*.html.
95. Iglesias-Garcia J, Dominguez-Munoz JE, Abdulkader I, Larino-Noia J, Eugenyeva E, Lozano-Leon A, et al. Influence of on-site cytopathology evaluation on the diagnostic accuracy of endoscopic ultrasound-guided fine needle aspiration (EUS-FNA) of solid pancreatic masses. Am J Gastroenterol. 2011;106(9):1705-10.

96. Abdallah AA, Krige JE, Bornman PC. Biliary tract obstruction in chronic pancreatitis. HPB (Oxford). 2007;9(6):421-8.
97. van Gulik TM, Reeders JW, Bosma A, Moojen TM, Smits NJ, Allema JH, et al. Incidence and clinical findings of benign, inflammatory disease in patients resected for presumed pancreatic head cancer. Gastrointest Endosc. 1997;46(5):417-23.
98. Krishna NB, Mehra M, Reddy AV, Agarwal B. EUS/EUS-FNA for suspected pancreatic cancer: influence of chronic pancreatitis and clinical presentation with or without obstructive jaundice on performance characteristics. Gastrointest Endosc. 2009;70(1):70-9.
99. Euscher ED, Marsh WL, Jr., Lucas JG, Frankel WL. Histologic and immunohistochemical changes in the stented common bile duct. Appl Immunohistochem Mol Morphol. 2007;15(3):299-304.
100. Fritscher-Ravens A, Brand L, Knofel WT, Bobrowski C, Topalidis T, Thonke F, et al. Comparison of endoscopic ultrasound-guided fine needle aspiration for focal pancreatic lesions in patients with normal parenchyma and chronic pancreatitis. Am J Gastroenterol. 2002;97(11):2768-75.
101. Stelow EB, Bardales RH, Lai R, Mallery S, Linzie BM, Crary GS, et al. The cytological spectrum of chronic pancreatitis. Diagn Cytopathol. 2005;32(2):65-9.
102. Harewood GC, Wiersema LM, Halling AC, Keeney GL, Salamao DR, Wiersema MJ. Influence of EUS training and pathology interpretation on accuracy of EUS-guided fine needle aspiration of pancreatic masses. Gastrointest Endosc. 2002;55(6):669-73.
103. Alsibai KD, Denis B, Bottlaender J, Kleinclaus I, Straub P, Fabre M. Impact of cytopathologist expert on diagnosis and treatment of pancreatic lesions in current clinical practice. A series of 106 endoscopic ultrasound-guided fine needle aspirations. Cytopathology. 2006;17(1):18-26.
104. Lin F, Staerkel G. Cytologic criteria for well differentiated adenocarcinoma of the pancreas in fine-needle aspiration biopsy specimens. Cancer. 2003;99(1):44-50.
105. Ylagan LR, Edmundowicz S, Kasal K, Walsh D, Lu DW. Endoscopic ultrasound guided fine-needle aspiration cytology of pancreatic carcinoma: a 3-year experience and review of the literature. Cancer. 2002;96(6):362-9.
106. Afify AM, al-Khafaji BM, Kim B, Scheiman JM. Endoscopic ultrasound-guided fine needle aspiration of the pancreas. Diagnostic utility and accuracy. Acta Cytol. 2003;47(3):341-8.
107. Woolf KM, Liang H, Sletten ZJ, Russell DK, Bonfiglio TA, Zhou Z. False-negative rate of endoscopic ultrasound-guided fine-needle aspiration for pancreatic solid and cystic lesions with matched surgical resections as the gold standard: one institution's experience. Cancer Cytopathol. 2013;121(8):449-58.
108. Williams DB, Sahai AV, Aabakken L, Penman ID, van Velse A, Webb J, et al. Endoscopic ultrasound guided fine needle aspiration biopsy: a large single centre experience. Gut. 1999;44(5):720-6.
109. Mishra G, Zhao Y, Sweeney J, Pineau BC, Case D, Ho C, et al. Determination of qualitative telomerase activity as an adjunct to the diagnosis of pancreatic adenocarcinoma by EUS-guided fine-needle aspiration. Gastrointest Endosc. 2006;63(4):648-54.
110. Siddiqui AA, Kowalski TE, Shahid H, O'Donnell S, Tolin J, Loren DE, et al. False-positive EUS-guided FNA cytology for solid pancreatic lesions. Gastrointest Endosc. 2011;74(3):535-40.
111. Gleeson FC, Kipp BR, Caudill JL, Clain JE, Clayton AC, Halling KC, et al. False positive endoscopic ultrasound fine needle aspiration cytology: incidence and risk factors. Gut. 2010;59(5):586-93.
112. de la Fuente SG, Ceppa EP, Reddy SK, Clary BM, Tyler DS, Pappas TN. Incidence of benign disease in patients that underwent resection for presumed pancreatic cancer diagnosed by endoscopic ultrasonography (EUS) and fine-needle aspiration (FNA). J Gastrointest Surg. 2010;14(7):1139-42.
113. Raut CP, Grau AM, Staerkel GA, Kaw M, Tamm EP, Wolff RA, et al. Diagnostic accuracy of endoscopic ultrasound-guided fine-needle aspiration in patients with presumed pancreatic cancer. J Gastrointest Surg. 2003;7(1):118-26; discussion 27-8.
114. Lachter J, Rosenthal Y, Kluger Y. A multidisciplinary survey on controversies in the use of EUS-guided FNA: assessing perspectives of surgeons, oncologists and gastroenterologists. BMC Gastroenterol. 2011;11:117.
115. Harewood GC, Wiersema MJ. Endosonography-guided fine needle aspiration biopsy in the evaluation of pancreatic masses. Am J Gastroenterol. 2002;97(6):1386-91.
116. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther. 2001;69(3):89-95.
117. Henry NL, Hayes DF. Cancer biomarkers. Molecular Oncology. 2012;6(2):140-6.
118. Frangogiannis NG. Biomarkers: hopes and challenges in the path from discovery to clinical practice. Transl Res. 2012;159(4):197-204.
119. Lee JW, Figeys D, Vasilescu J. Biomarker assay translation from discovery to clinical studies in cancer drug development: quantification of emerging protein biomarkers. Adv Cancer Res. 2007;96:269-98.
120. Smith NR, Womack C. A matrix approach to guide IHC-based tissue biomarker development in oncology drug discovery. J Pathol. 2014;232(2):190-8.

121. Sim SC, Ingelman-Sundberg M. Pharmacogenomic biomarkers: new tools in current and future drug therapy. Trends in Pharmacological Sciences. 2011;32(2):72-81.
122. O'Hurley G, Sjostedt E, Rahman A, Li B, Kampf C, Ponten F, et al. Garbage in, garbage out: A critical evaluation of strategies used for validation of immunohistochemical biomarkers. Mol Oncol. 2014;8(4):783-98.
123. Allred DC. Issues and updates: evaluating estrogen receptor-alpha, progesterone receptor, and HER2 in breast cancer. Mod Pathol. 2010;23 Suppl 2:S52-9.
124. Debiec-Rychter M, Wasag B, Stul M, De Wever I, Van Oosterom A, Hagemeijer A, et al. Gastrointestinal stromal tumours (GISTs) negative for KIT (CD117 antigen) immunoreactivity. J Pathol. 2004;202(4):430-8.
125. Shah RB, Zhou M, LeBlanc M, Snyder M, Rubin MA. Comparison of the basal cell-specific markers, 34betaE12 and p63, in the diagnosis of prostate cancer. Am J Surg Pathol. 2002;26(9):1161-8.
126. Duffy MJ. Carcinoembryonic antigen as a marker for colorectal cancer: is it clinically useful? Clin Chem. 2001;47(4):624-30.
127. Barry MJ. PSA screening for prostate cancer: the current controversy--a viewpoint. Patient Outcomes Research Team for Prostatic Diseases. Ann Oncol. 1998;9(12):1279-82.
128. Rifai N, Gillette MA, Carr SA. Protein biomarker discovery and validation: the long and uncertain path to clinical utility. Nat Biotechnol. 2006;24(8):971-83.
129. Paulovich AG, Whiteaker JR, Hoofnagle AN, Wang P. The interface between biomarker discovery and clinical validation: The tar pit of the protein biomarker pipeline. Proteomics Clin Appl. 2008;2(10-11):1386-402.
130. Hustinx SR, Cao D, Maitra A, Sato N, Martin ST, Sudhir D, et al. Differentially expressed genes in pancreatic ductal adenocarcinomas identified through serial analysis of gene expression. Cancer Biol Ther. 2004;3(12):1254-61.
131. Castronovo V, Wang YH, Musmeci D, Dumont B, Turtoi A. Proteomic analysis of human pancreas cancers for the identification of targetable biomarkers. FASEB Journal Conference: Experimental Biology. 2010;20100424(20100428).
