Modifying factors influencing genetic risk of pancreatic cancer

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor of Medicine by Jennifer Dorothy Law

October 2021

Acknowledgements

I would like to thank my supervisors Professor William Greenhalf, Professor Eithne Costello and Mr Christopher Halloran for their help and support during the course of my MD research.

I would also like to thank the pancreatic cancer research group at The University of Liverpool for their support. In particular Dr Anthony Evans, Dr Lawrence Barrera and Dr Jessica Hale for their assistance in scientific techniques, data analysis and reviewing. I would also like to thank Dr Jennifer Mitchell and Professor Ross Sibson for all of their help with genetics-based scientific techniques.

I am grateful for the support from The European Union who funded my research and thesis through Immunovia (Lund, Sweden). I also wish to thank the multiple patients who continue to be involved in both EUROPAC and the IMMPACT studies for their keen support, time and patience.

Lastly, but most importantly I wish to thank Scott and Robert for their loving support, patience and kindness.

Modifying factors influencing genetic risk of pancreatic cancer – Jennifer D Law

Abstract

Background: The European Registry of Hereditary Pancreatitis and Familial Pancreatic Cancer (EUROPAC) continues to recruit families with an increased risk of pancreatic cancer (PDAC) for both primary and secondary screening. Prospective risk of PDAC in families with multiple cases of the disease is sufficiently high to mean that they are suitable for screening for cancer, but most families have no known causative mutation. Some EUROPAC families have mutations in genes which cause cancer syndromes known to increase the risk of PDAC above the population baseline such as Hereditary Breast and Ovarian Cancer syndrome (*BRCA1, BRCA2*), Familial Atypical Mole Melanoma (*p16, CDKN2A*), Peutz Jeghers syndrome (*STK11*) and Lynch syndrome (HNPCC – *MLH1, MSH2, MSH6, PMS2, EPCAM*). However, the risk for mutation carriers taken as a whole is much less than would be predicted for prospective risk in families with multiple cases. This thesis addresses whether yield from screening can be improved by risk stratification and targeting screening to highest risk individuals. This includes addressing the hypothesis that a particular predisposing mutation may have different risks for cancer depending on the particular family history (context specific risk).

Methods: Populations registered with EUROPAC were evaluated to allow description and comparison of individuals who attended for secondary screening with those who did not. Subsequently the population with confirmed pathogenic *BRCA2* mutations was analysed to allow further description and to assess links between the familial risk of PDAC and other pheno-genetic characteristics. As part of the thesis a method was developed and optimised to allow full sequencing of the *BRCA2* gene and surrounding bases to assess whether single-nucleotide variant (SNV) phenotypes may be linked to different cancer risks between *BRCA2* families. This method included optimisation of cell plug formation, pulsed-field gel electrophoresis, quantitative PCR and Nanopore sequencing.

Results: Participation in screening was found to be influenced by many factors including level of perceived risk. Lack of a specific mutation was a major element in reducing uptake, patients with known mutations perceived their risk as greater. However, the actual risk in *BRCA2* carriers varied according to family history with individuals with fewer previous cases of PDAC in their family having risk that would probably not justify screening. From this we developed the hypothesis that whilst specific single mutations are necessary for an elevated risk of PDAC in some families, there are other factors that are necessary for these mutations to be penetrant. These other factors may be combinations of multiple polymorphisms in other genes, high risk combinations being ubiquitous in some families and absent in others. A method for carrying out long range haplotyping analysis was developed which can be taken forward to address this hypothesis.

Contents

Tables

Figures

Abbreviations

BMI	Body mass index
CA19-9	Carbohydrate antigen 19-9
CAPS Consortium	The International Cancer of the Pancreas Screening
cDNA	Complementary Deoxyribonucleic Acid
CIMBA	Consortium of Investigators of Modifiers of BRCA1/2
CRISPR	Clustered regularly interspaced short palindromic repeats
CT	Computed tomography
DMEM	Dulbecco Modified Eagle Medium
DSB	Double-strand break
EUROPAC Pancreatic Cancer	The European Registry of Hereditary Pancreatitis and Familial
EUS	Endoscopic ultrasound
FAMMM	Familial atypical multiple mole melanoma
FAP	Familial adenomatous polyposis
FHx	Family History
FOLFIRINOX	5-fluorouracil, folinic acid, irinotecan and oxaliplatin
FPC	Familial pancreatic cancer
gDNA	Genomic Deoxyribonucleic Acid
GWAS	Genome-wide association studies
HBOC	Hereditary Breast and Ovarian Cancer Syndrome
HNPCC	Hereditary non-polyposis Colorectal cancer
HP	Hereditary pancreatitis
HPC	Hereditary pancreatic cancer
IMMPACT	The Clinical validation of a serum protein biomarker signature for the early diagnosis of pancreatic cancer and cystic neoplasms
IPMN	Intraductal papillary mucinous neoplasm

MCN	Mucinous cystic neoplasms
MRI	Magnetic resonance Imaging
NGS	Next generation sequencing
NICE	National Institute for Clinical Excellence
NFW	Nuclease free water
PanIN	Pancreatic intraepithelial neoplasia
PBMC	Peripheral blood mononuclear cells
PC	Pancreatic cancer
PCUK	Pancreatic Cancer UK
PDAC	Pancreatic adenocarcinoma
PFGE	Pulsed field gel electrophoresis
PJS	Peutz Jeghers syndrome
qPCR	Quantitative polymerase chain reaction
RCA	Rolling-circle amplification
SNV	Single nucleotide variant
UK	United Kingdom of Great Britain and Northern Ireland
US	Ultrasound
USA	United States of America

1. Introduction

1.1 Pancreatic cancer

Pancreatic cancer has a reported incidence of between 3 and 12 per 100,000 population per year (1,2). Despite the low incidence pancreatic cancer is the 4th leading cause of cancer-related deaths in the United States of America (USA) with an annual death rate of 41,780 (2–4). In the United Kingdom (UK) the Office for National Statistics data reported 8,650 cancer-related deaths due to malignant neoplasm of the pancreas in 2019, and unlike many other cancers there continues to be an increasing annual trend (5). Projections suggest that pancreatic cancer will become the 2nd leading cause of cancer-related deaths by 2030 (6).

Pancreatic adenocarcinoma (PDAC) makes up the majority of diagnoses of pancreatic cancer (7). Due to lack of early symptoms, late diagnosis at presentation and difficulties with identification of 'high risk' individuals; diagnosis of disease is often late-stage (6). These factors plus an often poor response to chemotoxic agents means 5-year survival is poor (3.2-14.6%) and this has not improved in line with that of other cancers (1,6,8). Disease that is early-stage and is amenable to a complete surgical resection has an improved prognosis (9,10).

1.1.1 Precursor lesions

Three precursors to PDAC are well described; pancreatic intraepithelial neoplasia (PanIN), intraductal papillary mucinous neoplasm (IPMN) and mucinous cystic neoplasms (MCN) (11–13). PanINs are the most common precursor lesions and it has been projected that it takes between 11 and 12 years for a PanIN3 (likened to carcinoma-in-situ) lesion to progress to PDAC (11). All PanINs are non-invasive lesions and the accumulation of genetic changes leads to low-grade lesions (PanIN-1) developing into high-grade (PanIN-3) which may eventually progress to invasive lesions (14). Considerable difficulties exist with radiological detection of PanIN lesions despite them sharing several features concerning for PDAC (15). Lastly,

debate continues about treatment of PanIN lesions found at the frozen section margin during pancreatic resection for PDAC. Distler and colleagues suggest further resection if PanIN-3 is found within the margin however, they feel that intermediate or low grade lesions may be left due to the increased morbidity associated with further resection (14).

IPMNs are well described and there is considerable evidence that the risk of malignancy is significantly higher in main duct (mean frequency 61.6%) compared with branch duct IPMNs (mean frequency 25.5.%) (12,16). These lesions can be difficult to distinguish radiologically from other cystic lesions of the pancreas pre-operatively (14). Generally, it is recommended that all main duct IPMNs and those branch duct IPMNs with 'worrisome features' are resected where possible (16). Surveillance of low-risk lesions is recommended initially at short intervals to establish stability (16).

MCNs are the least common of the precursor lesions and are significantly more common in women (14). Lesions can grow very large and despite only one third becoming invasive it is recommended, due to their presence in young patients, that they are resected regardless (16). Studies assessing resected MCNs found that between 13 and 17.5% contain invasive malignancy (13,16). Prognosis for patients undergoing a resection for invasive MCN has been reported favourably, certainly in comparison to PDAC however, this is to be accepted with caution given the limited long-term follow-up data and the low numbers included in the study (13). Surveillance has been suggested in individuals who are more frail however, the aim of this surveillance along with the clinical and patient benefit remains unclear (16).

There is considerable evidence that disease treated at an early stage with a clear resection margin (R0) has a significantly better outcome however, to date preoperative detection of these precursor lesions remains difficult (9,10,17).

1.1.2 Risk factors

Multiple risk factors for PDAC have been assessed and these can be subdivided into modifiable and non-modifiable. Non-modifiable risk factors include; age, genetic factors and family history, sex, blood group, race, diabetes mellitus and chronic pancreatitis.

PDAC risk tends to increase with age. 90% of disease presents in people over 55 years of age and the majority in the 6th and 7th decades of life (17,18). In general those with an inherited increase risk of PDAC are affected earlier than sporadic cases (19). Familial disease and genetic factors will be discussed later but are reported to play a key part in up to 10% of cases of PDAC (20).

Diagnoses of PDAC are more common in men than women. A 2009 literature review assessing the possible causes for this found no link between differences in hormonal exposure in women and pancreatic cancer. The group went on to suggest that this difference may be linked to sex-based environmental exposures rather than an inherent difference in predisposition (21). It remains difficult to assess the validity of this conclusion given the group were assessing alterations in hormonal exposure in women; such as the contraceptive pill, pregnancy and early and late menarche however, none of these exposures mimic the differences in hormonal exposures between men and women. Increasing awareness and use of gender-affirming hormone therapy may allow monitoring of cancer risk in individuals taking exogenous hormones which could provide valuable insights into the effects of these hormones on cancer development (22).

Multiple population-based studies have shown that PDAC is more common in non-O blood groups and the Pancreatic Cancer Cohort Consortium proposed that around 19.5% of pancreatic cancers may be linked to blood group (23,24). It is difficult to assess, but it seems possible that this variation could be caused by genetic factors inherited alongside blood group.

A large population study within the USA found 45% more deaths due to PDAC in African-American individuals compared with Caucasians (25,26). There is the suggestion that some of this discrepancy relates to differences in lifestyle exposures, such as tobacco smoking and body mass index (BMI) however, this does not account for all of this difference (25). Interestingly Pernick and colleagues (2003) found that there were considerable differences in the underlying pathology of pancreatic tumours that may account for some survival variances by race (26). Coupled with this, certain ethnic groups are linked to a higher prevalence of germline mutations which may predispose to an increased risk of PDAC, such as individuals of Ashkenazi Jewish heritage being more likely to have a mutation in either of the *BRCA* genes (27–29).

There remains considerable bidirectional correlation between PDAC and diabetes (30). Individuals with long-standing type 1 or type 2 diabetes have up to a two-fold increase in PDAC (30,31). Whilst several biological mechanisms have been suggested to account for this increase, to date none have been cemented (30). Alongside this, a reduction in glucose tolerance is often seen prior to diagnosis of PDAC and more recently this change has been suggested to be a possible avenue for early detection of pancreatic cancer, with the potential for individuals presenting with new onset diabetes being targeted for early detection of PDAC (32,33). This however, does currently pose some difficulties given there are reported to be over 500,000 people living with undiagnosed type 2 diabetes in the UK alone and an increasing prevalence worldwide (34). Radiological screening of this population would not be feasible due to large numbers of new diagnoses of diabetes with roughly only 10% of this population having pancreato-genic diabetes (type 3c) and 90% of the type 3c diabetes group being due to chronic pancreatitis. Currently multiple research projects are assessing whether blood biomarkers could differentiate between diabetes types and therefore narrow the population requiring radiological screening (35,36).

Chronic pancreatitis carries a well-researched risk of PDAC. Overall there is a reported 13-fold increase in the lifetime risk of PDAC in individuals with chronic

pancreatitis and this is often linked to disease severity, length of time affected and underlying aetiology (37). Controversy remains about whether this population should undergo regular screening for PDAC given that the actual risk for an individual is about 5%. At present due to the large numbers affected and relatively low risk for the majority of individuals with chronic pancreatitis screening is not recommended and is not offered in the UK unless there is a known hereditary component (37).

Modifiable risk factors for PDAC include tobacco smoking, alcohol intake, BMI, diet and *helicobacter pylori (H.pylori)* gastric colonisation.

Tobacco smoking carries a considerable (>70%) increased risk of PDAC and this risk has been shown to persist for at least 10 years after smoking cessation (38,39). Smoking has been shown to increase inflammatory cell infiltration of the pancreas, leading to chronic inflammation and fibrosis which is a common precursor to PDAC, although the connecting mechanism is unknown (40). Smoking also carries considerable correlation with both acute and chronic pancreatitis, and multiple studies have suggested that this may be an accumulative risk with alcohol (39). Tobacco smoking has been shown not only to increase the risk of PDAC in individuals with chronic pancreatitis, but also leads to development of PDAC as much as 20 years earlier (41). The emergence of e-cigarettes has provided new avenues for study within cancer biology but as yet the long-term effects are unknown. Advice from the National Institute for Clinical Excellence (NICE) and Public Health England is that individuals should be encouraged to stop smoking completely, including e-cigarettes, as they may cause unknown harm (42,43). More pragmatically, this possible harm is balanced with the high likelihood that they are less harmful or at least less carcinogenic than cigarette smoking

Alcohol consumption is a considerable and well-known risk factor for both acute and chronic pancreatitis. Links between PDAC and alcohol consumption are less consistent. Multiple published large cohort studies have reported a variation in correlation between alcohol intake and PDAC ranging from none to a strongly

positive correlation (44–47). Rohrmann and colleagues suggest that some of this variation may be due to differences in reporting, sample size differences, confounding variables (such as smoking), differences in risk between alcohol types (e.g. spirits compared with wine) and the difference between measuring current and lifetime alcohol consumption (44).

A large meta-analysis of prospective studies assessing the correlation between BMI and PDAC found a roughly 10% increased risk per 5 units increase in BMI (48). Interestingly the group found that this correlation persisted even in non-smokers with a normal BMI. Whilst this risk did persist when adjusted for lifestyle factors such as smoking and diabetes diagnosis, BMI remains heavily linked with glucose intolerance and therefore it is difficult to fully separate the risk of BMI from diabetes and other metabolic syndromes (48). Interestingly, whilst a high BMI makes PDAC more difficult to diagnose and treat, its link to overall disease survival is debated (49).

Multiple dietary factors have been suggested to increase the risk of PDAC. Studies into dietary factors are usually large cohort studies and therefore are subject to significant confounding. Multiple cohort studies have found a possible link between red meat, processed meat and well-done meat with PDAC (50–52). Interestingly a further large cohort study reported this association was dose-related but only found in men (53). Other studies have suggested a link between sugary drinks, especially those containing fructose, and PDAC (54,55). Schernhammer and colleagues found that this relationship was only present in women with an increased BMI (56). The group postulated that this was related to glucose intolerance in this group regardless of a lack of diagnosis of diabetes mellitus.

Infection with certain *H.pylori* strains has been found in a case-control study to be correlated with PDAC risk specifically in individuals with non-O blood groups with a possibility of a link with gastric acidity (24,57). This result was not echoed in a large cohort study which suggested no link between *H.pylori* gastric colonisation and PDAC (58).

1.1.3 Current diagnosis and treatment

PDAC is often diagnosed at an advanced stage, with 80-85% patients presenting with disease where a curative resection is not possible (59). Early disease usually has no symptoms and is often found incidentally on radiological imaging. In later disease, symptoms often reflect local invasion or metastatic disease (59). Non-specific, 'common' symptoms plus low incidence of disease mean patients often seek medical attention more frequently than normal. A 2014 study found patients had a median of 18 GP visits in the year leading up to a diagnosis of PDAC (60).

Best practice for diagnosis and treatment follows a multi-disciplinary team approach. Complete surgical resection with a pancreatico-duodenectomy, distal or total pancreatectomy continues to be the only curative management for PDAC (17). Debate continues as to whether anastomotic technique, minimally invasive surgery and vascular resection impacts long-term outcomes after a resection with curative intent (61–63). Surgical treatment alone is insufficient, and without further treatment or a combination of treatments most PDACs relapse and eventually lead to death (9).