132. Logsdon C, Simeone D, Binkley C, Arumugam T, Greenson J, Giordano T, et al. Molecular profiling of pancreatic adenocarcinoma and chronic pancreatitis identifies multiple genes differentially regulated in pancreatic cancer. Cancer Res. 2003;63:2649 - 57.
133. Iacobuzio-Donahue CA, Maitra A, Olsen M, Lowe AW, van Heek NT, Rosty C, et al. Exploration of global gene expression patterns in pancreatic adenocarcinoma using cDNA microarrays. Am J Pathol. 2003;162(4):1151-62.
134. Chen JH, Ni RZ, Xiao MB, Guo JG, Zhou JW. Comparative proteomic analysis of differentially expressed proteins in human pancreatic cancer tissue. Hepatobiliary and Pancreatic Diseases International. 2009;8(2):193-200.
135. Chen R, Pan S, Brentnall TA, Aebersold R. Proteomic profiling of pancreatic cancer for biomarker discovery. Mol Cell Proteomics. 2005;4(4):523-33.
136. Shen J, Person MD, Zhu J, Abbruzzese JL, Li D. Protein expression profiles in pancreatic adenocarcinoma compared with normal pancreatic tissue and tissue affected by pancreatitis as detected by two-dimensional gel electrophoresis and mass spectrometry. Cancer Res. 2004;64(24):9018-26.
137. Chen J, Chen LJ, Yang RB, Xia YL, Zhou HC, Wu W, et al. Expression and clinical significance of apolipoprotein E in pancreatic ductal adenocarcinoma. Med Oncol. 2013;30(2):583.
138. Cummings J, Raynaud F, Jones L, Sugar R, Dive C. Fit-for-purpose biomarker method validation for application in clinical trials of anticancer drugs. Br J Cancer. 2010;103(9):1313-7.
139. Zapata M, Cohen C, Siddiqui MT. Immunohistochemical expression of SMAD4, CK19, and CA19-9 in fine needle aspiration samples of pancreatic adenocarcinoma: Utility and potential role. Cytojournal. 2007;4(13).
140. Gao L, Antic T, Hyjek E, Gong C, Mueller J, Waxman I, et al. Immunohistochemical analysis of E-cadherin and zeste homolog 2 expression in endoscopic ultrasound-guided fine-needle aspiration of pancreatic adenocarcinoma. Cancer Cytopathol. 2013;121(11):644-52.
141. Ali S, Cohen C, Little JV, Sequeira JH, Mosunjac MB, Siddiqui MT. The utility of SMAD4 as a diagnostic immunohistochemical marker for pancreatic adenocarcinoma, and its expression in other solid tumors. Diagn Cytopathol. 2007;35(10):644-8.
142. Drucker E, Krapfenbauer K. Pitfalls and limitations in translation from biomarker discovery to clinical utility in predictive and personalised medicine. Epma j. 2013;4(1):7.
143. Gnanapragasam VJ. Unlocking the molecular archive: the emerging use of formalin-fixed paraffin-embedded tissue for biomarker research in urological cancer. BJU Int. 2010;105(2):274-8.
144. Fairley JA, Gilmour K, Walsh K. Making the most of pathological specimens: molecular diagnosis in formalin-fixed, paraffin embedded tissue. Curr Drug Targets. 2012;13(12):1475-87.

145. Ioannidis JP, Panagiotou OA. Comparison of effect sizes associated with biomarkers reported in highly cited individual articles and in subsequent meta-analyses. Jama. 2011;305(21):2200-10.
146. Taylor CR. Standardization in immunohistochemistry: the role of antigen retrieval in molecular morphology. Biotech Histochem. 2006;81(1):3-12.
147. Matos LL, Trufelli DC, de Matos MG, da Silva Pinhal MA. Immunohistochemistry as an important tool in biomarkers detection and clinical practice. Biomark Insights. 2010;5:9-20.
148. Jambhekar NA, Chaturvedi AC, Madur BP. Immunohistochemistry in surgical pathology practice: a current perspective of a simple, powerful, yet complex, tool. Indian J Pathol Microbiol. 2008;51(1):2-11.
149. Chang DK, Jamieson NB, Johns AL, Scarlett CJ, Pajic M, Chou A, et al. Histomolecular phenotypes and outcome in adenocarcinoma of the ampulla of vater. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2013;31(10):1348-56.
150. Jamieson NB, Morran DC, Morton JP, Ali A, Dickson EJ, Carter CR, et al. MicroRNA molecular profiles associated with diagnosis, clinicopathologic criteria, and overall survival in patients with resectable pancreatic ductal adenocarcinoma. Clin Cancer Res. 2012;18(2):534-45.
151. Weinstein MH, Signoretti S, Loda M. Diagnostic utility of immunohistochemical staining for p63, a sensitive marker of prostatic basal cells. Mod Pathol. 2002;15(12):1302-8.
152. Urruticoechea A, Smith IE, Dowsett M. Proliferation marker Ki-67 in early breast cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2005;23(28):7212-20.
153. Chang F, Vu C, Chandra A, Meenan J, Herbert A. Endoscopic ultrasound-guided fine needle aspiration cytology of pancreatic neuroendocrine tumours: cytomorphological and immunocytochemical evaluation. Cytopathology. 2006;17(1):10-7.
154. Gu M, Ghafari S, Lin F, Ramzy I. Cytological diagnosis of endocrine tumors of the pancreas by endoscopic ultrasound-guided fine-needle aspiration biopsy. Diagn Cytopathol. 2005;32(4):204-10.
155. Acs G, Lawton TJ, Rebbeck TR, LiVolsi VA, Zhang PJ. Differential expression of E-cadherin in lobular and ductal neoplasms of the breast and its biologic and diagnostic implications. American journal of clinical pathology. 2001;115(1):85-98.
156. Jung S, Wu C, Eslami Z, Tanguay S, Aprikian A, Kassouf W, et al. The role of immunohistochemistry in the diagnosis of flat urothelial lesions: a study using CK20, CK5/6, P53, Cd138, and Her2/Neu. Ann Diagn Pathol. 2014;18(1):27-32.
157. Brunnstrom H, Johansson L, Jirstrom K, Jonsson M, Jonsson P, Planck M. Immunohistochemistry in the differential diagnostics of primary lung cancer: an investigation within the Southern Swedish Lung Cancer Study. American journal of clinical pathology. 2013;140(1):37-46.
158. Kosarac O, Takei H, Zhai QJ, Schwartz MR, Mody DR. S100P and XIAP expression in pancreatic ductal adenocarcinoma: potential novel biomarkers as a diagnostic adjunct to fine needle aspiration cytology. Acta Cytol. 2011;55(2):142-8.
159. Kosarac O, Takei H, Zhai QJ, Schwartz MR, Mody DR. S100P and XIAP expression in pancreatic ductal adenocarcinoma: potential novel biomarkers as a diagnostic adjunct to fine needle aspiration cytology. Acta Cytologica. 2011;55(2):142-8.
160. Bartlett JM, Brookes CL, Robson T, van de Velde CJ, Billingham LJ, Campbell FM, et al. Estrogen receptor and progesterone receptor as predictive biomarkers of response to endocrine therapy: a prospectively powered pathology study in the Tamoxifen and Exemestane Adjuvant Multinational trial. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2011;29(12):1531-8.
161. Harvey JM, Clark GM, Osborne CK, Allred DC. Estrogen receptor status by immunohistochemistry is superior to the ligand-binding assay for predicting response to adjuvant endocrine therapy in breast cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 1999;17(5):1474-81.
162. Fuzery AK, Levin J, Chan MM, Chan DW. Translation of proteomic biomarkers into FDA approved cancer diagnostics: issues and challenges. Clin Proteomics. 2013;10(1):13.
163. Cuzick J, Dowsett M, Pineda S, Wale C, Salter J, Quinn E, et al. Prognostic value of a combined estrogen receptor, progesterone receptor, Ki-67, and human epidermal growth factor receptor 2 immunohistochemical score and comparison with the Genomic Health recurrence score in early breast cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2011;29(32):4273-8.
164. Agarwal B, Ludwig OJ, Collins BT, Cortese C. Immunostaining as an adjunct to cytology for diagnosis of pancreatic adenocarcinoma. Clin Gastroenterol Hepatol. 2008;6(12):1425-31.
165. Awadallah NS, Shroyer KR, Langer DA, Torkko KC, Chen YK, Bentz JS, et al. Detection of B7-H4 and p53 in pancreatic cancer: potential role as a cytological diagnostic adjunct. Pancreas. 2008;36(2):200-6.

166. Lin F, Shi J, Liu H, Hull ME, Dupree W, Prichard JW, et al. Diagnostic utility of S100P and von Hippel-Lindau gene product (pVHL) in pancreatic adenocarcinoma - With implication of their roles in early tumorigenesis. American Journal of Surgical Pathology. 2008;32(1):78-91.
167. Liu H, Shi J, Anandan V, Wang HL, Diehl D, Blansfield J, et al. Reevaluation and identification of the best immunohistochemical panel (pVHL, Maspin, S100P, IMP-3) for ductal adenocarcinoma of the pancreas. Arch Pathol Lab Med. 2012;136(6):601-9.
168. Kopelman Y, Marmor S, Ashkenazi I, Fireman Z. Value of EUS-FNA cytological preparations compared with cell block sections in the diagnosis of pancreatic solid tumours. Cytopathology. 2011;22(3):174-8.
169. Iwao T, Tsuchida A, Hanada K, Eguchi N, Kajiyama G, Shimamoto F. Immunocytochemical detection of p53 protein as an adjunct in cytologic diagnosis from pancreatic duct brushings in mucin-producing tumors of the pancreas. Cancer. 1997;81(3):163-71.
170. Chhieng DC, Benson E, Eltoum I, Eloubeidi MA, Jhala N, Jhala D, et al. MUC1 and MUC2 expression in pancreatic ductal carcinoma obtained by fine-needle aspiration. Cancer. 2003;99(6):365-71.
171. Bhardwaj A, Marsh WL, Jr., Nash JW, Barbacioru CC, Jones S, Frankel WL. Double immunohistochemical staining with MUC4/p53 is useful in the distinction of pancreatic adenocarcinoma from chronic pancreatitis: a tissue microarray-based study. Arch Pathol Lab Med. 2007;131(4):556-62.
172. Agarwal B, Ludwig OJ, Collins BT, Cortese C. Immunostaining as an Adjunct to Cytology for Diagnosis of Pancreatic Adenocarcinoma. Clinical Gastroenterology and Hepatology. 2008;6(12):1425-31.
173. Jhala N, Jhala D, Vickers SM, Eltoum I, Batra SK, Manne U, et al. Biomarkers in Diagnosis of pancreatic carcinoma in fine-needle aspirates. American journal of clinical pathology. 2006;126(4):572-9.
174. Toll AD, Witkiewicz AK, Bibbo M. Expression of K homology domain containing protein (KOC) in pancreatic cytology with corresponding histology. Acta Cytol. 2009;53(2):123-9.
175. Ligato S, Zhao H, Mandich D, Cartun RW. KOC (K homology domain containing protein overexpressed in cancer) and S100A4-protein immunoreactivity improves the diagnostic sensitivity of biliary brushing cytology for diagnosing pancreaticobiliary malignancies. Diagn Cytopathol. 2008;36(8):561-7.
176. Giorgadze TA, Peterman H, Baloch ZW, Furth EE, Pasha T, Shiina N, et al. Diagnostic utility of mucin profile in fine-needle aspiration specimens of the pancreas: an immunohistochemical study with surgical pathology correlation. Cancer. 2006;108(3):186-97.
177. Deng H, Shi J, Wilkerson M, Meschter S, Dupree W, Lin F. Usefulness of S100P in diagnosis of adenocarcinoma of pancreas on fine-needle aspiration biopsy specimens. American journal of clinical pathology. 2008;129(1):81-8.
178. Phillips KA, Van Bebber S, Issa AM. Diagnostics and biomarker development: priming the pipeline. Nat Rev Drug Discov. 2006;5(6):463-9.
179. Shin JH, Bae JH, Lee A, Jung CK, Yim HW, Park JS, et al. CK7, CK20, CDX2 and MUC2 Immunohistochemical staining used to distinguish metastatic colorectal carcinoma involving ovary from primary ovarian mucinous adenocarcinoma. Jpn J Clin Oncol. 2010;40(3):208-13.
180. Yantiss RK, Woda BA, Fanger GR, Kalos M, Whalen GF, Tada H, et al. KOC (K homology domain containing protein overexpressed in cancer): a novel molecular marker that distinguishes between benign and malignant lesions of the pancreas. Am J Surg Pathol. 2005;29(2):188-95.
181. Issaq HJ, Waybright TJ, Veenstra TD. Cancer biomarker discovery: Opportunities and pitfalls in analytical methods. Electrophoresis. 2011;32(9):967-75.
182. Jakobsen JN, Santoni-Rugiu E, Ravn J, Sorensen JB. Intratumour variation of biomarker expression by immunohistochemistry in resectable non-small cell lung cancer. Eur J Cancer. 2013;49(11):2494-503.
183. Jakobsen JN, Sorensen JB. Intratumor heterogeneity and chemotherapy-induced changes in EGFR status in non-small cell lung cancer. Cancer Chemother Pharmacol. 2012;69(2):289-99.
184. Duxbury MS, Matros E, Clancy T, Bailey G, Doff M, Zinner MJ, et al. CEACAM6 is a novel biomarker in pancreatic adenocarcinoma and PanIN lesions. Ann Surg. 2005;241(3):491-6.
185. Okami J, Yamamoto H, Fujiwara Y, Tsujie M, Kondo M, Noura S, et al. Overexpression of cyclooxygenase-2 in carcinoma of the pancreas. Clinical Cancer Research. 1999;5(8):2018-24.
186. Wachter DL, Schlabrakowski A, Hoegel J, Kristiansen G, Hartmann A, Riener MO. Diagnostic value of immunohistochemical IMP3 expression in core needle biopsies of pancreatic ductal adenocarcinoma. American Journal of Surgical Pathology. 2011;35(6):873-7.
187. Lee CS, Rush M, Charalambous D, Rode J. Immunohistochemical demonstration of the p53 tumour suppressor gene product in cancer of the pancreas and chronic pancreatitis. J Gastroenterol Hepatol. 1993;8(5):465-9.
188. Wang Y, Gao J, Li Z, Jin Z, Gong Y, Man X. Diagnostic value of mucins (MUC1, MUC2 and MUC5AC) expression profile in endoscopic ultrasound-guided fine-needle aspiration specimens of the pancreas. Int J Cancer. 2007;121(12):2716-22.

189. Zlobec I, Steele R, Terracciano L, Jass JR, Lugli A. Selecting immunohistochemical cut-off scores for novel biomarkers of progression and survival in colorectal cancer. Journal of clinical pathology. 2007;60(10):1112-6.
190. Budczies J, Klauschen F, Sinn BV, Gyorffy B, Schmitt WD, Darb-Esfahani S, et al. Cutoff Finder: a comprehensive and straightforward Web application enabling rapid biomarker cutoff optimization. PLoS One. 2012;7(12):e51862.
191. Burrell RA, McGranahan N, Bartek J, Swanton C. The causes and consequences of genetic heterogeneity in cancer evolution. Nature. 2013;501(7467):338-45.
192. Verma M. Pancreatic cancer biomarkers and their implication in cancer diagnosis and epidemiology. Cancers (Basel). 2010;2(4):1830-7.
193. Anagnostou VK, Welsh AW, Giltnane JM, Siddiqui S, Liceaga C, Gustavson M, et al. Analytic variability in immunohistochemistry biomarker studies. Cancer Epidemiol Biomarkers Prev. 2010;19(4):982-91.
194. Emoto K, Yamashita S, Okada Y. Mechanisms of heat-induced antigen retrieval: does pH or ionic strength of the solution play a role for refolding antigens? J Histochem Cytochem. 2005;53(11):1311-21.
195. Vassallo J, Pinto GA, Alvarenga JM, Zeferino LC, Chagas CA, Metze K. Comparison of immunoexpression of 2 antibodies for estrogen receptors (1D5 and 6F11) in breast carcinomas using different antigen retrieval and detection methods. Appl Immunohistochem Mol Morphol. 2004;12(2):177-82.
196. Hermansen SK, Christensen KG, Jensen SS, Kristensen BW. Inconsistent immunohistochemical expression patterns of four different CD133 antibody clones in glioblastoma. J Histochem Cytochem. 2011;59(4):391-407.
197. McCabe A, Dolled-Filhart M, Camp RL, Rimm DL. Automated Quantitative Analysis (AQUA) of In Situ Protein Expression, Antibody Concentration, and Prognosis. Journal of the National Cancer Institute. 2005;97(24):1808-15.
198. Carden CP, Sarker D, Postel-Vinay S, Yap TA, Attard G, Banerji U, et al. Can molecular biomarker-based patient selection in Phase I trials accelerate anticancer drug development? Drug Discov Today. 2010;15(3-4):88-97.
199. Wagner PD, Srivastava S. New paradigms in translational science research in cancer biomarkers. Transl Res. 2012;159(4):343-53.
200. Neoptolemos JP, Dunn JA, Stocken DD, Almond J, Link K, Beger H, et al. Adjuvant chemoradiotherapy and chemotherapy in resectable pancreatic cancer: a randomised controlled trial. Lancet. 2001;358(9293):1576-85.
201. Chiang SC, Han CL, Yu KH, Chen YJ, Wu KP. Prioritization of cancer marker candidates based on the immunohistochemistry staining images deposited in the human protein atlas. PLoS One. 2013;8(11):e81079.