The current standard of care is for adjuvant chemotherapy with modified FOLFIRINOX (infusion 5-fluorouracil (5-FU), folinic acid, irinotecan, and oxaliplatin) in patients who are fit enough post-operatively (64). Individuals who have a performance score of 2 or more (therefore not as fit) may receive combination gemcitabine and capecitabine (gem-cap) chemotherapy (64). Adjuvant chemotherapy has been widely researched for PDAC and therapy has continued to develop and change over the last 20 years. The latest research suggests that modified FOLFIRINOX therapy after recovery from surgical resection for PDAC has a significantly improved median overall survival compared with gemcitabine, at the expense of significantly increased complications (65). To date there doesn't appear to be comparison between combination modified FOLFIRINOX and gem-cap as adjuvant therapy however, extrapolation from study comparison with gemcitabine alone suggests the former is superior (17,65). The high rates of complications with

modified FOLFIRINOX therapy does mean this is only suitable for the most fit patients. Many patients (up to 40%) are unable to have any adjuvant chemotherapy due to post-operative complications and this has led to research into neo-adjuvant treatment for PDAC (66). It is likely, given medical advances that pre and post-operative treatment for PDAC will become more individualised both to the patient and to the individual's tumour biology (67).

Neoadjuvant treatment aims to reduce micro-metastasis and reduce the size of the primary tumour (68). A 2017 meta-analysis suggested that there was no significant benefit to neo-adjuvant therapy in resectable disease but that individuals with borderline resectable disease did benefit (either with chemotherapy or radiotherapy) (68). The group did caution that therapy pre-operatively could delay surgery and lead to increased surgical difficulty due to fibrosis. There were difficulties with this analysis given that many studies only had a single arm (with no control group) and there was significant variation in the studies. A phase III randomised controlled trial has just completed recruitment in the Netherlands assessing the benefit of neo-adjuvant FOLFIRINOX compared with neoadjuvant gemcitabine chemoradiotherapy in PDAC in both resectable and borderline resectable disease (69).

Non-resectable disease is treated with palliative chemotherapy where tolerated, otherwise with best supportive care (9). Disease treated in the early stages has a better prognosis with 5-year survival after resection ranging from 31% in stage I disease to 3% for stage IV, and to this end much research is ongoing aiming to improve early diagnosis of disease (70).

1.2 Pancreatic cancer screening and genetics

The lifetime risk of PDAC within the general population is low (roughly 1%), therefore population screening for early detection is not feasible, both due to cost, risks of invasive screening modalities and high ratio of false to true positives. General consensus suggests that screening for PDAC should not occur unless the individuals' risk is above 5% (71,72). It is estimated that up to 10% of PDAC has an inherited component (73).

1.2.1 Inherited pancreatic cancer risk – who to screen

Individuals with a higher risk of PDAC (>5%) includes both those with hereditary pancreatic cancer (HPC) and familial pancreatic cancer (FPC). HPC encompasses individuals with a known genetic syndrome which increases their risk of PDAC along with other cancers and is estimated to make up to 15-20% of inherited pancreatic cancer risk (73). A list of the most common syndromes which carry a significant increased risk of PDAC, their predisposing mutations and predicted cancer risks is found in Table 1.1. The lifetime risk figures are based on an average increased risk, in this thesis I will propose that this risk is context specific. *BRCA2* mutations make up the largest proportion of the HPC group (2).

Debate continues as to whether the risk is well defined enough in Lynch syndrome to deny individuals pancreatic screening and at present expert consensus suggests these individuals should still undergo radiological screening (74). Radiological screening for individuals with *BRCA1* mutations is also debated, with many experts suggesting this group should undergo regular PDAC radiological screening if they have a first-degree relative affected with PDAC (74).

Development of PDAC within an individual with HPC is suggested to follow the 'two-hit' hypothesis, where an underlying germline mutation leads to cancer after developing further somatic mutations in the same gene. This hypothesis is not without issue; for example Yurgelun and colleagues found that fewer than half of

the individuals within an HPC cohort that developed PDAC actually had a somatic mutation in the same gene (3).

Table 1.1: Genetic syndromes known to increase PDAC risk, predisposing mutations and cancer risks (syndromes with PDAC risk below 5% are shown in grey)

Cancer syndrome	Associated mutation(s)	PDAC risk (Lifetime risk)	Other Cancer risks
Hereditary breast and ovarian cancer (HBOC)	*BRCA1*	1-3%(75)	Breast, ovarian, prostate
	BRCA2	2-7%(71,73,75)	Breast, ovarian, prostate, stomach, melanoma
Peutz Jeghers syndrome (PJS)	*STK11/LKB1*	8-36% (76–78)	Colorectal, breast, small bowel, gastric
Hereditary pancreatitis (HP)	*PRSS1, SPINK1*	7-54% (41,79,80)	
Hereditary non-polyposis Colorectal cancer (Lynch, HNPCC)	*MLH1, MSH2, MSH6, PMS2*	3.7% (poorly defined risk) (71,73,81)	Colorectal, endometrial, small bowel, ureter/renal pelvis
Familial atypical multiple mole melanoma (FAMMM)	*P16/CDKN2A*	17% (71,73,81)	Melanoma, breast, lung, endometrium
Familial adenomatous polyposis (FAP)	*APC*	1.7% (73,81)	Colorectal, hepatoblastoma, duodenal, thyroid, bile duct, brain

Screening for HPC individuals carries challenges, including that multiple studies have found that a significant proportion of individuals who develop seemingly sporadic PDAC actually have an underlying genetic mutation but with little or no family history of cancer (3,72). Some groups have gone on to suggest that germline mutation testing should be performed for predisposing mutations in all individuals who develop a 'sporadic' PDAC (72). Whilst this may aid family members of those who develop PDAC understand and manage their risk by encouraging them to seek radiological screening, this does not provide a method for detection of germline mutations which increase the risk of PDAC in those with limited or no family history. Unfortunately this highlights the difficulties with the pleiotropic nature of many of

these mutations along with variation in penetrance leading to a complex set of cancer syndromes (82).

The remainder of the population with inherited pancreatic cancer risk is thought to be individuals with FPC. FPC has been extensively researched and at present there still remains no gene(s) found to be responsible for this inherited risk. It is thought that FPC is of autosomal dominant inheritance with a reducing age at presentation for PDAC with each successive generation (73,82). This group brings considerable complexity given; the causative gene is unknown, there is the possibility of multiple sporadic cancers within families, and an unknown penetrance which means that individuals and possibly families without this predisposing mutation would still undergo high risk PDAC screening. This carries the risks of undergoing invasive screening, considerable anxiety and the risk of false negatives along with lowering the yield of malignant or pre-malignant lesions within the screening programme.

The International Cancer of the Pancreas Screening (CAPS) Consortium has reported that the lifetime risk of PDAC in an individual with two affected first-degree relatives is 8-12% and most agree that this group should undergo high-risk screening for PDAC (71). More controversy exists about whether individuals should be screened who have two family members affected with PDAC but only one is a first-degree relative however, still 77.5% recommended this group be screened (71).

1.2.2 When and how to screen

Whilst there is a general consensus that individuals with greater than 5% lifetime risk of PDAC undergo screening there is little consensus about the age at which screening should start, when it should finish, what screening modality should be used and the screening schedule.

The CAPS consortium discussed age at which PDAC screening should commence for individuals with FPC, and 68% agreed that this should be at 50 years old or within 10 years of the youngest PDAC affected relative (74). 67% of the group also suggested HPC screening should commence at 40 years old or within 10 years of the youngest

PDAC affected relative (74). Interestingly in an earlier CAPS consensus publication there was the suggestion that individuals who smoke tobacco should start screening 5 years earlier (55.1% agreed) however, this is not included in the majority of PDAC screening programme criteria (71). Whilst lifestyle factors play a considerable role in cancer risk it is impossible to quantify on an individual basis to allow adaptation of cancer screening.

PDAC screening cessation remains a contentious issue. There remains no consensus within the literature and between experts of when to stop radiological screening for high-risk individuals (71,74). At present The European Registry of Hereditary Pancreatitis and Familial Pancreatic Cancer (EUROPAC) group within the UK advocates cessation when the individual is deemed to be not fit for a pancreatic resection however, this is based on individual clinician perspectives.

Multiple radiological imaging methods are available for PDAC; including endoscopic ultrasound (EUS), computed tomography (CT), magnetic resonance imaging (MRI) and ultrasound (US) however, there remains positive and negative factors for the use of each as a tool for early detection. Whilst the debate continues over whether MRI, EUS or both is best for continued screening, most agree that ongoing imaging with either of these modalities is beneficial and avoids the considerable ionising radiation produced from repeated CT imaging (71,74). The CAPS consortium agreed (90%) that pancreatic screening should occur at 12 monthly intervals if normal (71,74).

Multiple blood markers have been suggested for use alongside high-risk PDAC screening. Carbohydrate antigen 19-9 (CA19-9) is the most well-known and well-researched PDAC blood marker. There continue to be significant problems with the use of CA19-9 as a screening marker alone for PDAC. Elevation in CA19-9 is known to occur prior to PDAC diagnosis however, it has a very low positive predictive value, is not produced in some individuals and can be raised in other pathologies, such as pancreatitis (74,83). It has been suggested therefore that CA19-9 be used in the event of abnormal imaging, rather than as a single screening marker (74,83).

Multiple other biomarkers have been suggested and researched for detecting PDAC however, at present none are validated for use for early detection of PDAC (84). Research is ongoing in this area and multiple groups have suggested that a combination of blood markers rather than a single indicator may have better sensitivity and specificity for early detection of PDAC (84,85).

The goal of pancreatic screening is to "prevent death from pancreatic cancer and prevent its emergence by identifying and treating precursor lesions" (74). Multiple studies to date have examined the benefit of radiological screening for PDAC in individuals deemed high risk. Prospective long-term follow-up studies have suggested a benefit to PDAC screening in detecting lesions at a resectable stage, therefore improving long-term survival (86,87). Radiological screening detected neoplastic lesions in between 1 and 7% of the screened individuals in these studies however, this was heavily dependent on the population screened and the length of follow-up (86,87). Both studies warn of the risk of false positives leading to individuals having a pancreatic resection with no underlying malignant or pre-malignant lesion, especially when aiming to treat at the precursor stage (86,87). These individuals undergo significant, potentially life-threatening surgery which may be unnecessary. In a meta-analysis of cohort studies of PDAC screening Corral and colleagues suggested that 135 high risk individuals needed to be screened to find one patient with a high risk lesion (88).

Surveillance for branch duct IPMNs and other cystic lesions remains contentious. Sheel and colleagues found that presence of branch duct IPMNs were independent of genetic predisposition to cancer and therefore should be treated as if found in the general population (89). International consensus guidelines suggest that individuals with branch duct IPMNs should initially undergo surveillance with MRI 3-6 monthly then screening interval should be based on the size of the lesion (16).

1.3 EUROPAC

The European Registry of Hereditary Pancreatitis and Familial Pancreatic Cancer (EUROPAC) study was established in 1997 in Liverpool, UK (90). This continues as a multi-centre study providing information and screening to individuals deemed at high risk of PDAC, alongside collecting samples and data to progress research in this area. The group split recruitment of individuals into either the FPC registry (which encompasses FPC and HPC as discussed above) or the HP registry allowing more targeted radiological screening of these individuals. EUROPAC is funded by Pancreatic Cancer UK (PCUK) (90).

1.4 PanFam-1 & IMMPACT studies

A prospective, multi-center investigational study of IMMray™ PanCan-d diagnostic platform for early detection of pancreatic ductal adenocarcinoma in high-risk populations (PanFam-1) is currently analysing data from individuals deemed at high risk of PDAC (85). Immunovia (Lund, Sweden) have completed recruitment of individuals from multiple sites worldwide to assess a multi-marker platform using blood markers to detect early PDAC. This platform will be compared with the radiological screening results from high-risk individuals already within radiological screening programmes and aims to validate the multi-marker platform for early PDAC diagnosis in this group. Recruitment was completed in October 2020 with results expected in 2022 (85).

The Clinical validation of a serum protein biomarker signature for the early diagnosis of pancreatic cancer and cystic neoplasms (IMMPACT) study was created in Liverpool in October 2017 in collaboration with Immunovia aiming to recruit high risk individuals undergoing radiological PDAC screening. This study includes recruitment of individuals suitable for the PanFam-1 study and indeed there is

agreement to share anonymised data and samples between the University of Liverpool and Immunovia.

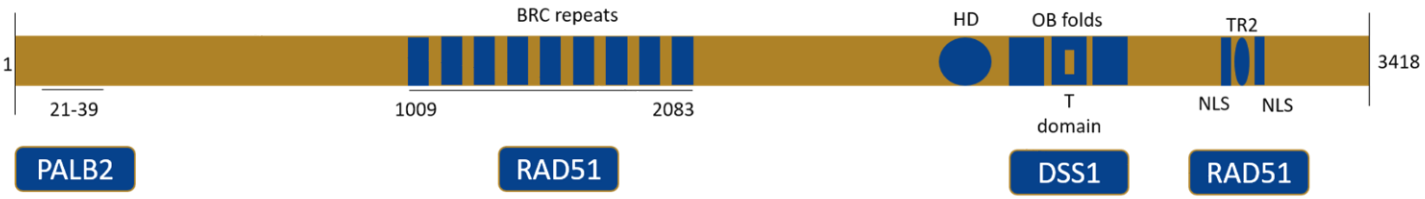
1.5 BRCA2

Germline mutations in the *BRCA2* gene are a relatively recent discovery. These were first made by a group in 1995 who found that a mutation on chromosome 13 appeared to be linked to breast, ovarian, prostate and pancreatic cancers (91). At an early stage the patterns of disease were noted to be consistent with a mutation in a tumour suppressor gene within this region. The *BRCA2* gene is found on the long arm of chromosome 13 (13q12.3) and is comprised of 27 exons, with exon 10 and 11 making up 60% of the coding region (92,93). It is generally thought that *BRCA2* mutations lead to increased rates of cancer following the 'two-hit hypothesis' (94).

1.5.1 Protein Structure

BRCA2 protein is comprised of 3418 amino acids which make up several important structural elements that are vital for the complete functioning of the protein (93). A schematic diagram of the BRCA2 protein is shown in figure 1.1. Firstly the BRC repeats are known to effect the binding with RAD51 which is essential for homologous recombination DNA repair (93,95). The DNA binding domain encompasses the helical domain (HD), three oligo-nucleotide binding folds (OB folds) and the tower domain (T domain). BRCA2 can act to protect DNA (particularly single stranded DNA at replication forks) as well as to repair double strand breaks once they have occurred. Surprisingly previous research has suggested that BRCA2 protein without the DNA binding domain does not impact the protection of stalled DNA forks (95). The TR2 domain also binds to RAD51 but has been linked to DNA protection rather than homologous recombination DNA repair and mutations in this region have been linked to increased tumorigenesis (95).

Figure 1.1: Structure of the BRAC2 protein (adapted from Fradet-Turcotte *et al.*)
(Boxes show proteins known to bind at specific regions of BRCA2)



1.5.2 Protein Function

BRCA2 is a multifunctional, predominantly nuclear protein. The most important function is as a mediator of double-strand break (DSB) DNA repair (93). DSBs are known to be highly cytotoxic, especially when occurring during DNA replication. BRCA2 protein aids in DSB repair by homologous recombination, using the undamaged DNA template for repair during the S or G2 phases of the cell cycle (93,96). Without functioning BRCA2 protein DNA undergoes classical non-homologous end joining which is error prone and can cause deletions and translocations, but is intrinsically much faster than homologous recombination (93,96,97).

BRCA2 protein also protects degradation of replication forks when stalled after oxidative stress. This allows for recruitment of proteins to facilitate DNA repair (95).

1.5.3 Impact of BRCA2 mutation

Mutations in the *BRCA2* gene are of variable penetrance and found in population studies to affect roughly one in 500 people (97,98). *BRCA2* and *BRCA1* mutations are linked to a syndrome with increased breast and ovarian cancer risk, HBOC. An individual with a *BRCA2* mutation has a lifetime risk of breast cancer between 40 and 84% and ovarian cancer between 11 and 27% (99). Families with *BRCA2* mutations have a higher incidence of male breast cancer and have also been linked to cancers of the prostate, pancreas, gall bladder, stomach and skin (97,100).

At present all individuals with a mutation in the *BRCA2* gene are treated as though they have a similar cancer predisposition. Multiple studies have found that there is a considerable variation in risk of cancers within individuals and families with *BRCA2* germline mutations (98,100,101). This suggests there are underlying modifiers of this risk, which will be further examined in this thesis. The Consortium of Investigators of Modifiers of *BRCA1/2* (CIMBA) have carried out multiple genome-wide association studies (GWAS) and have linked single nucleotide variants (SNVs) at individual alleles to an increased risk of ovarian cancer and others to breast

cancer in individuals with *BRCA2* mutations (102–104). Though attention has been paid to the risk variation in individuals with *BRCA2* mutations there is no current integration of this into personalisation of risk discussions or screening and whilst this would benefit individuals with *BRCA2* mutations it may be too complex to fully describe. To date no such investigation has been performed assessing the impact of SNVs on the development of PDAC in individuals with *BRCA2* mutations.

Treatment of PDAC in individuals with germline *BRCA2* mutations is broadly similar to that in the population with the possible exception of choice of specific chemotherapeutic agents. PDAC in such individuals is thought to be more susceptible to platinum-based chemotherapeutic agents and poly-ADP-ribose polymerase inhibitors however, as yet there is ongoing debate about whether these agents confer any increased survival (105). The most convincing evidence for the benefit of personalised medicine based on *BRCA2* mutations is in favour of continued use of platinum-based chemotherapeutic agents for metastatic PDAC in individuals with germline *BRCA2* mutations (106).

1.5.4 Management of individuals with BRCA2 mutations

NICE recommends genetic counselling and support for both individuals suspected of HBOC and individuals with a confirmed pathogenic mutation along with information about modifiable risk factors and risk reduction (107).

At present there are multiple prophylactic options open to individuals with confirmed pathogenic *BRCA2* mutations. Chowdhury and colleagues discussed the benefit of personalised prophylaxis for individuals with pathogenic mutations and divided this into primary prophylaxis, with use of endocrine drugs (tamoxifen or anastrazole) and adaptation of lifestyle factors and secondary prophylaxis, with radiological screening for cancers (107,108). There is also the possibility of prophylactic surgery such as mastectomy and bilateral salpingo-oophorectomy. The risk of breast and ovarian cancer is known to be substantially reduced in individuals who undergo a bilateral salpingo-oophorectomy, and this is known to be reduced further if this occurs before the age of 45 years old (98).

Radiological screening is available for individuals with *BRCA2* mutations. Firstly, as previously mentioned this is not adapted to actual individual risk, but rather based on the presence or absence of the pathological mutation itself. Chowdhury and colleagues discuss the benefits of risk stratified cancer screening being able to reduce cost as well as reduce false positives and over or under treatment (108). PDAC screening is carried out for individuals with *BRCA2* mutations with a first-degree relative affected with PDAC. NICE recommends breast surveillance with annual MRI scans between 30 and 49 years old and mammographic surveillance between 40 and 69 years old for individuals with a confirmed mutation (107). Some groups also advocate ovarian screening with regular Ca125 blood testing and transvaginal ultrasound scans however, this has to be provided alongside caution that this is not a guaranteed method of early identification of cancer (109).

1.5.5 Testing individuals for BRCA2 mutations

Current guidelines for offering individuals *BRCA1/2* testing vary considerably internationally. Toss and colleagues tested each of the most commonly used criteria in a retrospective study of individuals with PDAC (75). The group noted that the American National Comprehensive Cancer Network Guidelines allowed for the testing of 100% of the individuals with HBOC (given the criteria include testing individuals with PDAC as a matter of course) with a detection rate of 21.3%, whereas the Modena Criteria only allowed for testing of 55.4% of the individuals with a 24.5% detection rate (75). The group go on to discuss that the difference in criteria is a balance between cost and having testing available as widely as possible.

In the UK the NICE criteria for *BRCA* mutation testing are very stringent, and this perhaps further reflects a balance of cost alongside appropriate testing (107). The NICE guidelines are listed in the Materials and Methods section.

Interestingly several groups have suggested that the current strict *BRCA* mutation testing criteria are reflective of the era where mutation testing was very expensive. It may be that with the emergence of multiple more modern sequencing platforms that are cheaper and can process a high volume of samples more quickly, that more

individuals could be tested for pathogenic mutations as a matter of course. The NICE guidelines however, were reviewed in 2019 and do not include any such change in practice (107). Again, these guidelines assume a direct causative association between a pathogenic *BRCA2* mutation and cancer risk. This thesis will challenge this and indicate that context needs to be taken into account as well as primary mutation status.

1.6 Third generation sequencing

'Next generation' sequencing (NGS) broadly describes any high throughput DNA sequencing platform, this has now been subdivided to include second and third generation sequencing platforms. Single molecule and the possibility of ultra-long DNA reads are what distinguishes third generation sequencing (110,111). To date two platforms are available for third generation sequencing; Single molecule real time sequencing (Pacific Bioscience) and nanopore sequencing (Oxford Nanopore Technologies) (110).

Third generation sequencing generally requires much less expensive equipment, less computational space and is able to sequence much longer molecules at one sitting compared to second generation sequencing (110,111). This does however, come at the expense of accuracy (112,113). When sequencing short strands of DNA it is likely that second generation sequencing is preferable, given the error rate is significantly lower. However, when sequencing long strands, short-range sequencing platforms require multiple fragments to be sequenced and then pieced together, which is much more labour intensive, expensive and can be incomplete and, of particular relevance to this thesis, does not allow direct linkage of variants to a single strand of DNA (112).

1.6.1 Nanopore sequencing

Oxford Nanopore Technologies (ONS) released the MinION platform in 2014 and nanopore sequencing functions by measuring the changes in current across a pore whilst bases are passing through (114). At present, when compared with platforms from Pacific Bioscience, ONS platforms allow for much longer read lengths (longest read for ONS 2Mb, for Pacific Bioscience ~20Kb) along with real-time base calling which allows for continual analysis (110). This potentially means that when creating haplotypes of multiple SNVs it is more accurate with the ONS platforms as sequencing of long DNA strands is possible, rather than piecing together smaller segments. Alongside this the MinION is a very small sequencing platform which is much cheaper than the Pacific Bioscience platforms (110).

Multiple studies have been carried out using the MinION for long DNA reads and have suggested that, with specialised software and data correction, accuracy can be increased to 99.8% (112,115). This does still require caution however, given that for ultra-long reads this can still be a considerable actual number of errors. Coupled with this it has also been shown that errors from the MinION are generally concentrated in specific regions, which therefore cannot be corrected by repeated sequencing of the same fragment. To account for this Watson and Warr suggest that short-range platforms should be used to check areas with mutations (112). This has been echoed by other researchers who have dubbed this ‘hybrid sequencing’ and suggests this allows for the long-range sequencing from third generation platforms with the accuracy of second generation platforms (114).

2. Aims

To determine the high-risk population that are more likely to attend for pancreatic cancer screening on the EUROPAC registry

To describe the families with *BRCA2* mutations currently on the EUROPAC registry

To optimise a method for enrichment, amplification and sequencing of *BRCA2* and surrounding genes

To define and investigate haplotypes amongst *BRCA2* mutation positive individuals in DNA surrounding the gene and how this may correspond to familial risk of pancreatic cancer

3. Objectives

Recruiting individuals at higher than population risk of pancreatic cancer to the IMMPACT study along with blood sample collection from these individuals

Co-ordinating the IMMPACT study (including submissions to Health Research Authority (HRA) and ethics committee)

To evaluate the population recruited to the EUROPAC study, assessing individuals who are likely to attend for screening and those who are not

To create a set of clinical criteria to subdivide families with *BRCA2* mutations to allow stratified analysis of cancer risk

To fully analyse whether *BRCA2* mutation families carry different risk profiles for pancreatic and breast cancer

To assess possible confounding factors in the relationship between *BRCA2* families and cancer risk

To optimise and test a method for enrichment of required genome sequence prior to sequencing

To assess the presence of genome SNVs local to *BRCA2* and their effectiveness in further defining subgroups of *BRCA2* mutation carriers

4. Materials and methods

4.1 Recruitment to EUROPAC

The EUROPAC study was commenced in 1997 at The University of Liverpool. The EUROPAC study allows for both self-referral and clinical referral (such as from a General Practitioner (GP) or geneticist) to the team. Data is then collected about the demographics, clinical and family history of the individual and their family. Individuals referred have or are perceived to have an increased risk of pancreatic cancer. Reported history of pancreatic cancer is verified where possible using cancer databases, death certificates or patient notes. This may not always be possible for a diagnosis of pancreatic cancer abroad or prior to established accurate cancer diagnosis record keeping. Individuals are then assessed to deem whether or not they should be offered secondary screening. This assessment is completed by a clinical team to assess against the accepted high-risk criteria for pancreatic cancer and the inclusion criteria for EUROPAC alongside a discussion with the individual about the risks and benefits of screening.

This includes individuals with:

1. Genetic mutations known to increase the risk of pancreatic cancer (*BRCA2, HNPCC* mutations (*MLH1, MSH2, MSH6, PMS2*), Peutz-Jeghers syndrome (*STK11*), Familial Atypical Multiple Mole Melanoma Syndrome (*CDKN2A*)) alongside a family history of pancreatic cancer (one confirmed affected relative) – HPC individuals
2. Two or more first or second-degree relatives with confirmed pancreatic cancer- FPC individuals
3. Hereditary pancreatitis – confirmed *PRSS1* or *CFTR* mutations alongside episodes of recurrent pancreatitis

A referral from the individuals GP is sought prior to commencement of secondary screening.

Individuals fitting criteria 1 and/or 2 who are aged over 40 years, or less than ten years younger than any affected relative, are suitable for secondary screening with an initial computed tomography (CT) scan and annual endoscopic ultrasound examination (EUS), clinical examination, Ca19.9 and glycosylated haemoglobin (HbA1c) tests. Individuals fitting criteria 3 aged over 40 are suitable for secondary screening annually with a CT scan, clinical examination, Ca19.9 and HbA1c tests. All individuals suitable for secondary screening were first seen by a clinician at The Royal Liverpool University Hospital and were then referred on to their closest secondary screening centre if this was their preference. Secondary screening and clinical follow-up then occur at this preferred site, with results of scans and blood tests being sent to the EUROPAC team centrally. At the time of the database searches EUROPAC had ethical approval for the FPC database (individuals with FPC and HPC) from the North West Multi Centre Research Ethics Committee (24/09/03, MREC 03/08/069) and for the FPC screening from Warwickshire Research Ethics (30/10/07, MREC 07/H1211/96). The hereditary pancreatitis (HP) database had ethical approval from Scotland Research Committee (21/04/04, MREC 04/0/010) and for screening from Central Manchester Research Ethics Committee (26/10/07, MREC 07/H1008/153). This was amalgamated into a single ethics approval before the time of writing by the Yorkshire and Humber Research Ethics Committee (25/07/19, MREC 19/YH/0250)

4.2 Search and Analysis of EUROPAC database

The EUROPAC database is stored on a single hard-drive, not connected to the internet which is password protected to ensure confidentiality. The database software is Progeny© version 8 (Progeny Genetics LLC) which is specifically designed to allow storage of demographic data, genetic data and clinical data linked to familial relationships thus allowing construction of complex family trees, as well as pedigree dependent data searches. For ease of data storage, the EUROPAC data is stored on two separate databases entitled FPC and HP databases. The FPC

database contains information about individuals and families who are designated FPC or HPC. The HP database contains information about individuals with hereditary pancreatitis.

The EUROPAC database was searched on the 13th October 2017 using the following parameters to find individuals suitable for screening (separate searches each for FPC and HP databases):

1. Deceased status – Does not equal 1 (i.e. individual currently alive)
2. Risk category – high
3. Screening consent completed
    a. Is blank - for the non-screened population
    b. Is not blank – for the screened population

The database then provided demographic details, postcode (where available), genetic information, health and social information and family history for each individual. This information was checked alongside the paper health records held by the EUROPAC team to ensure information was correct.

Missing data was analysed to ensure that data gaps were fully explored to allow assessment of data bias.

Demographic information collected included; age, self-reported gender and postcode.

Health and social information collected included; family and personal history of cancer, with level of confirmation of diagnosis of pancreatic cancer (ranging from histological confirmation to cancer registry confirmation and beyond this to unconfirmed, i.e. word of mouth), number of family members with a confirmed diagnosis of pancreatic cancer, smoking history, alcohol history and any diagnosis of diabetes mellitus.

Distance from the closest screening centre was assessed using Google Maps (www.google.co.uk/maps). The available centres for EUROPAC screening (at the time of writing) are The Royal Liverpool University Hospital (L7 8XP), University

College Hospital (London, NW1 2BU), Queen's Medical Centre (Nottingham, NG7 2UH), St. James's University Hospital (Leeds, LS9 7TF), Bristol Royal Infirmary (BS2 8HW), Southampton General Hospital (SO16 6YD), Glasgow Royal Infirmary (G4 0SF), Freeman Hospital (Newcastle, NE7 7DN) and Morriston Hospital (Swansea, SA6 6NL). Initially, due to service provision within the NHS and the EUROPAC study individuals could only be screened at their local site or Liverpool, they could not choose other sites. Alongside this, some screening centres only allow secondary screening for individuals within their local vicinity, leading to some individuals having to undergo their screening at Liverpool. Since the release of the NICE guideline in 2018 suggesting individuals at high risk from pancreatic cancer should undergo screening more centres aside from those registered with EUROPAC have offered secondary screening (116).

A sociodemographic score was calculated from figures using StreetCheck (www.streetcheck.co.uk) which collects data from the Office for National Statistics. The StreetCheck shows the number of people in a specific postcode in each social grade; AB – higher and intermediate managerial, administrative, or professional positions, C1 – supervisory, clerical, and junior managerial/administrative/professional positions, C2 – skilled manual workers, DE – Semi-skilled and unskilled manual workers; those on state benefit/unemployed, & lowest grade workers. Postcodes were allocated a score based on the percentage of individuals in each group with 1 being the most and 4 being the least affluent.

Body mass index (BMI) was calculated by Progeny as per the latest weight (Kg)/height(m)^2. Individuals were then grouped according to the accepted criteria for BMI; <18.5 underweight, 18.5-24.9 normal, 25-29.9 overweight, 30-39.9 obese, ≥40 morbidly obese.

Statistical analysis was performed using JMP© Student Edition 14 (SAS). Differences in distributions of age and distance from centre (Km) between screened and non-screened groups were described using violin plots (JMP© Student Edition 14 (SAS)) and assessed using the Kolmogorov-Smirnov Test. Comparison between two binary

nominal variables, such as self-reported gender and whether individuals were screened, were assessed using a 2-Tail Fisher's Exact Test. Comparison between multiple group ordinal and binary nominal variables, such as BMI and whether individuals were screened were assessed using a Chi-squared-test. Bar charts were created using Microsoft Excel (Microsoft Office and Home Student 2019).

4.3 IMMPACT study

The Clinical validation of a serum protein biomarker signature for the early diagnosis of pancreatic cancer and cystic neoplasms (IMMPACT) study was developed as a collaboration between multiple international sites and commenced recruitment in March 2017. There was also an agreement that this would allow collaboration with Immunovia (Lund, Sweden) and allow samples to be processed as part of their international study to assess a diagnostic platform (A prospective, multi-centre investigational study of IMMray© PanCan-d diagnostic platform for early detection of pancreatic ductal adenocarcinoma in high-risk populations (PANFAM-1)). Inclusion criteria for the IMMPACT study closely mirrored those of the PANFAM-1 study to facilitate this collaboration. Individuals deemed at highest risk from developing pancreatic cancer during the study period (2 years follow-up) and who currently undergo secondary screening with EUROPAC were eligible. The study aimed to compare blood markers at 6 monthly intervals (5 samples in total) with image screening results. Individuals were included if they:

- Were able and willing to provide informed consent for the study and follow-up
- Attended for secondary screening for pancreatic cancer

AND

- Had two or more first degree relatives with a confirmed diagnosis of pancreatic cancer (aged over 50 years old or within 10 years of age of onset of pancreatic cancer in family member)

OR

- Three or more family members with a confirmed diagnosis of pancreatic cancer (aged over 50 years old or within 10 years of age of onset of pancreatic cancer in family member)

OR

- Any mutation confirmed pathogenic or likely pathogenic with a first or second-degree relative with pancreatic cancer (*BRCA1, BRCA2, PALB2, ATM, MLH1, MSH2, MSH6, PMS2, EPCAM*) (aged over 50 years old or within 10 years of age of onset of pancreatic cancer in family member)

OR

- Familial atypical multiple mole melanoma (FAMMM) with a confirmed pathogenic or likely pathogenic mutation in *p16* or *CDKN2A* (aged over 50 years old or within 10 years of age of onset of pancreatic cancer in family member)

OR

- Known *STK11* mutation carrier (Peutz Jegher's syndrome) (aged ≥ 35 years)

From 27th November 2018 individuals were also included if they had hereditary pancreatitis with a confirmed *PRSS1* mutation (aged over 50 years old or within 10 years of age of onset of pancreatic cancer in family member).

Individuals were contacted by post with a covering letter and a copy of the latest patient information sheet. Those who were keen to be involved were seen at The Royal Liverpool University Hospital and informed consent was gained after a discussion with an appropriate clinician. Individuals were able to decline the study or withdraw from the study at any point without repercussions or any alterations with their involvement with EUROPAC.

Once consented for the study blood samples were obtained from the individuals and processed at the University of Liverpool by GCP trained laboratory staff within 30 minutes of collection where possible. The individual was then contacted by their preferred method (email or telephone) to attend at 6 monthly intervals for further blood samples until 2-year follow-up. Image results were obtained through The

Royal Liverpool Hospital system where screening occurred locally and from the EUROPAC individual notes where screening occurred away from Liverpool. Efforts were made to recruit and collect samples from other EUROPAC sites however, due to sample processing times and facilities at other sites this was not possible. Ethical approval was obtained from the Liverpool East Ethics Committee (04/04/17, MREC 17/NW/0170). Ethical approval was also obtained to access previously collected and stored samples where the appropriate consent was available.

Data was stored on a separate central, secure database and did not contain patient identifiable information. Description of the patient data was performed and compared with the screened group of the EUROPAC database to ensure that the group was representative. Statistical analysis was performed in line with the analysis of the EUROPAC database.