202. Sauerland S, Seiler CM. Role of systematic reviews and meta-analysis in evidence-based medicine. World J Surg. 2005;29(5):582-7.
203. Yuan Y, Hunt RH. Systematic reviews: the good, the bad, and the ugly. Am J Gastroenterol. 2009;104(5):1086-92.
204. Nocito A, Kononen J, Kallioniemi OP, Sauter G. Tissue microarrays (TMAs) for high-throughput molecular pathology research. Int J Cancer. 2001;94(1):1-5.
205. Kononen J, Bubendorf L, Kallioniemi A, Barlund M, Schraml P, Leighton S, et al. Tissue microarrays for high-throughput molecular profiling of tumor specimens. Nat Med. 1998;4(7):844-7.
206. Packeisen J, Korsching E, Herbst H, Boecker W, Buerger H. Demystified...tissue microarray technology. Mol Pathol. 2003;56(4):198-204.
207. Camp RL, Charette LA, Rimm DL. Validation of tissue microarray technology in breast carcinoma. Lab Invest. 2000;80(12):1943-9.
208. Mobasheri A, Airley R, Foster CS, Schulze-Tanzil G, Shakibaei M. Post-genomic applications of tissue microarrays: basic research, prognostic oncology, clinical genomics and drug discovery. Histol Histopathol. 2004;19(1):325-35.
209. Watanabe A, Cornelison R, Hostetter G. Tissue microarrays: applications in genomic research. Expert Rev Mol Diagn. 2005;5(2):171-81.
210. Ali A, Brown V, Denley S, Jamieson N, Morton J, Nixon C, et al. Expression of KOC, S100P, mesothelin and MUC1 in pancreatico-biliary adenocarcinomas: development and utility of a potential diagnostic immunohistochemistry panel. BMC Clinical Pathology. 2014;14(1):35.
211. Maitra A, Adsay NV, Argani P, Iacobuzio-Donahue C, De Marzo A, Cameron JL, et al. Multicomponent analysis of the pancreatic adenocarcinoma progression model using a pancreatic intraepithelial neoplasia tissue microarray. Mod Pathol. 2003;16(9):902-12.

212. Winter JM, Tang LH, Klimstra DS, Brennan MF, Brody JR, Rocha FG, et al. A novel survival-based tissue microarray of pancreatic cancer validates MUC1 and mesothelin as biomarkers. PLoS One. 2012;7(7):e40157.
213. Biankin AV, Kench JG, Colvin EK, Segara D, Scarlett CJ, Nguyen NQ, et al. Expression of S100A2 calcium-binding protein predicts response to pancreatectomy for pancreatic cancer. Gastroenterology. 2009;137(2):558-68, 68.e1-11.
214. Hassan S, Ferrario C, Mamo A, Basik M. Tissue microarrays: emerging standard for biomarker validation. Curr Opin Biotechnol. 2008;19(1):19-25.
215. Bhutani MS, Hawes RH, Baron PL, Sanders-Cliette A, van Velse A, Osborne JF, et al. Endoscopic ultrasound guided fine needle aspiration of malignant pancreatic lesions. Endoscopy. 1997;29(9):854-8.
216. Erickson RA, Sayage-Rabie L, Beissner RS. Factors predicting the number of EUS-guided fine-needle passes for diagnosis of pancreatic malignancies. Gastrointest Endosc. 2000;51(2):184-90.
217. Yantiss RK, Cosar E, Fischer AH. Use of IMP3 in identification of carcinoma in fine needle aspiration biopsies of pancreas. Acta Cytologica. 2008;52(2):133-8.
218. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol. 2005;58(10):982-90.
219. Di Nisio M, Van Sluis GL, Bossuyt PM, Buller HR, Porreca E, Rutjes AW. Accuracy of diagnostic tests for clinically suspected upper extremity deep vein thrombosis: a systematic review. J Thromb Haemost. 2010;8(4):684-92.
220. Cochrane Collaboration [cited 2012 August 1]. Available from: http://ims.cochrane.org/revman.
221. R Development Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.
222. Philipp Doebler (2012). mada: Meta-Analysis of Diagnostic Accuracy (mada). R package version 0.4.1. http://CRAN.R-project.org/package=mada.
223. Zamora J, Abraira V, Muriel A, Khan K, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. BMC Med Res Methodol. 2006;6:31.
224. Denley SM, Jamieson NB, McCall P, Oien KA, Morton JP, Carter CR, et al. Activation of the IL-6R/Jak/stat pathway is associated with a poor outcome in resected pancreatic ductal adenocarcinoma. J Gastrointest Surg. 2013;17(5):887-98.
225. Morton JP, Timpson P, Karim SA, Ridgway RA, Athineos D, Doyle B, et al. Mutant p53 drives metastasis and overcomes growth arrest/senescence in pancreatic cancer. Proc Natl Acad Sci U S A. 2010;107(1):246-51.
226. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159-74.
227. Karanjawala ZE, Illei PB, Ashfaq R, Infante JR, Murphy K, Pandey A, et al. New markers of pancreatic cancer identified through differential gene expression analyses: Claudin 18 and annexin A8. American Journal of Surgical Pathology. 2008;32(2):188-96.
228. Ali S, Cohen C, Little JV, Sequeira JH, Mosunjac MB, Siddiqui MT. The utility of SMAD4 as a diagnostic immunohistochemical marker for pancreatic adenocarcinoma, and its expression in other solid tumors. Diagnostic Cytopathology. 2007;35(10):644-8.
229. Harsha HC, Kandasamy K, Ranganathan P, Rani S, Ramabadran S, Gollapudi S, et al. A compendium of potential biomarkers of pancreatic cancer. PLoS Med. 2009;6(4):e1000046.
230. Jimeno A, Hidalgo M. Molecular biomarkers: their increasing role in the diagnosis, characterization, and therapy guidance in pancreatic cancer. Mol Cancer Ther. 2006;5(4):787-96.
231. Singh P, Srinivasan R, Wig JD. Major molecular markers in pancreatic ductal adenocarcinoma and their roles in screening, diagnosis, prognosis, and treatment. Pancreas. 2011;40(5):644-52.
232. Cao D, Zhang Q, Wu LSF, Salaria SN, Winter JW, Hruban RH, et al. Prognostic significance of maspin in pancreatic ductal adenocarcinoma: Tissue microarray analysis of 223 surgically resected cases. Modern Pathology. 2007;20(5):570-8.
233. Lim YJ, Lee JK, Jang WY, Song SY, Lee KT, Paik SW, et al. Prognostic significance of maspin in pancreatic ductal adenocarcinoma. Korean Journal of Internal Medicine. 2004;19(1):15-8.
234. Maass N, Hojo T, Ueding M, Luttges J, Kloppel G, Jonat W, et al. Expression of the tumor suppressor gene Maspin in human pancreatic cancers. Clinical Cancer Research. 2001;7(4):812-7.
235. Nash JW, Bhardwaj A, Wen P, Frankel WL. Maspin is useful in the distinction of pancreatic adenocarcinoma from chronic pancreatitis: a tissue microarray based study. Appl Immunohistochem Mol Morphol. 2007;15(1):59-63.
236. Argani P, Iacobuzio-Donahue C, Ryu B, Rosty C, Goggins M, Wilentz RE, et al. Mesothelin is overexpressed in the vast majority of ductal adenocarcinomas of the pancreas: Identification of a new

pancreatic cancer marker by serial analysis of gene expression (SAGE). Clinical Cancer Research. 2001;7(12):3862-8.
237. Hassan R, Laszik ZG, Lerner M, Raffeld M, Postier R, Brackett D. Mesothelin is overexpressed in pancreaticobiliary adenocarcinomas but not in normal pancreas and chronic pancreatitis. American journal of clinical pathology. 2005;124(6):838-45.
238. Ordonez NG. Application of Mesothelin Immunostaining in Tumor Diagnosis. American Journal of Surgical Pathology. 2003;27(11):1418-28.
239. Glass JP, Parasher G, Arias-Pulido H, Donohue R, Prossnitz ER, Cerilli LA. Mesothelin and GPR30 staining among a spectrum of pancreatic epithelial neoplasms. Int J Surg Pathol. 2011;19(5):588-96.
240. Toll AD, Witkiewicz AK, Bibbo M. Expression of K homology domain containing protein (KOC) in pancreatic cytology with corresponding histology. Acta Cytologica. 2009;53(2):123-9.
241. Yantiss RK, Woda BA, Fanger GR, Kalos M, Whalen GF, Tada H, et al. KOC (K homology domain containing protein overexpressed in cancer): A novel molecular marker that distinguishes between benign and malignant lesions of the pancreas. American Journal of Surgical Pathology. 2005;29(2):188-95.
242. Lu Z, Hu L, Evers S, Chen J, Shen Y. Differential expression profiling of human pancreatic adenocarcinoma and healthy pancreatic tissue. Proteomics. 2004;4(12):3975-88.
243. Maitra A, Iacobuzio-Donahue C, Rahman A, Sohn TA, Argani P, Meyer R, et al. Immunohistochemical validation of a novel epithelial and a novel stromal marker of pancreatic ductal adenocarcinoma identified by global expression microarrays: Sea urchin fascin homolog and heat shock protein 47. American journal of clinical pathology. 2002;118(1):52-9.
244. Maitra A, Ashfaq R, Gunn CR, Rahman A, Yeo CJ, Sohn TA, et al. Cyclooxygenase 2 expression in pancreatic adenocarcinoma and pancreatic intraepithelial neoplasia: an immunohistochemical analysis with automated cellular imaging. American journal of clinical pathology. 2002;118(2):194-201.
245. Niijima M, Yamaguchi T, Ishihara T, Hara T, Kato K, Kondo F, et al. Immunohistochemical analysis and in situ hybridization of cyclooxygenase-2 expression in intraductal papillary-mucinous tumors of the pancreas. Cancer. 2002;94(5):1565-73.
246. Tanaka M, Shibahara J, Fukushima N, Shinozaki A, Umeda M, Ishikawa S, et al. Claudin-18 is an early-stage marker of pancreatic carcinogenesis. Journal of Histochemistry and Cytochemistry. 2011;59(10):942-52.
247. Ohuchida K, Mizumoto K, Ishikawa N, Fujii K, Konomi H, Nagai E, et al. The role of S100A6 in pancreatic cancer development and its clinical implication as a diagnostic marker and therapeutic target. Clinical Cancer Research. 2005;11(21):7785-93.
248. Rosty C, Ueki T, Argani P, Jansen M, Yeo CJ, Cameron JL, et al. Overexpression of S100A4 in pancreatic ductal adenocarcinomas is associated with poor differentiation and DNA hypomethylation. American Journal of Pathology. 2002;160(1):45-50.
249. Argani P, Rosty C, Reiter RE, Wilentz RE, Murugesan SR, Leach SD, et al. Discovery of new markers of cancer through serial analysis of gene expression: prostate stem cell antigen is overexpressed in pancreatic adenocarcinoma. Cancer Research. 2001;61(11):4320-4.
250. Yagn WB, Cai F, Cheng CT, Cao G, Qin ZY. Expression of prostate stem cell antigen and claudin-4 in human pancreatic carcinoma. [Chinese]. Acta Academiae Medicinae Sinicae. 2008;30(6):728-31.
251. Strickland LA, Ross J, Williams S, Ross S, Romero M, Spencer S, et al. Preclinical evaluation of carcinoembryonic cell adhesion molecule (CEACAM) 6 as potential therapy target for pancreatic adenocarcinoma. Journal of Pathology. 2009;218(3):380-90.
252. Swartz MJ, Batra SK, Varshney GC, Hollingsworth MA, Yeo CJ, Cameron JL, et al. MUC4 expression increases progressively in pancreatic intraepithelial neoplasia. American journal of clinical pathology. 2002;117(5):791-6.
253. Ligato S, Zhao H, Mandich D, Cartun RW. KOC (K homology domain containing protein overexpressed in cancer) and S100A4-protein immunoreactivity improves the diagnostic sensitivity of biliary brushing cytology for diagnosing pancreaticobiliary malignancies. Diagnostic Cytopathology. 2008;36(8):561-7.
254. Baruch AC, Wang H, Staerkel GA, Evans DB, Hwang RF, Krishnamurthy S. Immunocytochemical study of the expression of mesothelin in fine-needle aspiration biopsy specimens of pancreatic adenocarcinoma. Diagnostic Cytopathology. 2007;35(3):143-7.
255. McCarthy DM, Maitra A, Argani P, Rader AE, Faigel DO, Van Heek NT, et al. Novel markers of pancreatic adenocarcinoma in fine-needle aspiration: mesothelin and prostate stem cell antigen labeling increases accuracy in cytologically borderline cases. Appl Immunohistochem Mol Morphol. 2003;11(3):238-43.
256. Stewart CJR, Burke GM. Value of p53 immunostaining in pancreatico-biliary brush cytology specimens. Diagnostic Cytopathology. 2000;23(5):308-13.

257. Ishimaru S, Itoh M, Hanada K, Tsuchida A, Iwao T, Kajiyama G. Immunocytochemical detection of p53 protein from pancreatic duct brushings in patients with pancreatic carcinoma. Cancer. 1996;77(11):2233-9.
258. Li ZS, Liu F, Xu GM, Sun ZX, Zhou GX, Man XH. Value of the p53 protein for diagnosing cancer in pancreatic cells obtained by endoscopic pancreatic duct brushing. Chinese Journal of Digestive Diseases. 2002;3(3):107-10.
259. Villanacci V, Cestari R, Giulini S, Cengia P, Missale G, Berenzi A, et al. Immunocytochemical assessment of p53 protein to detect malignancy in increased cell-yield brush cytology from the biliopancreatic tree. Digestive diseases and sciences. 2009;54(4):789-92.
260. Gatsonis C, Paliwal P. Meta-analysis of diagnostic and screening test accuracy evaluations: methodologic primer. AJR Am J Roentgenol. 2006;187(2):271-81.
261. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol. 2003;3:25.
262. Sun CY, Wang BL, Hu CQ, Peng RY, Gao YB, Gu QY, et al. Expression of the bcl-2 gene and its significance in human pancreatic carcinoma. Hepatobiliary and Pancreatic Diseases International. 2002;1(2):306-8.
263. Wente MN, Gaida MM, Mayer C, Michalski CW, Haag N, Giese T, et al. Expression and potential function of the CXC chemokine CXCL16 in pancreatic ductal adenocarcinoma. International Journal of Oncology. 2008;33(2):297-308.
264. Boltze C, Schneider-Stock R, Aust G, Mawrin C, Dralle H, Roessner A, et al. CD97, CD95 and Fas-L clearly discriminate between chronic pancreatitis and pancreatic ductal adenocarcinoma in perioperative evaluation of cryocut sections. Pathol Int. 2002;52(2):83-8.
265. Abe N, Watanabe T, Masaki T, Mori T, Sugiyama M, Uchimura H, et al. Pancreatic duct cell carcinomas express high levels of high mobility group I(Y) proteins. Cancer Research. 2000;60(12):3117-22.
266. Shimoyama S, Gansauge F, Gansauge S, Widmaier U, Oohara T, Beger HG. Overexpression of intercellular adhesion molecule-1 (ICAM-1) in pancreatic adenocarcinoma in comparison with normal pancreas. Pancreas. 1997;14(2):181-6.
267. Xie MJ, Motoo Y, Iovanna JL, Su SB, Ohtsubo K, Matsubara F, et al. Overexpression of pancreatitis-associated protein (PAP) in human pancreatic ductal adenocarcinoma. Dig Dis Sci. 2003;48(3):459-64.
268. Akashi T, Oimomi H, Nishiyama K, Nakashima M, Arita Y, Sumii T, et al. Expression and diagnostic evaluation of the human tumor-associated antigen RCAS1 in pancreatic cancer. Pancreas. 2003;26(1):49-55.
269. Koshiba T, Hosotani R, Miyamoto Y, Wada M, Lee JU, Fujimoto K, et al. Immunohistochemical analysis of cyclooxygenase-2 expression in pancreatic tumors. Int J Pancreatol. 1999;26(2):69-76.
270. Oh YL, Song SY, Ahn G. Expression of maspin in pancreatic neoplasms: application of maspin immunohistochemistry to the differential diagnosis. Appl Immunohistochem Mol Morphol. 2002;10(1):62-6.
271. Zhao H, Mandich D, Cartun RW, Ligato S. Expression of K homology domain containing protein overexpressed in cancer in pancreatic FNA for diagnosing adenocarcinoma of pancreas. Diagn Cytopathol. 2007;35(11):700-4.
272. La Rosa S, Franzi F, Marchet S, Finzi G, Clerici M, Vigetti D, et al. The monoclonal anti-BCL10 antibody (clone 331.1) is a sensitive and specific marker of pancreatic acinar cell carcinoma and pancreatic metaplasia. Virchows Arch. 2009;454(2):133-42.
273. Hosoda W, Sasaki E, Murakami Y, Yamao K, Shimizu Y, Yatabe Y. BCL10 as a useful marker for pancreatic acinar cell carcinoma, especially using endoscopic ultrasound cytology specimens. Pathol Int. 2013;63(3):176-82.
274. Hosoda W, Takagi T, Mizuno N, Shimizu Y, Sano T, Yamao K, et al. Diagnostic approach to pancreatic tumors with the specimens of endoscopic ultrasound-guided fine needle aspiration. Pathol Int. 2010;60(5):358-64.