Power calculation for the PANFAM-1 study assumes the incidence of 2.4% over three years with a sensitivity and specificity of the IMMRAY© platform of 94% and 85% respectively. With a greater than 80% power the enrolment goal over 3 years was enrolment of 2000 subjects. A separate power calculation for the IMMPACT study was not performed due to this being a local contributor to the PANFAM-1 study.

4.3.1 Blood sample collection & processing

Venepuncture was performed with a 21G needle and 60ml blood collected. This included collection into a K^+EDTA tube for plasma and cell pellet, a K^+EDTA for peripheral blood mononuclear cells (PBMCs) and serum tubes.

K^+EDTA tubes were inverted 8-10 times immediately after collection and were processed within 15 minutes of collection where possible. The whole blood was then transferred into pre-labelled sterile Eppendorf tubes and centrifuged at 16000g for 1 minute. The top plasma layer was then carefully placed into cryovials and the cellular component left in the tubes. Samples were frozen at -80°C until required.

Serum tubes were inverted 6 times to activate clot formation then allowed to stand for 45 minutes prior to processing. Tubes were then centrifuged at 1200g for 10 minutes at room temperature. The serum layer above the clotted blood was then carefully removed and transferred into cryovials before being stored at -80°C until required.

Processing of study blood samples was performed by appropriately trained, designated GCLP technicians within the Liverpool GCP lab facility.

4.4 BRCA2 family analysis

The EUROPAC database was searched on the 7th March 2019 for families with:

1. BRCA2 mutation free text = 'is not blank'

OR

2. BRCA2 mutation = positive

Additional families in folders deemed '?HBOC' were also included. Families were then searched through manually to ensure that there was evidence of at least one individual having a confirmed BRCA2 mutation.

To ensure that only the families with pathogenic mutations were included each mutation was searched using multiple mutation databases and publications to ensure that any mutation was confirmed to be pathogenic. Any non-pathogenic mutations were excluded. For families with a confirmed *BRCA2* mutation in medical notes but with no clear record of the mutation type several attempts were made to ascertain the exact mutation details, this included contacting the GP and genetics team of any individual who had tested positive and did not have the information available.

4.4.1 Division into family groups

Family trees were assessed individually to ensure they were an accurate representation of the family history found in the paper EUROPAC notes. Pseudo-anonymised printouts of each family were provided to each investigator (two clinicians and one geneticist) along with ages of diagnosis for any cancers within the families where the information was available. The investigators individually assessed each family tree against the criteria for Familial Pancreatic Cancer (Pancreatic cancer criteria) and hereditary breast and ovarian cancer (HBOC).

HBOC criteria: adapted from NICE clinical guideline (CG164, updated 20th November 2019)(107)

1. Metachronous breast and ovarian cancer
2. Breast cancer diagnosed ≤35 years old
3. Bilateral breast cancer (with the first cancer diagnosed ≤40 years old)
4. Triple negative breast cancer diagnosed ≤50 years old
5. High grade non-mucinous ovarian cancer (or fallopian or primary peritoneal cancer)
6. ≥2 first degree relatives with:
    a. Bilateral breast cancer and another breast cancer diagnosed ≤50 years old
    b. Male breast cancer
    c. 2 breast cancers diagnosed ≤50 years old
    d. ≥3 first-degree relatives with breast or ovarian cancer

Pancreatic cancer criteria: adapted from FPC EUROPAC inclusion criteria

1. Two or more first or second-degree relatives with confirmed pancreatic cancer

Families were then grouped into those fitting:

1. HBOC criteria alone
2. Pancreatic cancer criteria alone
3. Both the HBOC AND pancreatic cancer criteria
4. Neither the HBOC NOR pancreatic cancer criteria

The assessors divided the families into the above groups with agreement being reached by discussion and consensus.

4.4.2 Survival analysis for the family groups

The family trees were assessed, with information collected about each family and each individual. Firstly, the families as a whole were analysed. Information was collected about the mutation type within the family, the number of individuals within the family with pancreatic cancer, breast cancer, ovarian cancer or prostate cancer. Information was also collected where possible about the age of diagnosis of each type of cancer. This information was averaged within the family to provide the average age at diagnosis for each cancer type. The different family types were compared to assess whether the type of mutation (i.e. frameshift or single-nucleotide variant) was more common to a specific family group. Specific shared mutations were also compared to assess whether this could predict the family group. A step-wise analysis of the *BRCA2* mutation position was performed using Kaplan-Meier survival curves (JMP© Student Edition 14 (SAS)) to assess whether mutation position lead to an increased predisposition to each family group.

Average age of pancreatic cancer diagnosis, average age of breast cancer diagnosis, average number of pancreatic cancers per family and average number of breast cancers per family were assessed and compared between the family groups.

Survival analysis was firstly performed using only those individuals with a diagnosis of pancreatic, breast, ovarian or prostate cancer. This compared the family groups by age at diagnosis of the cancer. Kaplan Meier survival curves were created using JMP© Student Edition 14 (SAS). This was used to provide comparison of whether at risk individuals in each family group were diagnosed with specific cancer types at different ages.

Secondly, to assess the risk profile of the different cancer types between the family groups every individual with direct family linkage were included in the analysis to assess whether certain families were more predisposed to each cancer type. Again

Kaplan-Meier survival curves were created (JMP© Student Edition 14 (SAS)) and censored for those individuals with a diagnosis of pancreatic, breast, ovarian or prostate cancer.

4.4.3 Lifestyle data analysis for the family groups

Cigarette smoking and diagnosis of diabetes mellitus data was then analysed from the family groups to assess for possible confounding of the survival profiles of each BRCA2 family group. Firstly, data completeness was analysed to assess discrepancies in data availability between the family groups. Then the numbers of cancers in the smoking and non-smoking groups were compared as well as those with and without diabetes.

A survival analysis (Kaplan-Meier survival curves (JMP© Student Edition 14 (SAS))) was then performed comparing all of the BRCA2 individuals who smoked and those who did not; censored for pancreatic cancer or for breast cancer. This analysis was repeated for individuals with and without a diagnosis of diabetes. This was to assess whether there was any significant difference in survival from pancreatic cancer or diagnosis of breast cancer in these individuals.

Lastly the survival analyses for pancreatic and breast cancers were repeated in the BRCA2 family groups but with censorship for current and ex-smokers.

4.5 Enrichment for BRCA2 chromosomal locus

4.5.1 Cell culture

Fibroblast cell culture was performed using cells from a cancer patient recruited in the Royal Liverpool Hospital (patient R2797). Fibroblasts were used from passage 14 (c/o Lawrence Barrera-Briceno) activated fibroblasts in clean conditions in a tissue culture hood. Cells were thawed from -80°C storage, immediately placed into 5ml warmed Dulbecco Modified Eagle Medium (DMEM, high glucose 4.5g/L, GlutaMAX,

Life Technologies, 1966047) with 10% Foetal Bovine Serum (FBS, Life technologies, 10270106) and centrifuged at 200g for 5 minutes. The supernatant was then discarded and the cell pellet resuspended in warmed DMEM with 10% FBS. Cells were initially transferred into a T75 flask and cultured in 10ml media. Once the cell cultures were established larger cultures were grown in T175 flasks with 30ml media. Cells were cultured in an incubator at 37°C, humidified with 5% CO_2.

To detach cells 2ml warmed 0.05% trypsin (Life technologies, 25200056) was added to the T175 flask (1ml for T75 flask) and placed into an incubator (37°C, humidified with 5% CO_2) for 5 minutes. Cultures were viewed under a microscope to assess detachment from the flask. 8ml warmed DMEM with 10% FBS was added along the side of the flask to precipitate detachment. The fibroblast cultures were then split into culture flasks as deemed appropriate or used directly for formation of agarose cell plugs. Cell counts were performed manually using a haemocytometer.

4.5.2 Formation of agarose cell plugs

Agarose cell plugs were formed using a CHEF mammalian genomic DNA plug kit (BioRad, 170359) as per manufacturer's instructions. Plugs were formed at 1% agarose concentration. Initially 2 x 10^6 R2797 fibroblasts per plug were used however, due to the large cellular content these were not stable, therefore this was reduced to 1 x10^6 cells per plug. When forming cell plugs with BRCA2 individual PBMC samples 1.5 x10^6 cells per plug were used. Cells were first washed with phosphate buffered saline (PBS x1, pH7.4, Life technologies, 10010056) and centrifuged at 300g for 5 minutes before being resuspended in 10ml PBS. 2% Cleancut agarose (BioRad) was melted and equilibrated to 50°C in a water bath. Cells were centrifuged at 1000g for 5 minutes at 4°C, the supernatant was discarded and then resuspended in suspension buffer (concentration as required for 1% plug concentration) and equilibrated to 50°C. The appropriate quantity of agarose was then added and carefully mixed before transferring into plug moulds (75µl per plug). Moulds were then placed at 4°C for 30 minutes. Solidified plugs were placed

into 1.5ml Eppendorf tubes containing 30µl Proteinase K (BioRad) stock solution and 750µl Proteinase K reaction buffer (BioRad) overnight incubated at 50°C.

4.5.3 Cell lysis within plugs

Plugs were then washed three times in 10ml 1xTE buffer, manually shaking for 10 seconds. Buffer was discarded between washes. Plugs were then incubated with 50µl RNaseA (100mg/ml, Quiagen, 19101) in 2.5ml 1xTE buffer with occasional shaking for 1 hour at 37°C. Plugs were then washed four times with 1x wash buffer (1ml/plug) with the second wash also containing 40µl 1mM PMSF (Thermofisher, 36978) each wash for 1 hour with gentle agitation at room temperature. Plugs were then stored in 1x TE buffer at 4°C until required.

4.5.4 Choosing cleavage sites and designing oligonucleotides

The desired fragment of DNA was decided to be roughly one recombination distance each side of the *BRCA2* gene (~1Mb total). Clustered-regularly interspaced short palindromic repeats-Cas9 (CRISPR-Cas9) cleavage of DNA has been used for DNA cleavage in previous studies (117). Given that CRISPR-Cas9 cleavage is known to be more accurate if within an exon this was chosen to be within the gene preceding this distance (*RXFP2*) and the gene following (*PDS5B*) this distance from the *BRCA2* gene. Thus, cleavage would create a roughly 1Mb fragment containing the *BRCA2* gene plus the surrounding gDNA (see figure 4.1).

Using the UCSC genome browser the *RXFP2* and *PDS5B* genes were examined for CRISPR target sites (118). Firstly this included all areas within this region with a protospacer adjacent motif (PAM) sequence (XGG). This was quickly narrowed using the UCSC genome browser CRISPR/Cas9 function allowing assessment of specificity, efficiency and off-target scores. The three best performing cleavage sites for each gene were selected and Alt-R CRISPR-Cas9 crRNA were designed for each of these sites (IDT).

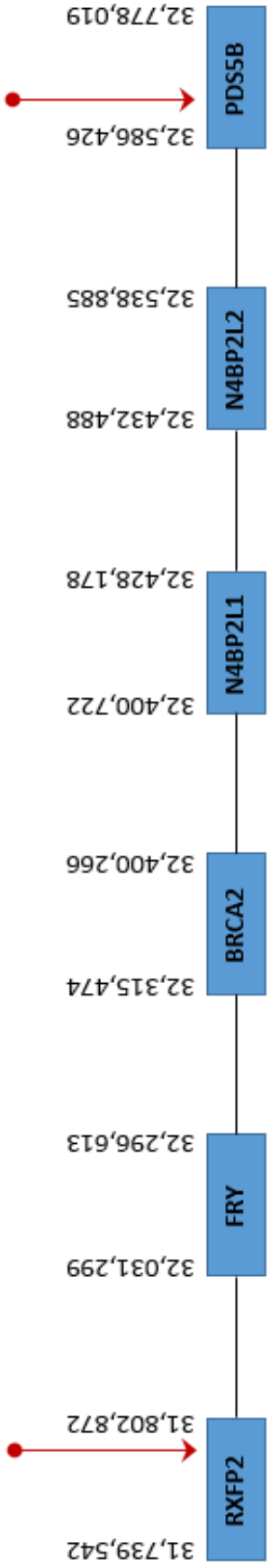


Figure 4.1: Diagrammatic representation of expected gDNA fragment. Numbers show start and end base numbers for each gene, red arrows show desired cleavage sites.

4.5.5 CRISPR-Cas9 in agarose cell plugs

The oligonucleotides were reconstituted as per manufacturer's instructions and stored as per manufacturer's instructions until required. Firstly, the crRNA oligonucleotides were assessed in pairs (one targeting *RXFP2* and one *PDS5B*) with a total of nine pairs compared. Once a pair of oligonucleotides was deemed best for target enrichment these oligonucleotides were used throughout for *BRCA2* mutation individual samples. Pairs of crRNA oligonucleotides were combined with tracrRNA (IDT, 1072533): 2µl each crRNA, 2µl tracrRNA, 12µl nuclease-free duplex buffer (IDT) and incubated at 95°C for 5 minutes. These hybridized RNA mixes were then incubated at room temperature for a further 10 minutes then aliquoted into smaller quantities to reduce freeze-thaw during use. Oligonucleotides were then stored at -20°C until required.

For plug digestion 0.33µl Cas9 nuclease (20µM, NEB, M0386T), 4µl Cas9 buffer (10x, NEB) and 30.67µl nuclease free water (NFW) per plug was combined with 5µl of the required hybridized RNA mix (crRNA and tracrRNA combined) and incubated at room temperature for 30 minutes. For controls 5µl NFW was added instead of the hybridized RNA mix. Agarose cell plugs were washed four times with 10ml 10mM Tris-HCl (pH8). Plugs were then washed with 500µl Cas9 buffer (1x, NEB) before being cut into two equal slices and submerged in 40µl per piece of the Cas9-RNA solution and incubated at 37°C for 2 hours. 3µl proteinase K (20mg/ml, Qiagen, 19131) was then added to each tube and incubated for a further 3 hours at 43°C. Plug pieces were then used immediately for pulsed field gel electrophoresis (PFGE).

4.5.6 PFGE

PFGE was performed using a CHEF DRIII (BioRad). Agarose gels were made at 1% with 1x TAE using SeaPlaque low melting point agarose (SLS, LZ50101). For optimisation of PFGE settings CHEF DNA size marker, *S.cerevisiae* Ladder (0.2-2.2Mb, BioRad, 1703605) and Lambda PFG ladder (48.5-1018Kb, NEB, N0341S) were

used. Once settings were optimised the agarose cell plugs were placed into the wells and sealed with 1% agarose.

Gels were post-stained with Sybr-safe (LifeSci, S33102) diluted to 1:10,000 in 1xTAE sufficient to cover the gel. This was then placed onto a rocker for 40 minutes protected from the light. Gels were then imaged using the appropriate filter with a ChemiDoc (BioRad). Images from the ChemiDoc were viewed and annotated using ImageLab software (BioRad).

4.5.7 DNA extraction from the agarose gel

The required agarose gel pieces were initially washed five times in 1x TE. These were then placed into a heat block at 70°C for 5 minutes to melt the agarose, then incubated at 43°C for 10 minutes. 2µl agarase (0.5U/ml, Thermofisher, EO0461) was then added to each piece and incubated at 43°C for 1.5 hours. 22µl 3M NaAc and 500µl cold isopropanol was then added to each tube which were then incubated at -20°C for 1h20. Tubes were then centrifuged for 30minutes at 14,000g. The supernatant was then discarded and the pellet resuspended in 500µl 75% ethanol and then centrifuged again at 14,000g for 5 minutes. The supernatant was again discarded and the pellets allowed to dry before resuspension in 20µl NFW. Samples were stored at -20°C until required.

4.5.8 qPCR for BRCA2 fragment

Custom designed *BRCA2* qPCR primers were used to detect the required fragment of gDNA (Forward sequence TCCCAAAGTGCTGGGATTAC, reverse sequence GGATCTCTTCCCGTCTCTATCT, 126bp fragment, IDT). Serial dilutions of human DNA (Promega, G1521) were used to optimise the qPCR primers.

Firstly, a qPCR master mix was created with 5µl Sybr green master mix (Roche, LightCycler 480 Sybr Green 1 kit, 4707516001), 1.5µl NFW and 0.5µl qPCR primers per sample to test. 7µl of this mix was then placed into the qPCR plate wells along with 3µl of the sample to be tested. NFW was used as a negative control and

purified human female DNA (G1521, Promega) diluted to 3.75ng/µl was used as a positive control. The plate was covered with a polyolefin StarSeal (Starlab, E2796-9795) and then centrifuged for 30 seconds at 2000rpm. The plate was then placed into a Lightcycler 480 (Roche) and the qPCR sequence run. The lightcycler protocol has a pre-incubation phase at 95°C (ramp rate 4.8°C/sec) then 45 amplification cycles (95°C for 10 seconds, 57.1°C for 10 seconds (calculated specific to the BRCA2 primers used), 72°C for 10 seconds), melting phase (95°C for 5 seconds, 65°C for 1 minute, 97°C to complete). The plate was then cooled to 40°C. Analysis was performed per sample using the LightCycler 480 software to assess the presence of the qPCR product. Amplification threshold was assessed for each DNA sample.

To compare enrichment for the desired DNA fragment, bands at the expected size were cut from the agarose gel and processed before qPCR. To ensure that there was no effect of differences in cell number within the agarose cell plugs a lane control was used for each sample using a visible DNA band from the same agarose gel lane. Oligonucleotides were assessed as having enriched for the desired fragment if the amplification threshold was reduced from the control to the expected fragment bands during qPCR. The oligonucleotide pair with the largest decrease in amplification threshold (therefore increase in enrichment) was chosen for the individual *BRCA2* sample CRISPR-Cas9.

4.5.9 Rolling-circle amplification (RCA)

RCA was performed by first circularising the DNA, then treating with exonucleases and then amplifying the DNA. To ligate the ends of the DNA 8.5µl of the sample was mixed with 8.5µl NFW, 2µl T4 DNA ligase buffer (10x, NEB) and 1µl T4 DNA ligase (NEB, E7546S) on ice. This was incubated at 16°C overnight. Heat inactivation at 65°C for 10 minutes was then performed. 0.9µl exonuclease 1 (NEB, M0293S) and 0.9µl exonuclease 3 (NEB, M0206S) was then added to remove any non-circularised DNA and incubated at 37°C for 90 minutes. Heat inactivation was again performed at 65°C for 10 minutes.

Amplification was then performed by adding Phi29 DNA polymerase (NEB M0269S), Phi29 buffer (NEB), dNTP (NEB, N0447S) and BSA (NEB). Quantities were optimised to allow amplification of the large fragment DNA with comparison of the RCA steps using qPCR.

4.5.10 DNA sequencing

DNA sequencing was performed using the MinION (Oxford Nanopore Technologies). DNA was processed prior to sequencing as per the SQK-LSK109 protocol (Oxford Nanopore Technologies) updated 23/05/18 using the Ligation Sequencing Kit (Oxford Nanopore Technologies) and Flow cell priming Kit (Oxford Nanopore Technologies). AMPure XP beads (Beckman Coulter, A63880) were used initially as per protocol for purification and clean-up prior to DNA sequencing.

Flow cells were checked prior to use by plugging the MinION into a computer with MiniKNOW software (Oxford Nanopore Technologies) and performing a flow cell test. Flow cells were only used if they had >800 available pores for sequencing.

For DNA repair and end-preparation 48µl gDNA, 7µl Ultra II end-prep reaction buffer (NEB), 3µl Ultra II end-prep enzyme mix (NEB, E7546) and 2µl NFW were added to a thin-walled PCR tube and incubated at 20°C for 5 minutes then 65°C for 5 minutes. The sample was then placed into a clean DNA Lo-Bind tube (Sigma, Z666548-250EA). AMPure XP beads were resuspended by vortexing and then 60µl added to the DNA mixture and mixed by pipetting slowly. This was then incubated on a rotator mixer for 5 minutes at room temperature. The sample was then centrifuged using a microcentrifuge and pelleted on a magnet. The supernatant was then removed. Keeping the tube on the magnet the beads were washed twice with 200µl newly prepared 70% ethanol without disturbing the pellet. The pellet was centrifuged again then placed back onto the magnet, any residual ethanol was removed and the pellet left to dry for 30 seconds. The tube was then removed from the magnet and the pellet resuspended in 61µl NFW and incubated for 2 minutes at

room temperature. The tube was returned to the magnet and the eluate transferred into a clean DNA Lo-Bind tube.

For adapter ligation and clean-up 25µl ligation buffer (LNB, Oxford Nanopore Technologies), 10µl NEBNext Quick T4 DNA ligase (NEB, M0202M) and 5µl adapter mix (AMX, Oxford Nanopore Technologies) was added to the gDNA and incubated at room temperature for 10 minutes. 40µl AMPure XP beads were added and mixed by pipetting before incubation on a rotator mixer for 5 minutes at room temperature. The sample was then centrifuged, pelleted on the magnet and the supernatant removed. Beads were then washed and resuspended in 250µl long-fragment buffer (LFB, Oxford Nanopore Technologies) before being returned to the magnet. The supernatant was removed. This step was repeated. The sample was then centrifuged, pelleted on the magnet, any residual supernatant was removed and allowed to dry for 30 seconds. The tube was then removed from the magnet and the pellet resuspended in 15µl elution buffer before incubation for 10 minutes at room temperature. The sample was returned to the magnet and the eluate removed to a clean DNA Lo-Bind tube.

The flow cell was primed as per manufacturer's instructions. For DNA library preparation 37.5µl sequencing buffer (SQB, Oxford Nanopore Technologies), 25.5µl loading beads (LB, Oxford Nanopore Technologies) and 12µl gDNA sample were mixed then added to the flow cell via the sample port in a dropwise fashion. Sequencing was then commenced. The MiniKNOW software allows real-time observation of the sequencing data. Sequencing was run for between 1 and 4 hours depending on the number and length of passed reads.

Read sequences were assessed using the U.S National Library of Medicine Blast software (119).

Flow cells were washed and stored using the manufacturer's instructions.

5. Results: Analysis of EUROPAC database and IMMPACT recruitment

5.1 Introduction

Since it was established in 1997 the EUROPAC team has continued to recruit individuals and families deemed to have an increased risk of pancreatic cancer above that of the general population. As described previously (see section 4.1: Materials & methods, Recruitment to EUROPAC) individuals with confirmed increased risk of pancreatic cancer are offered secondary screening with the aim of detecting cancer sufficiently early to allow curative management.

Description and analysis of the demographic details of individuals at risk not only allows for assessment of possible bias within the population who attend for secondary screening, but also may allow adaptation of the current EUROPAC pathways to improve uptake of secondary screening by such participants. This would benefit both at risk individuals and further pancreatic research, by implementing changes to increase the proportion of individuals attending for secondary screening and to allow an increase in the yield of pancreatic cancer detection without missing 'at risk' individuals.

5.2 Overview of population attending for secondary screening

Searching the EUROPAC Progeny database established that a total of 6272 individuals were identified as being suitable for secondary screening. 4958 were on the familial pancreatic cancer database (FPC) and 1314 on the hereditary pancreatitis database (HP). This signifies the group that has previously been invited to seek further information about and to attend for a clinical appointment to discuss secondary screening. There was a significant discrepancy in the demographics between HP and FPC screening suitable individuals. 55.67% of the FPC and 51.67% of the HP suitable individuals were female, this difference was statistically significant ($p=0.0106$). There was minimal age discrepancy between the

HP and FPC groups with the median age for the FPC group being 59 years (IQR 49-71) and for HP 57 years old (IQR 49-69).

A total of 378 individuals have enrolled and attend for secondary screening, the majority being FPC (345 FPC, 33 HP). Overall, 6% of suitable individuals attended for screening, with a larger proportion of the FPC than HP group suitable attending (7% of the FPC group compared with 2.5% HP, see table 5.2).

5.3 Data completeness

There were multiple problems with missing data specifically in the group not attending for secondary screening, as these individuals have had limited interaction with screening services. Information about gender and age was available for all of the individuals suitable for secondary screening. There were significant amounts of missing data for the rest of the other descriptive statistics, with data analysis performed with the remaining available data. (See Table 5.1 for description of data availability)

Table 5.1: Description of EUROPAC individual data availability

	Secondary screening (available/total)	**Not secondary screening (available/total)**	**Overall (available/total)**
Distance from centre (Km)	377/378 (99.7%)	959/5894 (16.3%)	1336/6272 (21.3%)
Demographic data	376/378 (99.5%)	949/5894 (16.1%)	1325/6272 (21.1%)
Smoking	370/378 (97.9%)	971/5894 (16.5%)	1341/6272 (21.4%)
Alcohol	368/378 (97.4%)	872/5894 (14.8%)	1240/6272 (19.8%)
Body mass index (BMI)	245/378 (64.8%)	650/5894 (11.0%)	895/6272 (14.3%)

The missing data were found overwhelmingly to be in the group who did not attend for secondary screening for pancreatic cancer with around 16% data availability for distance from centre, demographic, smoking and alcohol intake information.

5.4 Demographic data

The demographic details of both the screened and not screened population has been described in Table 5.2. The number of secondary screening suitable individuals was greater for the FPC than the HP group and a greater proportion of the FPC registered individuals compared with the HP attended for secondary screening (7% Vs 2.5%). There were also a larger proportion of females than males registered with the database, and a greater proportion of the secondary screened individuals were female.

Table 5.2: Description of demographic details of individuals registered with EUROPAC

	Secondary Screening	**Not Secondary Screening**	**Total**
FPC database	345 (7.0%)	4613 (93.0%)	4958
HP Database	33 (2.5%)	1281 (97.5%)	1314
Overall	378 (6.0%)	5894 (94.0%)	6272
Female	243 (64.3%)	3196 (54.2%)	3439 (54.8%)
Age (years, median (IQR)	57 (49-66)	59 (49-71)	58 (49-70)

A total of 1396 individuals had postcode details available (377 screened and 1019 non-screened). 60 of the non-screened individuals were excluded from the distance from centre and sociodemographic analysis due to being located outside of the UK. The closest centre to the individuals' post-code was assessed. Liverpool and London centres accounted for over 50% of the screened and non-screened individuals. Table 5.3 shows the closest screening centre for both the screened and non-screened individuals.

The majority of the individuals attending for secondary screening with postcode information available attended their closest screening centre (349/377, 92.6%).

Twenty-eight of the screened individuals attended for secondary screening at a site that was not their closest. 18/28 of these were in favour of the lead centre (Liverpool), whilst 3/28 attended screening at their local non-EUROPAC hospital. The remaining (7/28) were due to patient preference.

Table 5.3: Nearest EUROPAC centre for both screened and non-screened individuals

Nearest screening centre	**Secondary screening N (% TOTAL)**	**Not secondary screening N (% TOTAL)**	**Total at each centre N (%)**
London	103 (27.3)	298 (31.1)	401 (30.0)
Liverpool	120 (31.8)	222 (23.1)	342 (25.6)
Nottingham	45 (11.9)	120 (12.5)	165 (12.4)
Leeds	35 (9.3)	70 (7.3)	105 (7.9)
Bristol	16 (4.2)	85 (8.8)	101 (7.6)
Southampton	27 (7.2)	50 (5.2)	77 (5.8)
Glasgow	18 (4.8)	50 (5.2)	68 (5.1)
Newcastle	4 (1.1)	38 (4.0)	42 (3.1)
Morriston	9 (2.4)	26 (2.7)	35 (2.6)
TOTAL	377	959	1336

5.5 Comparison of EUROPAC population: screened versus non-screened

5.5.1 Self-reported gender

Statistical analysis of the population attending for secondary screening compared with those not attending highlighted a significant difference in the self-reported gender of the individuals. 243 of the screened (64.3%) compared with 3196 (54.2%) of the unscreened population reported as female. The difference in the proportion of female individuals within the groups was found to be statistically significant ($p = 0.0001$). When subdivided into FPC and HP the percentage of women in the

screened group remains roughly constant for the FPC (64.9%) but reduces for the HP (57.6%).

Overall, 4.77% of the males compared with 7.07% of the females suitable for secondary screening attended (p=0.0001). In the HP group this reduces to 2.2% of the males and 2.8% of the females and this difference is no longer significant (p=0.5974).

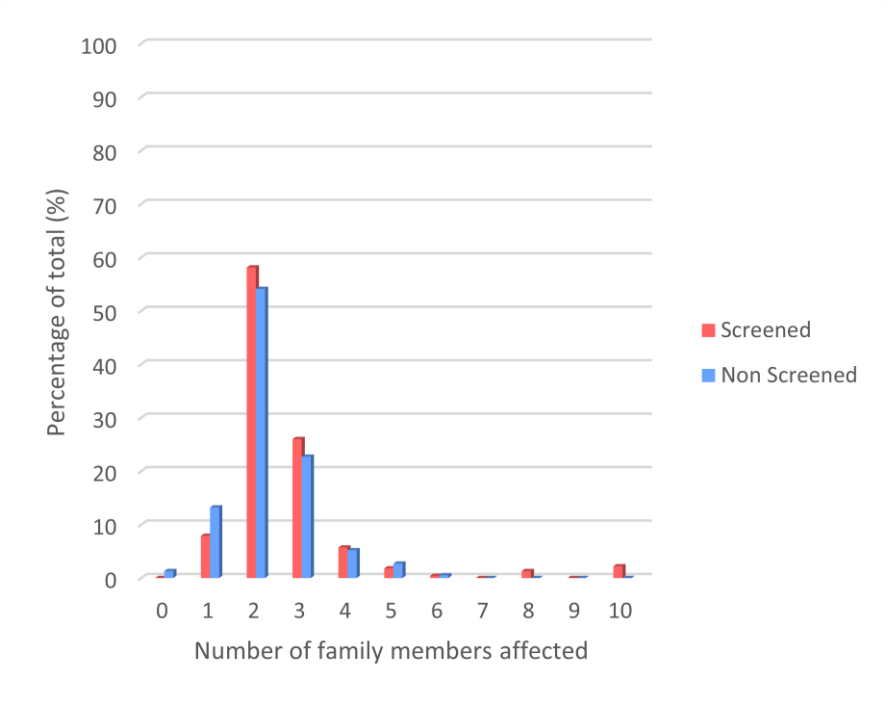
5.5.2 Number of relatives affected

The majority of both the screened and non-screened groups had 2 relatives with confirmed pancreatic cancer diagnoses, as expected due to the secondary screening criteria. 128 of the group attending for screening (33.9%) and 1868 (31.7%) of those not attending for screening were found to have two or more family members with a confirmed diagnosis of pancreatic cancer. This could be used as a surrogate marker for risk awareness amongst the groups however, this difference was found to not be statistically significant (p=0.3008). The distribution of the number of affected family members was found to be similar in both the screened and non-screened groups. Figure 5.1 and Table 5.4 show the distribution of the number of family members affected with pancreatic cancer comparing the screened to the non-screened population.

Table 5.4: Distribution of family members with confirmed diagnoses of pancreatic cancer

Family members affected with PC	**Screened (% of screened)**	**Non-screened (% of non-screened)**	**Total (% of total)**
0	0	42 (1.3%)	42 (1.2%)
1	18 (7.9%)	423 (13.2%)	441 (12.8%)
2	132 (58.1%)	1738 (54.1%)	1870 (54.3%)
3	59 (26.0%)	731 (22.7%)	790 (23.0%)
4	13 (5.7%)	168 (5.2%)	181 (5.3%)
5	4 (1.8%)	88 (2.7%)	92 (2.7%)
6	1 (0.4%)	17 (0.5%)	18 (0.5%)
8	3 (1.3%)	0	3 (0.1%)
10	5 (2.2%)	0	5 (0.1%)
Total	227	3215	3442

Figure 5.1: Percentage distribution of family members with confirmed diagnoses of pancreatic cancer between screened and non-screened population



5.5.3 Confirmed pathogenic mutations

Another potential surrogate marker for risk awareness amongst the groups is the presence or absence of mutations found to increase the risk of pancreatic cancer (see Table 5.5 for subdivisions of the confirmed pathogenic mutations). Overall, 269 (4.56%) of the non-screened and 47 (12.43%) of the screened population had a confirmed pathogenic mutation. This can be subdivided into HP and FPC. 39 (0.85%) of the FPC non-screened and 16 (4.64%) of the FPC screened had a mutation. Within the HP group there is a much more demonstrable difference with 230 (17.95%) of the non-screened and 31 (93.94%) of the screened patient having a mutation. Individuals with mutations both overall and from the individual FPC and HP groups are more likely to attend for secondary screening ($p<0.0001$). Table 5.5 describes the pathogenic mutations which are present in the above groups. The majority of FPC mutations were *BRCA2* (HBOC) or *P16* (FAMMM), which overall contributed to 87.27% of the FPC mutations. There was a higher proportion of *BRCA2* mutations found in the screened compared with the non-screened populations (43.75% Vs 38.46%). The majority of the HP screened group had a pathogenic mutation (93.94%) and a large proportion of these (93.54%) were *PRSS1* mutations, with cystic fibrosis mutations (*CFTR*) making up a much smaller proportion (9.68%). CFTR mutations are included into the HP group due to the large number of presentations with pancreatitis rather than this being true HP. The relatively low screening uptake in these individuals suggests that individuals with CFTR mutations understand this is not an HP mutation. A much smaller proportion of the HP non-screened group had pathogenic mutations (17.95%).

Table 5.5: Number and description of gene mutations found overall, subdivided into HP and FPC groups.