275. Wang H, Song X, Logsdon C, Zhou G, Evans DB, Abbruzzese JL, et al. Proteasome-mediated degradation and functions of hematopoietic progenitor kinase 1 in pancreatic cancer. Cancer Res. 2009;69(3):1063-70.
276. Chu PG, Schwarz RE, Lau SK, Yen Y, Weiss LM. Immunohistochemical staining in the diagnosis of pancreatobiliary and ampulla of Vater adenocarcinoma: application of CDX2, CK17, MUC1, and MUC2. Am J Surg Pathol. 2005;29(3):359-67.
277. Haglund C, Lindgren J, Roberts PJ, Nordling S. Gastrointestinal cancer-associated antigen CA 19-9 in histological specimens of pancreatic tumours and pancreatitis. Br J Cancer. 1986;53(2):189-95.
278. Satomura Y, Sawabu N, Takemori Y, Ohta H, Watanabe H, Okai T, et al. Expression of various sialylated carbohydrate antigens in malignant and nonmalignant pancreatic tissues. Pancreas. 1991;6(4):448-58.

279. Erhuma M, Kobel M, Mustafa T, Wulfanger J, Dralle H, Hoang-Vu C, et al. Expression of neutral endopeptidase (NEP/CD10) on pancreatic tumor cell lines, pancreatitis and pancreatic tumor tissues. Int J Cancer. 2007;120(11):2393-400.
280. Dodge JE, Covens AL, Lacchetti C, Elit LM, Le T, Devries-Aboud M, et al. Preoperative identification of a suspicious adnexal mass: a systematic review and meta-analysis. Gynecol Oncol. 2012;126(1):157-66.
281. Brazzelli M, Sandercock PA, Chappell FM, Celani MG, Righetti E, Arestis N, et al. Magnetic resonance imaging versus computed tomography for detection of acute vascular lesions in patients presenting with stroke symptoms. Cochrane Database Syst Rev. 2009(4):CD007424.
282. Edwards J, Mukherjee R, Munro AF, Wells AC, Almushatat A, Bartlett JM. HER2 and COX2 expression in human prostate cancer. Eur J Cancer. 2004;40(1):50-5.
283. Kirkegaard T, Edwards J, Tovey S, McGlynn LM, Krishna SN, Mukherjee R, et al. Observer variation in immunohistochemical analysis of protein expression, time for a change? Histopathology. 2006;48(7):787-94.
284. Jamieson NB, Carter CR, McKay CJ, Oien KA. Tissue biomarkers for prognosis in pancreatic ductal adenocarcinoma: a systematic review and meta-analysis. Clin Cancer Res. 2011;17(10):3316-31.
285. Smith RA, Tang J, Tudur-Smith C, Neoptolemos JP, Ghaneh P. Meta-analysis of immunohistochemical prognostic markers in resected pancreatic cancer. Br J Cancer. 2011;104(9):1440-51.
286. Woo SM, Ryu JK, Lee SH, Yoo JW, Park JK, Kim YT, et al. Recurrence and prognostic factors of ampullary carcinoma after radical resection: comparison with distal extrahepatic cholangiocarcinoma. Ann Surg Oncol. 2007;14(11):3195-201.
287. Dabizzi E, Assef MS, Raimondo M. Diagnostic management of pancreatic cancer. Cancers. 2011;3(1):494-509.
288. Miura F, Takada T, Amano H, Yoshida M, Furui S, Takeshita K. Diagnosis of pancreatic cancer. HPB (Oxford). 2006;8(5):337-42.
289. Van Beers BE. Diagnosis of cholangiocarcinoma. HPB. 2008;10(2):87-93.
290. Ali A, Ul-Haq Z, Mohamed M, MacKay DF, Duthie F, Oien K. Abstract 1142: Systematic review and meta-analysis of immunohistochemical diagnostic markers for pancreatic ductal adenocarcinoma. Cancer Research. 2013;73(8 Supplement):1142.
291. Argani P, Shaukat A, Kaushal M, Wilentz RE, Su GH, Sohn TA, et al. Differing rates of loss of DPC4 expression and of p53 overexpression among carcinomas of the proximal and distal bile ducts. Cancer. 2001;91(7):1332-41.
292. Swierczynski SL, Maitra A, Abraham SC, Iacobuzio-Donahue CA, Ashfaq R, Cameron JL, et al. Analysis of novel tumor markers in pancreatic and biliary carcinomas using tissue microarrays. Hum Pathol. 2004;35(3):357-66.
293. Mueller-Pillasch F, Lacher U, Wallrapp C, Micha A, Zimmerhackl F, Hameister H, et al. Cloning of a gene highly overexpressed in cancer coding for a novel KH-domain containing protein. Oncogene. 1997;14(22):2729-33.
294. Liao B, Hu Y, Herrick DJ, Brewer G. The RNA-binding protein IMP-3 is a translational activator of insulin-like growth factor II leader-3 mRNA during proliferation of human K562 leukemia cells. J Biol Chem. 2005;280(18):18517-24.
295. Simon R, Bourne PA, Yang Q, Spaulding BO, di Sant'Agnese PA, Wang HL, et al. Extrapulmonary small cell carcinomas express K homology domain containing protein overexpressed in cancer, but carcinoid tumors do not. Hum Pathol. 2007;38(8):1178-83.
296. Walter O, Prasad M, Lu S, Quinlan RM, Edmiston KL, Khan A. IMP3 is a novel biomarker for triple negative invasive mammary carcinoma associated with a more aggressive phenotype. Hum Pathol. 2009;40(11):1528-33.
297. Zheng W, Yi X, Fadare O, Liang SX, Martel M, Schwartz PE, et al. The oncofetal protein IMP3: a novel biomarker for endometrial serous carcinoma. Am J Surg Pathol. 2008;32(2):304-15.
298. Li D, Yan D, Tang H, Zhou C, Fan J, Li S, et al. IMP3 Is a Novel Prognostic Marker that Correlates with Colon Cancer Progression and Pathogenesis. Annals of Surgical Oncology. 2009;16(12):3499-506.
299. Becker T, Gerke V, Kube E, Weber K. S100P, a novel Ca(2+)-binding protein from human placenta. cDNA cloning, recombinant protein expression and Ca2+ binding properties. Eur J Biochem. 1992;207(2):541-7.
300. Arumugam T, Simeone DM, Van Golen K, Logsdon CD. S100P promotes pancreatic cancer growth, survival, and invasion. Clin Cancer Res. 2005;11(15):5356-64.
301. Parkkila S, Pan P-w, Ward A, Gibadulinova A, Oveckova I, Pastorekova S, et al. The calcium-binding protein S100P in normal and malignant human tissues. BMC Clinical Pathology. 2008;8(1):2.
302. Hassan R, Ho M. Mesothelin targeted cancer immunotherapy. Eur J Cancer. 2008;44(1):46-53.
303. Tang Z, Qian M, Ho M. The role of mesothelin in tumor progression and targeted therapy. Anticancer Agents Med Chem. 2013;13(2):276-80.

304. Ordonez NG. Value of mesothelin immunostaining in the diagnosis of mesothelioma. Mod Pathol. 2003;16(3):192-7.
305. Gendler SJ. MUC1, the renaissance molecule. J Mammary Gland Biol Neoplasia. 2001;6(3):339-53.
306. Besmer DM, Curry JM, Roy LD, Tinder TL, Sahraei M, Schettini J, et al. Pancreatic ductal adenocarcinoma mice lacking mucin 1 have a profound defect in tumor growth and metastasis. Cancer Res. 2011;71(13):4432-42.
307. Kitamura H, Yonezawa S, Tanaka S, Kim YS, Sato E. Expression of mucin carbohydrates and core proteins in carcinomas of the ampulla of Vater: their relationship to prognosis. Jpn J Cancer Res. 1996;87(6):631-40.
308. Utsunomiya T, Yonezawa S, Sakamoto H, Kitamura H, Hokita S, Aiko T, et al. Expression of MUC1 and MUC2 mucins in gastric carcinomas: its relationship with the prognosis of the patients. Clin Cancer Res. 1998;4(11):2605-14.
309. Sagara M, Yonezawa S, Nagata K, Tezuka Y, Natsugoe S, Xing PX, et al. Expression of mucin 1 (MUC1) in esophageal squamous-cell carcinoma: its relationship with prognosis. Int J Cancer. 1999;84(3):251-7.
310. Matsukita S, Nomoto M, Kitajima S, Tanaka S, Goto M, Irimura T, et al. Expression of mucins (MUC1, MUC2, MUC5AC and MUC6) in mucinous carcinoma of the breast: comparison with invasive ductal carcinoma. Histopathology. 2003;42(1):26-36.
311. Bodenstine TM, Seftor RE, Khalkhali-Ellis Z, Seftor EA, Pemberton PA, Hendrix MJ. Maspin: molecular mechanisms and therapeutic implications. Cancer Metastasis Rev. 2012;31(3-4):529-51.