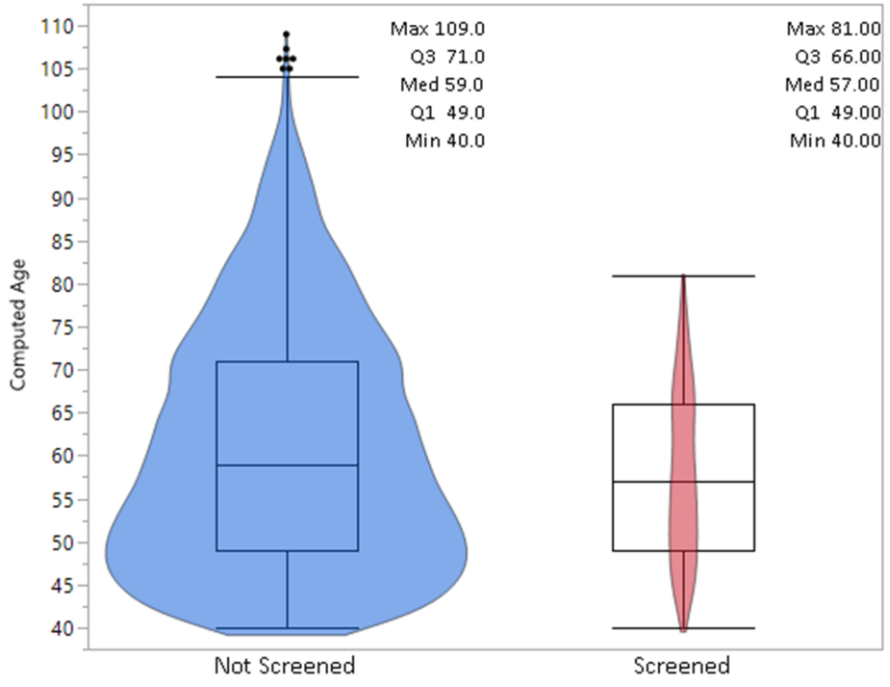
Database	Screened	Individuals with mutation (% of Total)	Mutation Type (% of Total)	Number with mutation (% of Mutations in group)
FPC	Yes	16 (4.64%)	*CDKN2A* (1.74%)	6 (37.5%)
			BRCA2 (2.03%)	7 (43.75%)
			HNPCC (0.58%)	2 (12.5%)
			CFTR (0.29%)	1 (6.25%)
FPC	No	39 (0.85%)	*CDKN2A* (0.44%)	20 (51.28%)
			BRCA2 (0.33%)	15 (38.46%)
			PJS (0.07%)	3 (7.69%)
			HNPCC (0.02%)	1 (2.56%)
HP	Yes	31 (93.94%)*	*PRSS1* (87.87%)	29 (93.54%)
			CFTR (9.09%)	3 (9.68%)
HP	No	230 (17.95%)	*PRSS1* (15.92%)	204 (88.70%)
			CFTR (2.03%)	26 (11.30%)

* One individual in this group had both *PRSS1* and *CFTR* mutations

5.5.4 Age

The median age of the individuals attending for screening was marginally younger than those who chose not to participate (57 years (IQR 49-66) and 59 years (IQR 49-71) respectively). There was a significant difference in age distribution between the two groups ($p<0.001$, see Figure 5.2)). The difference in sample size between the groups made further statistical analysis inappropriate.

Figure 5.2: Violin plot showing the age distribution of the screened and non-screened groups. Each plot size is relative to the number of individuals in each group (median and IQR are shown)



5.5.5 Body mass index (BMI)

It was noted that a greater proportion of the non-screened individuals were overweight, obese or morbidly obese compared to the screened individuals (50.16% Vs 46.12%) however, this was not statistically significant (P=0.1596) (See Figure 5.3). When the FPC data was analysed independently this relationship was found to be more pronounced however, remained not statistically significant (p=0.0597).

5.5.6 Tobacco smoking and alcohol intake

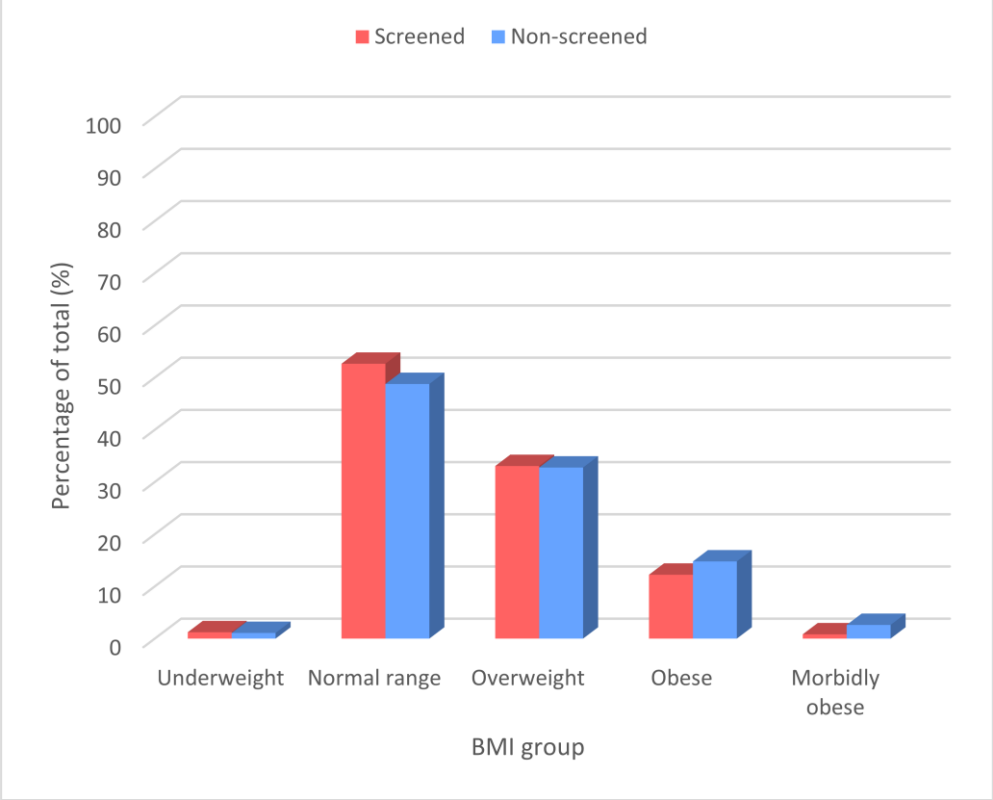
27 of the 370 (7.3%) individuals with smoking data in the screened group were found to currently smoke cigarettes. This is compared with 191 of the 971 (19.7%) in the non-screened group. This difference was found to be statistically significant (p<0.001). When subdivided into HP and FPC this difference remained, with 98 (18.85%) of the non-screened FPC compared with 15 of the screened FPC (6.85%) currently smoking. The HP group were more likely to currently smoke than the FPC and again the non-screened (n=86, 28.29%) were more likely than the screened (n=5, 15.63%) to smoke.

Information collected about the alcohol intake of individuals was found to show a statistically significant difference between the groups, with 312 (84.8%) of the screened compared with 629 (72.1%) of the non-screened group reporting regular alcohol intake (p<0.001 see Table 6). 805 (83.25%) of the FPC compared with 136 (49.82%) of the HP reported an ongoing regular alcohol intake (p>0.0001).

Table 5.6: Information about the self-reported alcohol intake for the screened and non-screened individuals

Regular alcohol intake	**Screened N (%)**	**Non-screened N (%)**	**Total N (%)**
None	27 (7.3%)	137 (15.7%)	164 (13.2%)
Previous	29 (7.9%)	106 (12.2%)	135 (10.9%)
Yes	312 (84.8%)	629 (72.1%)	941 (75.9%)

Figure 5.3: Percentage distribution of individuals BMI in screened and non-screened population



5.5.7 Diagnosis of diabetes

A diagnosis of diabetes was unsurprisingly more common for individuals in the HP group than the FPC, due to the effects of chronic pancreatitis. Overall, the presence of diabetes was more common in the screened compared with the non-screened group (7.67% screened Vs 5.40% non-screened). This was not statistically significant ($p = 0.0635$). When this was subdivided into FPC and HP however, the HP group did show a significant difference between the screened and non-screened groups ($p<0.0001$, 57.58% screened and 13.11% non-screened individuals with diagnosis of diabetes). This difference may reflect a difference in severity of HP between the screened and non-screened groups. The FPC group did not have a statistically significant difference ($p=0.8744$, 2.90% screened Vs 3.25% non-screened) between screened and non-screened individuals with a diabetes diagnosis.

5.5.8 Sociodemographic score

Postcode analysis was used to create a sociodemographic score, with a score of 1 being the most and 4 being the least affluent. A larger proportion of the screened individuals were scored in the two most affluent groups (73.9% screened Vs 66.6% non-screened). There was found to be a statistically significant difference between the sociodemographic scores of the individuals in the screened when compared to the non-screened groups ($p=0.0057$, figure 5.4). This persisted when the HP and FPC groups were separated. It was also noted that there was a statistically significant difference between the sociodemographic score in the FPC and HP groups, with more individuals being more affluent in the FPC group (70.84% FPC Vs 56.74% HP, $p<0.0001$).

5.5.9 Distance to screening centre

The median distance from the nearest screening centre was 47.2Km (IQR 20.9-84.5 Km) in the non-screened and 55.7 Km (IQR 21.9-95.0 Km) in the screened groups ($p<0.0001$). There was a statistically significant difference in the distribution of

distance from screening centre between the two groups (p=0.0452, Figure 5.5) however, due to large difference in sample size further data analysis of this distribution was not suitable. It was noted that the HP group (median 56.57Km, IQR 22.65-94.43Km) tended to have an increased distance to their screening centre than the FPC group (median 48.68Km, IQR 21.08-87.17Km) (p=0.0235). Nearest centre rather than screened centre was used for this analysis due to the small number of individuals not attending their closest centre.

Figure 5.4: Percentage distribution of sociodemographic score between screened and non-screened population

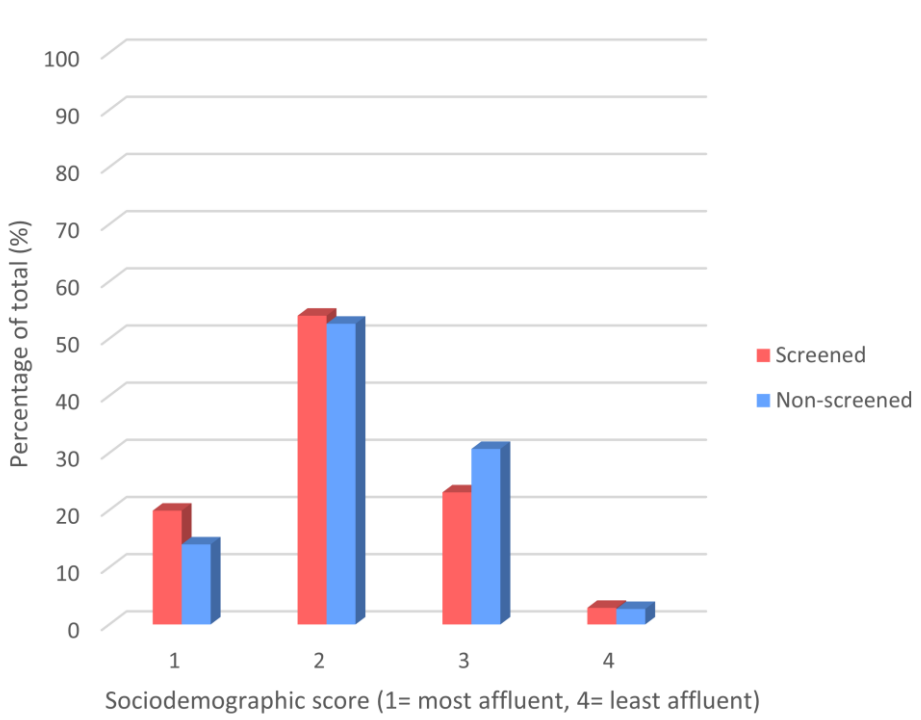
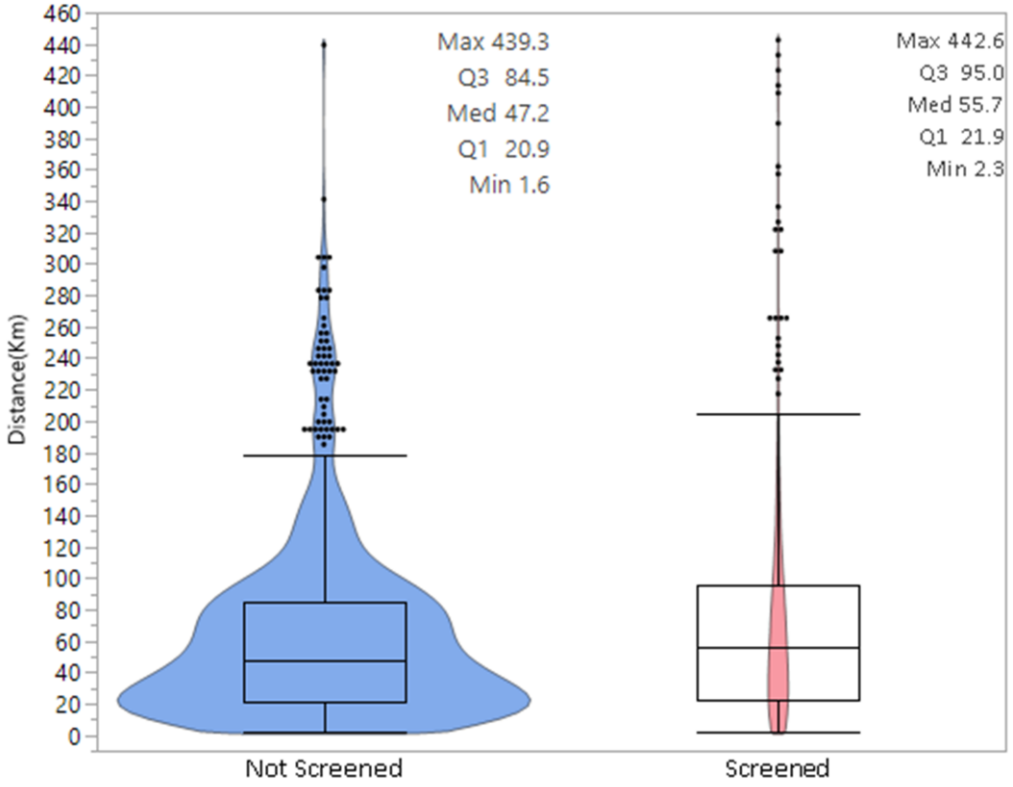


Figure 5.5: Violin plot showing distribution of distance from screening centre (Km) of the screened & non-screened groups. Each plot size is relative to the number of samples in each group (median and IQR are shown)



5.6 IMMPACT study recruitment

5.6.1 Recruitment criteria

The inclusion criteria for the IMMPACT study were based on initial research into pancreatic cancer risk which provides theoretical rather than empirical evidence. The IMMPACT study was created to allow detection of early pancreatic cancer biomarkers but also provide further data on cancer development in the population deemed 'at risk'. This will lead to refining of recruitment and screening strategies both for further studies and for clinical practice allowing us to overcome barriers to screening and also to increase screening yield. This would also mean a reduction in 'unnecessary' screening for those individuals who were originally deemed at increased risk for developing pancreatic cancer.

5.6.2 Individuals recruited

A letter and patient information sheet were sent to all of the FPC individuals attending secondary screening at Liverpool (120 individuals). Due to the recruitment criteria being changed to include the HP individuals at a later date this group were not contacted by letter. Individuals were also recruited from EUROPAC outpatient clinics directly. To date 85 individuals currently attending for secondary screening with EUROPAC have been recruited to the IMMPACT study. The majority of these (83/85) attend Liverpool for their secondary screening. 80 of the Liverpool secondary screening individuals were recruited (66.7% contacted individuals).

All of the IMMPACT samples collected were from the Liverpool secondary screening site and therefore the recruited individuals had to travel to that site for sample collection. The range of distance travelled was 3.7-344.4Km with a median of 33.7Km (IQR 12.6-61.15Km). The majority of samples were collected from individuals attending for their EUROPAC follow-up appointment or for their secondary screening appointment, therefore as expected the distance to site was reduced compared with the overall screened population however, this was not statistically significantly different between the screened and IMMPACT individuals (p=0.0984). A small number of individuals (n=2) attended an alternative site for

their secondary screening appointment but attended Liverpool specifically for the IMMPACT study.

The group recruited to the IMMPACT study includes 82 FPC and 3 HP (see Table 5.7 for further description of IMMPACT individuals).

Table 5.7: Recruited IMMPACT individuals described by recruitment criteria (fhx = family history of pancreatic cancer)

HP/FPC criteria	At risk Criteria met	Number of individuals
HP	PRSS1 mutation	3
FPC	BRCA1 mutation & fhx	1
	BRCA2 mutation & fhx	8
	HNPCC mutation & fhx	1
	FAMMM mutation & fhx	1
	PJS mutation & fhx	1
	2 close* relatives confirmed PDAC	57
	3 close* relatives confirmed PDAC	13

* close signifies first or second-degree relatives

58 of the recruited individuals were female (68%) and 27 were male, this was compared with 64.3% of the overall screened population and found not to be different (p=0.5301). At the time of writing the study was continuing to recruit and 21 individuals had provided a single sample, 22 two samples, 31 three samples, 10 four samples and one had completed the study with all 5 samples provided. The age range for the recruited individuals is 33-77, with a median age of 58 years old (IQR 51-64) compared with 57 years (IQR 49-66) for the overall screened group. There was not a statistically significant difference in ages between the groups (p=0.6943).