312. Maass N, Teffner M, Rosel F, Pawaresch R, Jonat W, Nagasaki K, et al. Decline in the expression of the serine proteinase inhibitor maspin is associated with tumour progression in ductal carcinomas of the breast. J Pathol. 2001;195(3):321-6.
313. Berardi R, Morgese F, Onofri A, Mazzanti P, Pistelli M, Ballatore Z, et al. Role of maspin in cancer. Clin Transl Med. 2013;2(1):8.
314. Boltze C. Loss of maspin is a helpful prognosticator in colorectal cancer: a tissue microarray analysis. Pathology - Research and Practice. 2005;200(11–12):783-90.
315. Machtens S, Serth J, Bokemeyer C, Bathke W, Minssen A, Kollmannsberger C, et al. Expression of the p53 and Maspin protein in primary prostate cancer: correlation with clinical features. Int J Cancer. 2001;95(5):337-42.
316. Lee MJ, Suh CH, Li ZH. Clinicopathological significance of maspin expression in breast cancer. J Korean Med Sci. 2006;21(2):309-14.
317. Takanami I, Abiko T, Koizumi S. Expression of maspin in non-small-cell lung cancer: correlation with clinical features. Clinical lung cancer. 2008;9(6):361-6.
318. Mazumdar M, Glassman JR. Categorizing a prognostic variable: review of methods, code for easy implementation and applications to decision-making about cancer treatments. Stat Med. 2000;19(1):113-32.
319. Weiss HL, Niwas S, Grizzle WE, Piyathilake C. Receiver operating characteristic (ROC) to determine cut-off points of biomarkers in lung cancer patients. Dis Markers. 2003;19(6):273-8.
320. Borlot VF, Biasoli I, Schaffel R, Azambuja D, Milito C, Luiz RR, et al. Evaluation of intra- and interobserver agreement and its clinical significance for scoring bcl-2 immunohistochemical expression in diffuse large B-cell lymphoma. Pathol Int. 2008;58(9):596-600.
321. Cohen DA, Dabbs DJ, Cooper KL, Amin M, Jones TE, Jones MW, et al. Interobserver agreement among pathologists for semiquantitative hormone receptor scoring in breast carcinoma. American journal of clinical pathology. 2012;138(6):796-802.
322. Hameed O, Adams AL, Baker AC, Balmer NE, Bell WC, Burford HN, et al. Using a higher cutoff for the percentage of HER2+ cells decreases interobserver variability in the interpretation of HER2 immunohistochemical analysis. American journal of clinical pathology. 2008;130(3):425-7.
323. Kim MJ, Shin HC, Shin KC, Ro JY. Best immunohistochemical panel in distinguishing adenocarcinoma from squamous cell carcinoma of lung: tissue microarray assay in resected lung cancer specimens. Ann Diagn Pathol. 2013;17(1):85-90.
324. Galgano MT, Castle PE, Atkins KA, Brix WK, Nassau SR, Stoler MH. Using biomarkers as objective standards in the diagnosis of cervical biopsies. Am J Surg Pathol. 2010;34(8):1077-87.
325. Cheuk W, Wong KO, Wong CS, Dinkel JE, Ben-Dor D, Chan JK. Immunostaining for human herpesvirus 8 latent nuclear antigen-1 helps distinguish Kaposi sarcoma from its mimickers. American journal of clinical pathology. 2004;121(3):335-42.
326. Lin F, Shi J, Liu H, Hull ME, Dupree W, Prichard JW, et al. Diagnostic utility of S100P and von Hippel-Lindau gene product (pVHL) in pancreatic adenocarcinoma-with implication of their roles in early tumorigenesis. The American journal of surgical pathology. 2008;32(1):78-91.

327. Maass N, Hojo T, Ueding M, Luttges J, Kloppel G, Jonat W, et al. Expression of the tumor suppressor gene Maspin in human pancreatic cancers. Clinical cancer research : an official journal of the American Association for Cancer Research. 2001;7(4):812-7.
328. Strickland LA, Ross J, Williams S, Ross S, Romero M, Spencer S, et al. Preclinical evaluation of carcinoembryonic cell adhesion molecule (CEACAM) 6 as potential therapy target for pancreatic adenocarcinoma. J Pathol. 2009;218(3):380-90.
329. Yamaguchi U, Hasegawa T, Sakurai S, Sakuma Y, Takazawa Y, Hishima T, et al. Interobserver variability in histologic recognition, interpretation of KIT immunostaining, and determining MIB-1 labeling indices in gastrointestinal stromal tumors and other spindle cell tumors of the gastrointestinal tract. Appl Immunohistochem Mol Morphol. 2006;14(1):46-51.
330. Wentzensen N, Schwartz L, Zuna RE, Smith K, Mathews C, Gold MA, et al. Performance of p16/Ki-67 immunostaining to detect cervical cancer precursors in a colposcopy referral population. Clin Cancer Res. 2012;18(15):4154-62.
331. Wolff AC, Hammond ME, Schwartz JN, Hagerty KL, Allred DC, Cote RJ, et al. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2007;25(1):118-45.
332. Polley MY, Leung SC, McShane LM, Gao D, Hugh JC, Mastropasqua MG, et al. An international Ki67 reproducibility study. J Natl Cancer Inst. 2013;105(24):1897-906.
333. Abe N, Watanabe T, Masaki T, Mori T, Sugiyama M, Uchimura H, et al. Pancreatic duct cell carcinomas express high levels of high mobility group I(Y) proteins. Cancer research. 2000;60(12):3117-22.
334. Crocetti E, Caldarella A, Ferretti S, Ardanaz E, Arveux P, Bara S, et al. Consistency and inconsistency in testing biomarkers in breast cancer. A GRELL study in cut-off variability in the Romance language countries. Breast. 2013;22(4):476-81.
335. Mallofre C, Castillo M, Morente V, Sole M. Immunohistochemical expression of CK20, p53, and Ki-67 as objective markers of urothelial dysplasia. Mod Pathol. 2003;16(3):187-91.
336. Perkins NJ, Schisterman EF. The inconsistency of "optimal" cutpoints obtained using two criteria based on the receiver operating characteristic curve. Am J Epidemiol. 2006;163(7):670-5.
337. Gosho M, Nagashima K, Sato Y. Study designs and statistical analyses for biomarker research. Sensors (Basel). 2012;12(7):8966-86.
338. Tzankov A, Zlobec I, Went P, Robl H, Hoeller S, Dirnhofer S. Prognostic immunophenotypic biomarker studies in diffuse large B cell lymphoma with special emphasis on rational determination of cut-off scores. Leuk Lymphoma. 2010;51(2):199-212.
339. Simon R. Clinical trial designs for evaluating the medical utility of prognostic and predictive biomarkers in oncology. (1741-0541 (Print)).
340. Fisher G, Yang ZH, Kudahetti S, Moller H, Scardino P, Cuzick J, et al. Prognostic value of Ki-67 for prostate cancer death in a conservatively managed cohort. Br J Cancer. 2013;108(2):271-7.
341. Zlobec I, Vuong T, Compton CC. The predictive value of apoptosis protease-activating factor 1 in rectal tumors treated with preoperative, high-dose-rate brachytherapy. Cancer. 2006;106(2):284-6.
342. Hirsch FR, Dziadziuszko R, Thatcher N, Mann H, Watkins C, Parums DV, et al. Epidermal growth factor receptor immunohistochemistry: comparison of antibodies and cutoff points to predict benefit from gefitinib in a phase 3 placebo-controlled study in advanced nonsmall-cell lung cancer. Cancer. 2008;112(5):1114-21.
343. Kobayashi T, Iwaya K, Moriya T, Yamasaki T, Tsuda H, Yamamoto J, et al. A simple immunohistochemical panel comprising 2 conventional markers, Ki67 and p53, is a powerful tool for predicting patient outcome in luminal-type breast cancer. BMC Clinical Pathology. 2013;13(1):5.
344. Kerkhof M, van Dekken H, Steyerberg EW, Meijer GA, Mulder AH, de Bruine A, et al. Grading of dysplasia in Barrett's oesophagus: substantial interobserver variation between general and gastrointestinal pathologists. Histopathology. 2007;50(7):920-7.
345. Lorinc E, Jakobsson B, Landberg G, Veress B. Ki67 and p53 immunohistochemistry reduces interobserver variation in assessment of Barrett's oesophagus. Histopathology. 2005;46(6):642-8.
346. Turner JK, Williams GT, Morgan M, Wright M, Dolwani S. Interobserver agreement in the reporting of colorectal polyp pathology among bowel cancer screening pathologists in Wales. Histopathology. 2013;62(6):916-24.
347. von Wasielewski R, Mengel M, Wiese B, Rudiger T, Muller-Hermelink HK, Kreipe H. Tissue array technology for testing interlaboratory and interobserver reproducibility of immunohistochemical estrogen receptor analysis in a large multicenter trial. American journal of clinical pathology. 2002;118(5):675-82.
348. Cross SS, Betmouni S, Burton JL, Dube AK, Feeley KM, Holbrook MR, et al. What levels of agreement can be expected between histopathologists assigning cases to discrete nominal categories? A study of the diagnosis of hyperplastic and adenomatous colorectal polyps. Mod Pathol. 2000;13(9):941-4.

349. Jaraj SJ, Camparo P, Boyle H, Germain F, Nilsson B, Petersson F, et al. Intra- and interobserver reproducibility of interpretation of immunohistochemical stains of prostate cancer. Virchows Arch. 2009;455(4):375-81.
350. Coras R, de Boer OJ, Armstrong D, Becker A, Jacques TS, Miyata H, et al. Good interobserver and intraobserver agreement in the evaluation of the new ILAE classification of focal cortical dysplasias. Epilepsia. 2012;53(8):1341-8.
351. Cascinu S, Staccioli MP, Gasparini G, Giordani P, Catalano V, Ghiselli R, et al. Expression of vascular endothelial growth factor can predict event-free survival in stage II colon cancer. Clin Cancer Res. 2000;6(7):2803-7.
352. Hoang MP, Sahin AA, Ordonez NG, Sneige N. HER-2/neu gene amplification compared with HER-2/neu protein overexpression and interobserver reproducibility in invasive breast carcinoma. American journal of clinical pathology. 2000;113(6):852-9.
353. Ogino J, Asanuma H, Hatanaka Y, Matsuno Y, Gotoda H, Muraoka S, et al. Validity and reproducibility of Ki-67 assessment in gastrointestinal stromal tumors and leiomyosarcomas. Pathol Int. 2013;63(2):102-7.
354. Thomson TA, Hayes MM, Spinelli JJ, Hilland E, Sawrenko C, Phillips D, et al. HER-2/neu in breast cancer: interobserver variability and performance of immunohistochemistry with 4 antibodies compared with fluorescent in situ hybridization. Mod Pathol. 2001;14(11):1079-86.
355. Park D, Kåresen R, Noren T, Sauer T. Ki-67 expression in primary breast carcinomas and their axillary lymph node metastases: clinical implications. Virchows Archiv. 2007;451(1):11-8.
356. Aumann K, Frey AV, May AM, Hauschke D, Kreutz C, Marx JP, et al. Subcellular mislocalization of the transcription factor NF-E2 in erythroid cells discriminates prefibrotic primary myelofibrosis from essential thrombocythemia. Blood. 2013;122(1):93-9.
357. Trivedi A, Cartun RW, Ligato S. Role of lymphovascular invasion and immunohistochemical expression of IMP3 in the risk stratification of superficially invasive pT1 esophageal adenocarcinoma. Pathol Res Pract. 2014.
358. Jo VY, Cibas ES, Pinkus GS. Claudin-4 immunohistochemistry is highly effective in distinguishing adenocarcinoma from malignant mesothelioma in effusion cytology. Cancer Cytopathol. 2014;122(4):299-306.
359. Betta PG, Andrion A, Donna A, Mollo F, Scelsi M, Zai G, et al. Malignant mesothelioma of the pleura. The reproducibility of the immunohistological diagnosis. Pathol Res Pract. 1997;193(11-12):759-65.
360. Atkinson R, Mollerup J, Laenkholm AV, Verardo M, Hawes D, Commins D, et al. Effects of the change in cutoff values for human epidermal growth factor receptor 2 status by immunohistochemistry and fluorescence in situ hybridization: a study comparing conventional brightfield microscopy, image analysis-assisted microscopy, and interobserver variation. Arch Pathol Lab Med. 2011;135(8):1010-6.
361. Fowler LJ, Lachar WA. Application of immunohistochemistry to cytology. Arch Pathol Lab Med. 2008;132(3):373-83.
362. Qin S-y, Zhou Y, Li P, Jiang H-x. Diagnostic Efficacy of Cell Block Immunohistochemistry, Smear Cytology, and Liquid-Based Cytology in Endoscopic Ultrasound-Guided Fine-Needle Aspiration of Pancreatic Lesions: A Single-Institution Experience. PLoS ONE. 2014;9(9):e108762.
363. Paulo JA, Lee LS, Banks PA, Steen H, Conwell DL. Proteomic analysis of formalin-fixed paraffin-embedded pancreatic tissue using liquid chromatography tandem mass spectrometry. Pancreas. 2012;41(2):175-85.
364. Dim DC, Jiang F, Qiu Q, Li T, Darwin P, Rodgers WH, et al. The usefulness of S100P, mesothelin, fascin, prostate stem cell antigen, and 14-3-3 sigma in diagnosing pancreatic adenocarcinoma in cytological specimens obtained by endoscopic ultrasound guided fine-needle aspiration. Diagn Cytopathol. 2011.
365. Brooks JD. Translational genomics: the challenge of developing cancer biomarkers. Genome Res. 2012;22(2):183-7.
366. Parikh NI, Vasan RS. Assessing the clinical utility of biomarkers in medicine. Biomark Med. 2007;1(3):419-36.
367. Gong Y, Symmans WF, Krishnamurthy S, Patel S, Sneige N. Optimal fixation conditions for immunocytochemical analysis of estrogen receptor in cytologic specimens of breast carcinoma. Cancer. 2004;102(1):34-40.
368. Fetsch PA, Abati A. The effects of antibody clone and pretreatment method on the results of HER2 immunostaining in cytologic samples of metastatic breast cancer: A query and a review of the literature. Diagn Cytopathol. 2007;35(6):319-28.
369. Shi SR, Cote RJ, Taylor CR. Antigen retrieval techniques: current perspectives. J Histochem Cytochem. 2001;49(8):931-7.
370. Mayeux R. Biomarkers: potential uses and limitations. NeuroRx. 2004;1(2):182-8.

371. Chatzipantelis P, Salla C, Konstantinou P, Karoumpalis I, Sakellariou S, Doumani I. Endoscopic ultrasound-guided fine-needle aspiration cytology of pancreatic neuroendocrine tumors: a study of 48 cases. Cancer. 2008;114(4):255-62.
372. Skov BG, Baandrup U, Jakobsen GK, Kiss K, Krasnik M, Rossen K, et al. Cytopathologic diagnoses of fine-needle aspirations from endoscopic ultrasound of the mediastinum. Cancer Cytopathology. 2007;111(4):234-41.
373. Doornewaard H, van der Schouw YT, van der Graaf Y, Bos AB, van den Tweel JG. Observer variation in cytologic grading for cervical dysplasia of Papanicolaou smears with the PAPNET testing system. Cancer Cytopathology. 1999;87(4):178-83.
374. Darragh TM, Tokugawa D, Castle PE, Follansbee S, Borgonovo S, LaMere BJ, et al. Interrater agreement of anal cytology. Cancer Cytopathol. 2013;121(2):72-8.
375. Khan SA, Lankes HA, Patil DB, Bryk M, Hou N, Ivancic D, et al. Ductal lavage is an inefficient method of biomarker measurement in high-risk women. Cancer Prev Res (Phila). 2009;2(3):265-73.
376. Salomao M, Remotti H, Allendorf JD, Poneros JM, Sethi A, Gonda TA, et al. Fine-needle aspirations of pancreatic serous cystadenomas: improving diagnostic yield with cell blocks and alpha-inhibin immunohistochemistry. Cancer Cytopathol. 2014;122(1):33-9.
377. Hu HY, Liu H, Zhang JW, Hu K, Lin Y. Clinical significance of Smac and Ki-67 expression in pancreatic cancer. Hepatogastroenterology. 2012;59(120):2640-3.
378. Shi C, Merchant N, Newsome G, Goldenberg DM, Gold DV. Differentiation of pancreatic ductal adenocarcinoma from chronic pancreatitis by PAM4 immunohistochemistry. Arch Pathol Lab Med. 2014;138(2):220-8.
379. Xiao W, Hong H, Awadallah A, Yu S, Zhou L, Xin W. CRABP-II is a highly sensitive and specific diagnostic molecular marker for pancreatic ductal adenocarcinoma in distinguishing from benign pancreatic conditions. Hum Pathol. 2014;45(6):1177-83.
380. Horn A, Chakraborty S, Dey P, Haridas D, Souchek J, Batra SK, et al. Immunocytochemistry for MUC4 and MUC16 is a useful adjunct in the diagnosis of pancreatic adenocarcinoma on fine-needle aspiration cytology. Arch Pathol Lab Med. 2013;137(4):546-51.
381. Pepe MS, Etzioni R, Feng Z, Potter JD, Thompson ML, Thornquist M, et al. Phases of biomarker development for early detection of cancer. J Natl Cancer Inst. 2001;93(14):1054-61.
382. Heckman-Stoddard BM. Oncology biomarkers: discovery, validation, and clinical use. Semin Oncol Nurs. 2012;28(2):93-8.