A smaller proportion of the IMMPACT recruited individuals compared with the overall screened population had a BMI which classed them as overweight, obese or morbidly obese (44.06% Vs 46.12%, see figure 5.6) however, this was not

statistically significant (p=0.8089). The IMMPACT recruited individuals were more likely to smoke compared with the overall screened group (11.76% Vs 7.3%), again this was not statistically significantly different between the two groups (p=0.1817). The IMMPACT individuals were less likely to currently drink alcohol (76.47% Vs 84.8%) however, this did not hold statistical significance between the groups (p=0.0794).

The overall screened population had a higher proportion of individuals in the two most affluent groups compared to the IMMPACT recruited population (73.9% screened Vs 71.76% IMMPACT). The distribution of individuals in each sociodemographic group was observed to be similar for the screened compared with the IMMPACT population and there was not a significant statistical difference between the two groups (p=0.5493) (see Figure 5.7).

The distribution of the number of relatives affected with pancreatic cancer was similar for both the overall screened and IMMPACT recruited individuals (see Figure 5.8).

Overall, this analysis suggests that the IMMPACT recruited individuals are representative of the overall population attending secondary screening, aside from the closer proximity to the Liverpool screening site.

To date two individuals were withdrawn from the IMMPACT study after recruitment, one due to patient preference, the other due to having a screening detected pancreatic abnormality which required treatment.

Figure 5.6: Proportion of overall screened population alongside IMMPACT recruited individuals' body mass index group (BMI)

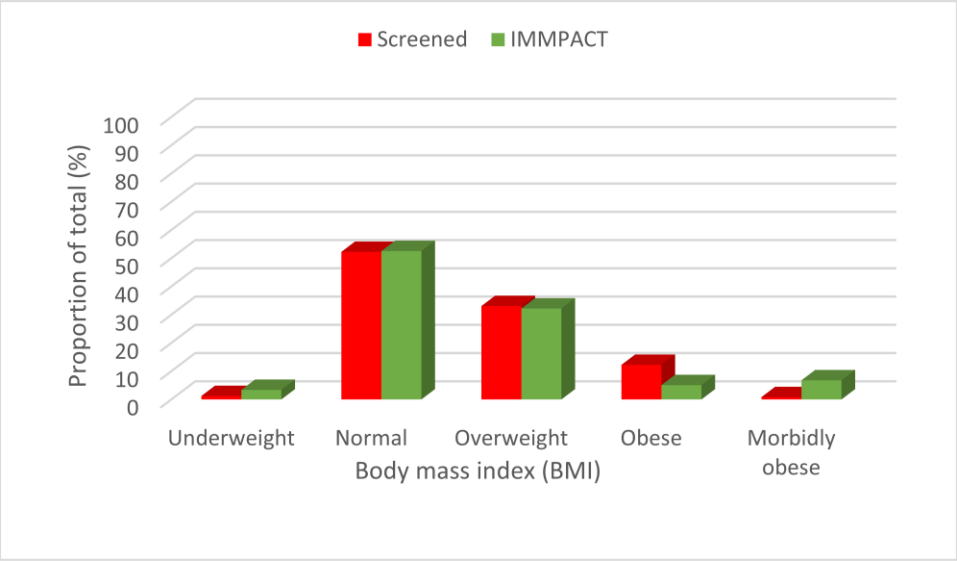


Figure 5.7: Proportion of individuals in the overall screened and IMMPACT individuals' sociodemographic score

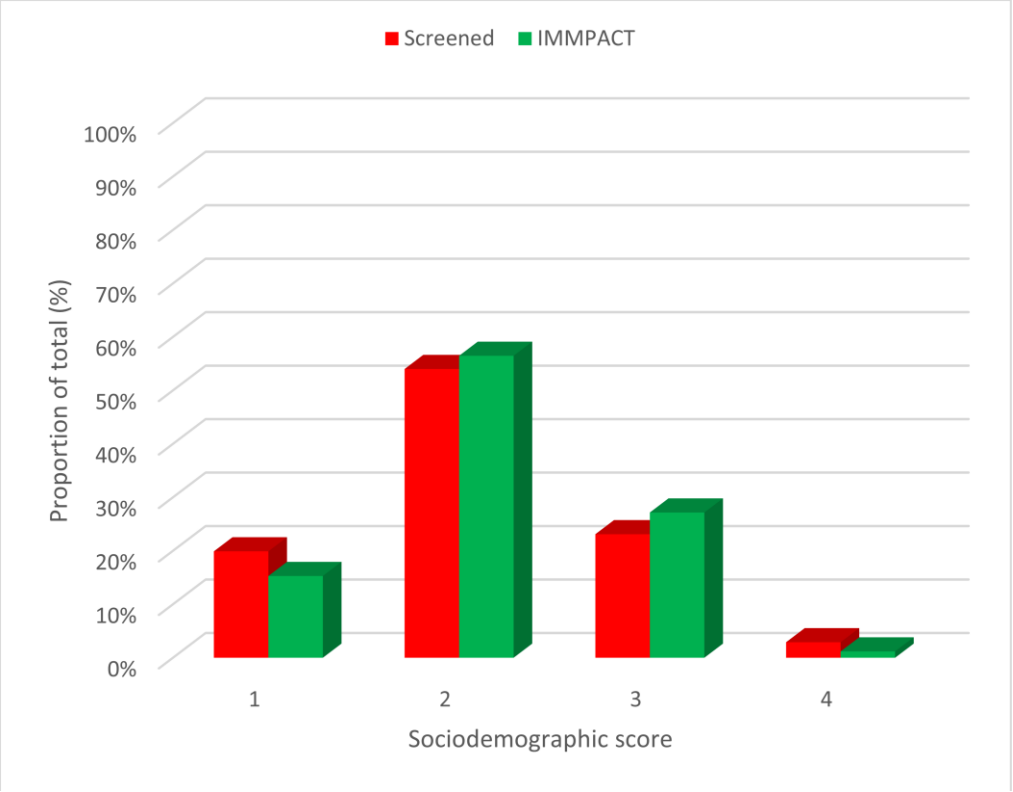
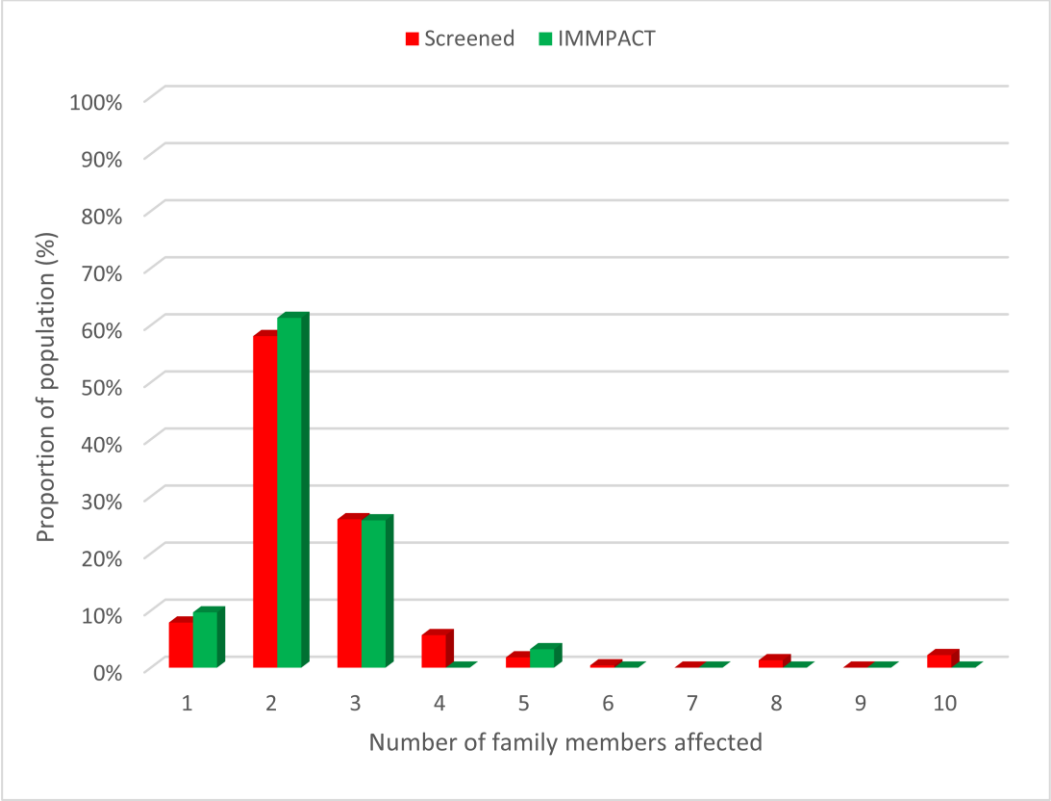


Figure 5.8: Proportion of the screened and IMMPACT recruited individuals with multiple family members affected with pancreatic cancer



5.7 Discussion

Analysis of the EUROPAC database shows that although over 6000 individuals were deemed as suitable for secondary screening based on current criteria but only 6% attended. A large proportion of missing data was also noted specifically in the non-screened group, which may lead to skewing of the data.

Overall, we found that individuals attending for secondary screening were more likely than the non-screened to be female, have a confirmed pathogenic mutation and be more affluent. Whilst there did appear to be a significant difference in age distribution and distance from screening centre travelled it was difficult to further assess the relationship between these parameters and attendance for secondary screening. It was also noted that individuals who attended for screening were less likely to smoke, perhaps linked by being more health conscious. This however, is contradicted by the screening group being more likely to currently drink alcohol, this may be due to a population difference in risk between drinking alcohol and smoking tobacco. It was also noted that the HP group that attended for screening were more likely to be diabetic. This suggests a possible link between disease severity, with severe or multiple episodes of pancreatitis being likely to lead to diabetes, and attendance for secondary screening. It is also possible that individuals may be aware of the link between diabetes and PDAC and therefore these individuals perceive their risk as higher, although this would be more likely in the FPC cohort, where no association was seen. These findings mirror those seen in other studies on screening for illness and will be discussed further in the discussion chapter.

Assessment of the IMMPACT recruited population showed no statistically significant differences between that and the overall screened population aside from being more likely to be closer to the Liverpool screening centre, as expected. This suggests that the IMMPACT recruited population is broadly representative of the overall screened population. There were fewer HP than FPC individuals recruited which

mirrors a change in the inclusion criteria during the study to include the HP group. A reasonable proportion of the Liverpool screened population were recruited (>60%).

The current criteria for secondary screening in those without a known mutation is based on family history however, some of these individuals will have relatives affected by chance rather than due to a primary causative mutation. Alongside this if the family does have a causative mutation there is a only 50% chance of this being passed to progeny however, the actual individuals' risk of pancreatic cancer is likely to be affected by more than this single mutation, such as environmental and other genetic factors (such as SNVs).

Relatively few of the EUROPAC 'at risk' individuals actually have a confirmed mutation, as expected due to the strict recruitment criteria, this was more common for the HP group. Interestingly there appears to be a dramatic difference in the actual risk of pancreatic cancer within these cancer syndrome families and to this end we will investigate this further. This also leads to the concern that some individuals who are at risk may not be screened, due to small numbers of family members. Alongside this some of the screened population may actually be low risk. Increasing attendance at screening for individuals deemed to be at high risk plus further identification of at-risk populations would be beneficial both to individuals and to researchers.

6. Results: BRCA2 family analysis

6.1 Introduction

Individuals and families are referred to EUROPAC due to their increased risk of pancreatic cancer. Amongst those regularly referred are those with mutations leading to cancer syndromes known to increase the risk of pancreatic cancer. This includes families with pathogenic *BRCA2* mutations and a relative with pancreatic cancer at any age. *BRCA2* pathogenic mutations are known to increase the risk of breast, ovarian and prostate cancer as well as pancreatic cancer. Although it is widely believed that all individuals with *BRCA2* mutations carry a similar risk of pancreatic cancer, a decision was made with EUROPAC to only screen individuals with a family history of PDAC, in line with NICE guidelines. The pathogenic *BRCA2* mutation positive population known to the EUROPAC team was analysed further to allow a more in-depth examination of cancer risk within this population to assess whether possible changes to the recruitment and screening criteria for these individuals would be beneficial.

6.2 EUROPAC families with *BRCA2* mutations

Overall a total of 63 families with known *BRCA2* mutations were found to be registered with EUROPAC as of March 2019. On further investigation two families were excluded from further analysis; one due to having a confirmed benign mutation (c. 6100C>T, p.Arg2034Cys) and the other having a variation of unknown significance (c.3326C>T, p.Ala1109Val). 61 families were included for further analysis, comprising of 830 individuals.

Out of 61 families with confirmed pathogenic *BRCA2* mutations 35 (57.38%) had information available to the EUROPAC team about which mutation they carried (see table 6.1). All of these mutations were confirmed pathogenic by multiple sources. Five of the specific mutations were shared by more than one family (see table 6.2). Nineteen of the mutations (65.52%) were frameshift and ten were single nucleotide

variants (SNV). Figure 6.1 (a copy of figure 1.1) shows the BRCA2 protein structure to allow collation of the amino acid changes within the protein.

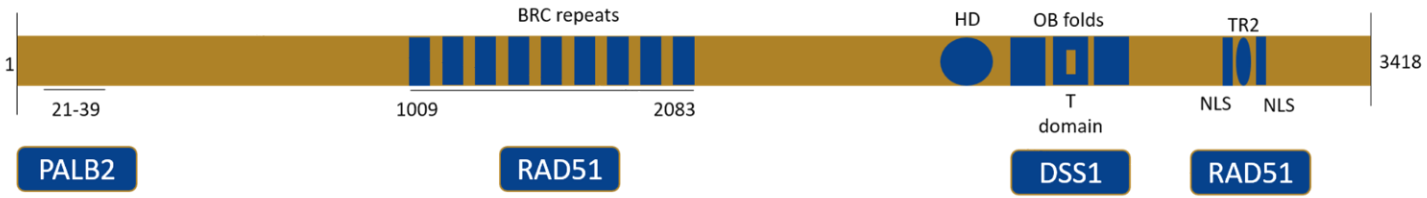
A stepwise analysis of survival within the families by affected *BRCA2* domain was performed which showed no statistically significant influence of the site of the mutation on individual survival from either pancreatic or breast cancer. We therefore sought to further stratify families to assess any other possible influences on survival.

Table 6.1: Mutations present in *BRCA2* positive EUROPAC families (Del = deletion, Dup = duplication, Indel = insertion and deletion)

Mutation	Protein change	Mutation type
c.314T>G	p.Leu105Ter	SNV
c.574_575delAT	p.Met192Valfs	Del→frameshift
c.631+2T>G	-	SNV
c.755_758delACAG	p.Asp252fs	Del→frameshift
c.1389_1390delAG	p.Val464fs	Del→frameshift
c.1929delG	p.Arg645fs	Del→frameshift
c.2330dupA	p.Asp777fs	Dup→frameshift
c.2409T>G	p.Tyr803Ter	SNV
c.2808_2811del	p.Ala938Profs	Del→frameshift
c.3847_3848delGT	p.Val1283Lysfs	Del→frameshift
c.4003G>T	p.Glu1335Ter	SNV
c.4859T>G	p.Leu1620Ter	SNV
c.4876_4877DelAA	p.Asn1626fs	Del→frameshift
c.5073dupA	p.Trp1692Metfs	Dup→frameshift
c.5351delA	p.Asn1784fs	Del→frameshift
c.5946delT**	p.Ser1982fs	Del →frameshift
c.6174delT*	p.Phe2058fs	Del→frameshift
c.6275_6276delTT*	p.Leu2092Profs	Del→frameshift
Del exons 14-16*	-	Del→frameshift
c.6502G>T	p.Gly2168Ter	SNV
c.6980del	p.Ser2326_Leu2327insTer	Del → frameshift
c.6996_7004delTGT ACCCTTins20	p.?	Indel → frameshift
c.7934del	p.Arg2645fs	Del→frameshift
c.8167G>C*	p.Asp2723His	SNV
c.8297delC	p.Thr2766fs	Del→frameshift
c.9117G>A	p.Pro3039=	SNV
c.9382C>T	p.Arg3128	SNV
c.9573G>A	p.Trp3191Ter	SNV
c.9699_9702del	p.Cys3233fs	Del→frameshift

*Mutation carried by two families, **Mutation carried by three families

Figure 6.1: Structure of the BRAC2 protein (adapted from Fradet-Turcotte *et al.*)
(Boxes show proteins known to bind at specific regions of BRCA2)



6.3 Categorising families with *BRCA2* mutations

Families were then grouped using the HBOC and pancreatic cancer (PC) criteria mentioned into those who fit either criterion, both or neither. This was confirmed by taking the opinion of three investigators independently (two clinicians and one geneticist). Independent scoring by the group agreed initially with 48 out of the 61 families (78.69%). The remaining 13 families were discussed within the clinical group and agreement was reached easily.

Twenty-two (36.07%) families were found to fit the HBOC criteria, 10 (16.39%) the criteria for PC, 20 (32.79%) fit the 'Neither' criteria and 9 (14.75%) families fit both the HBOC and pancreatic cancer criteria (an example of each family type can be seen in Figure 6.2)

The groups were then analysed based on the mutation data available.

Within the 35 families with *BRCA2* mutation locus available 12 (34.29%) were found to fit the HBOC criteria, 7 (20.0%) the criteria for pancreatic cancer, 9 (25.71%) fit neither criteria and 7 (20.0%) families fit both the HBOC and pancreatic cancer criteria.

In the groups 12/22 (54.55%) of the HBOC families had mutation data available and of these 9 (75.0%) were frameshift mutations. 9/20 (45.0%) of the families in neither group had mutation data available and 4 (45.44%) were frameshift mutations. 7/10 (70%) of the pancreatic cancer families had mutation data available and 6 (85.71%) were frameshift. Lastly, 7/9 (77.78%) of the families fitting both criteria had mutation data available and 4 (57.14%) were frameshift. Frameshift mutations were more commonly associated with families fitting the pancreatic cancer or HBOC criteria, whereas SNVs were more commonly associated with families fitting neither group (Figure 6.3). This finding was however, not statistically significant (p=0.2792).

There were five mutations shared between 11 families (see Table 6.2), again there was not seen to be any notable correlation between mutation locus and family

group however, the number of families with mutation data available made further analysis futile.

Table 6.2: A list of the pathogenic BRCA2 mutations shared by more than one family

Mutation	Mutation type	Family Group
Del exons 14-16	Frameshift	HBOC
		Neither
c.5946delT	Frameshift	PC
		HBOC
		HBOC
c.6174delT	Frameshift	PC
		Both
c.6275_6276delTT	Frameshift	HBOC
		HBOC
c.8167G>C	SNV	HBOC
		Neither

Figure 6.2: An example of a family tree of each *BRCA2* family type a. 'Neither' b. 'Both' c. PC d. HBOC

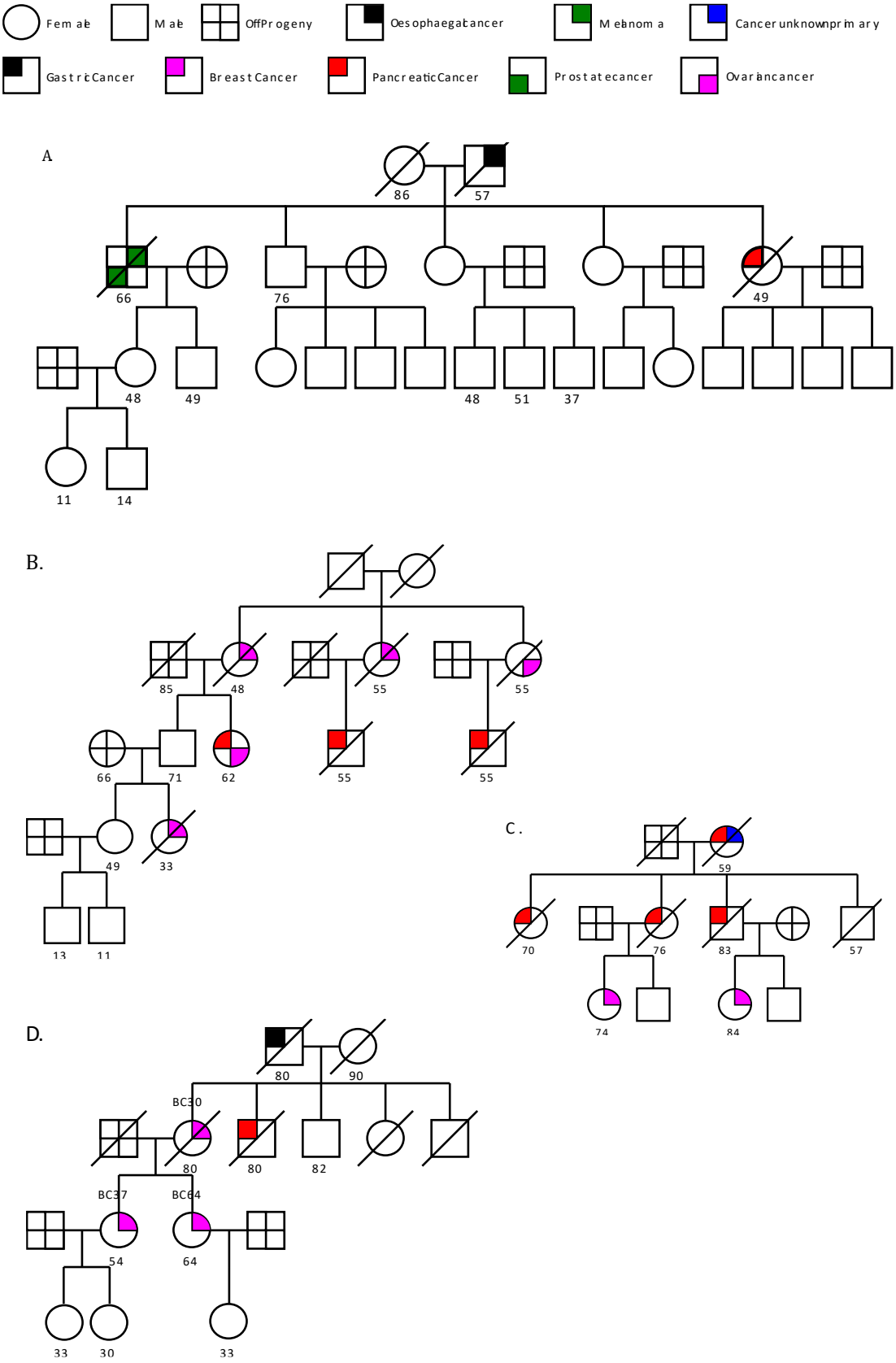
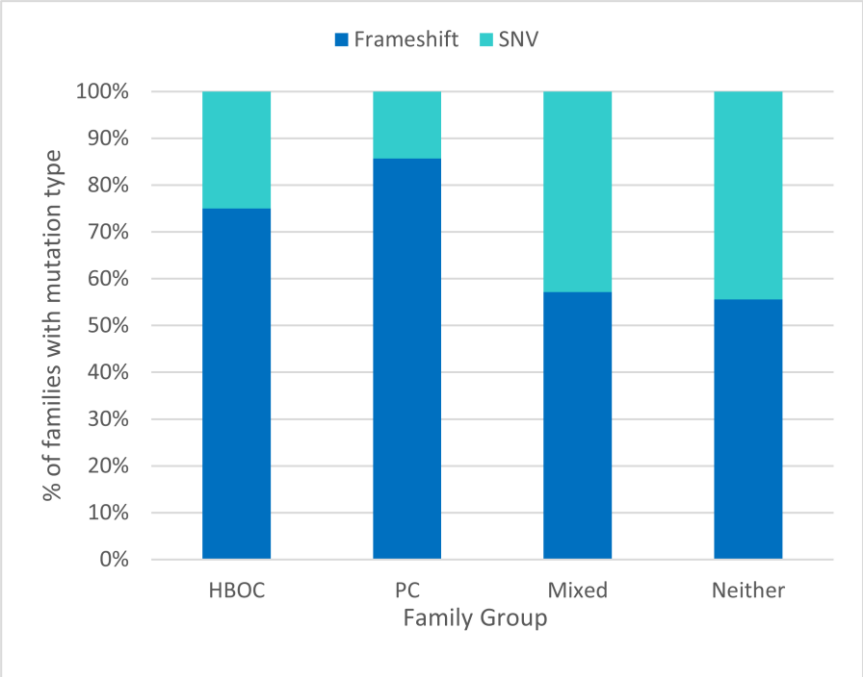


Figure 6.3: Proportion of each family grouping with frameshift mutations compared with SNVs



6.4 Survival analysis

Once the families were divided into PC, HBOC, Neither and Both categories the individuals within the families were analysed. Firstly, the individuals with a diagnosis of pancreatic cancer and/or breast cancer were assessed. To broadly describe the families the number and ages of the cancer diagnoses were calculated and analysed (Table 6.3). This suggested that in the individuals who suffered with pancreatic cancer or breast cancer there was no difference in the age at diagnosis between the families. There was found to be a difference in the average number of diagnoses of cancer per family (as expected) with the PC and 'Both' families having significantly more pancreatic cancer and the HBOC and 'Both' families having significantly more breast cancer. This suggests that whilst the 'at risk' individuals may still suffer with a cancer diagnosis at the same time point, there may be a different pattern and likelihood of being an 'at risk' individual within each family group.

Table 6.3: Overall analysis of PC, HBOC, 'Neither' and 'Both' category families with BRCA2 mutations

	PC families	**HBOC families**	**'Neither' families**	**'Both' families**	**P Value**
Pancreatic cancer diagnosis (average age, years)	61.96	61.95	59.47	60.64	0.9048
Breast cancer diagnosis (average age, years)	54	49.37	59.53	53.77	0.0979
Total number of pancreatic cancer diagnoses	33	24	19	25	-
Total number of breast cancer diagnoses	11	75	28	34	-
Average number of PC diagnoses per family	3.3	1.09	0.95	2.78	<0.001
Average number of BC diagnoses per family	1.1	3.41	1.4	3.78	<0.001

This was further echoed by survival analysis within the families. Firstly, survival analysis of the individuals with a diagnosis of pancreatic cancer showed no significant difference between the family groups for age at diagnosis (Figure 6.4a).

The Kaplan Meier survival curve shows multiple intersections between all of the family groups throughout its course. Secondly there was found to be no significant difference in the individuals' age at diagnosis of breast cancer between the family groups (Figure 6.4b). The Kaplan Meier curve of time to diagnosis of breast cancer again shows multiple intersection of all of the family types throughout. This reiterates that 'at risk' individuals tend to develop cancer at the same time point regardless of family type.

To analyse whether the risk profiles for pancreatic cancer and breast cancer were different between the family types all of the individuals in the BRCA2 families were analysed (a total of 830 individuals). In-depth analysis of all of the individuals showed that within the families there was a significant difference between the risk of pancreatic (Figure 6.5a) and breast cancer (Figure 6.5b). Diagnosis of pancreatic cancer was found to be more common in the pancreatic cancer group and least common in the 'neither' group ($p=0.0119$). Alongside this there was a higher risk of diagnosis of breast cancer in the HBOC group compared with the other family groups with pancreatic cancer families being least likely to suffer with breast cancer diagnoses ($p<0.001$). This suggests a significantly different cancer risk profile between the family groups.

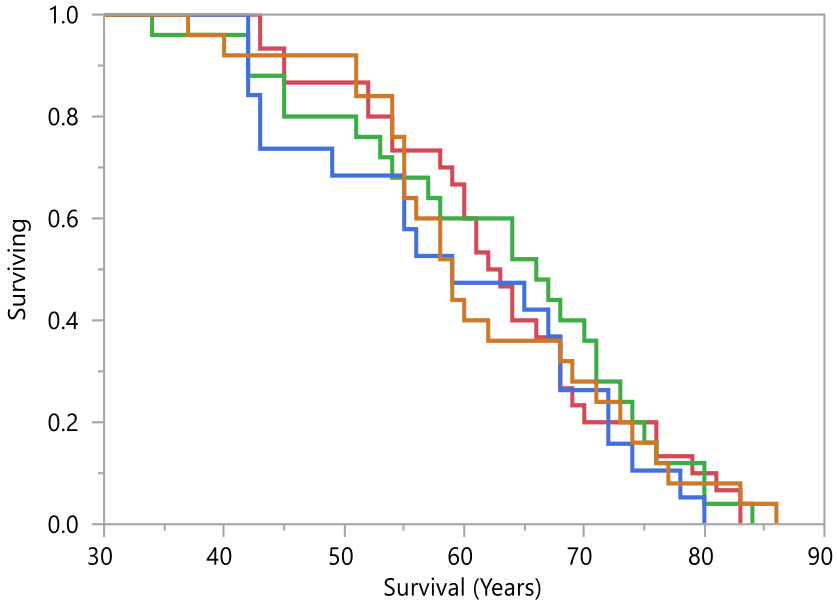
Overall, in the families with confirmed pathogenic BRCA2 mutations ovarian cancer and prostate cancer were much less common than pancreatic and breast cancer. 21 individuals throughout all of the BRCA2 families had suffered with a diagnosis of prostate cancer; 2 (9.25%) of these were in PDAC families, 9 each (42.86%) were in HBOC and 'neither' groups and 1 (4.7%) individual was in a 'both' family. Age at diagnosis was found to not be statically significantly different between the family groups ($p=0.7070$) however, small numbers made further analysis difficult. There was also found to be no statistically significant difference in risk of prostate cancer between the family groups ($p=0.0859$).

15 individuals suffered with ovarian cancer; 2 (13.33%) were in PDAC families, 4 (26.67%) each were in HBOC and 'both' families and 5 were in 'neither' families

(33.33%). Further analysis of the ages of diagnosis was difficult due to low numbers but analysis suggested no significant difference between the groups (p=0.4523). There was found to be no difference in risk of ovarian cancer between the family groups (p=0.7403).

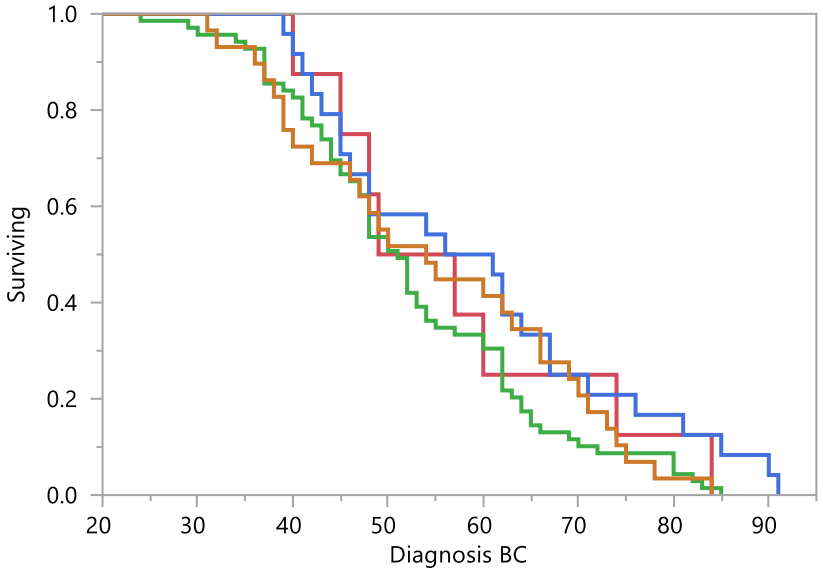
Figure 6.4: Kaplan-Meier plot for individual A. age at diagnosis of pancreatic cancer between families (p=0.8360) B. age at diagnosis of breast cancer (p=0.4575) (= PC family, HBOC family, 'neither' family, 'both' family). Table below shows at risk numbers for each time point.

A.



Time point (yrs)	0	20	40	60	80
PC	30	30	30	20	4
HBOC	25	25	25	16	3
Neither	19	19	19	10	1
Both	25	25	24	11	3

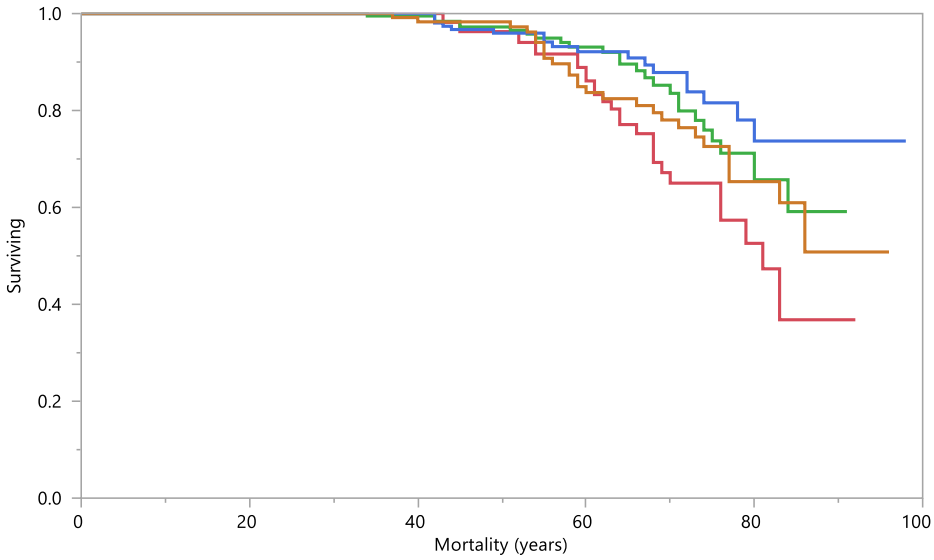
B.



Time point (yrs)	0	20	40	60	80
PC	8	8	8	3	8
HBOC	69	69	58	23	6
Neither	24	24	23	13	5
Both	29	29	22	13	2

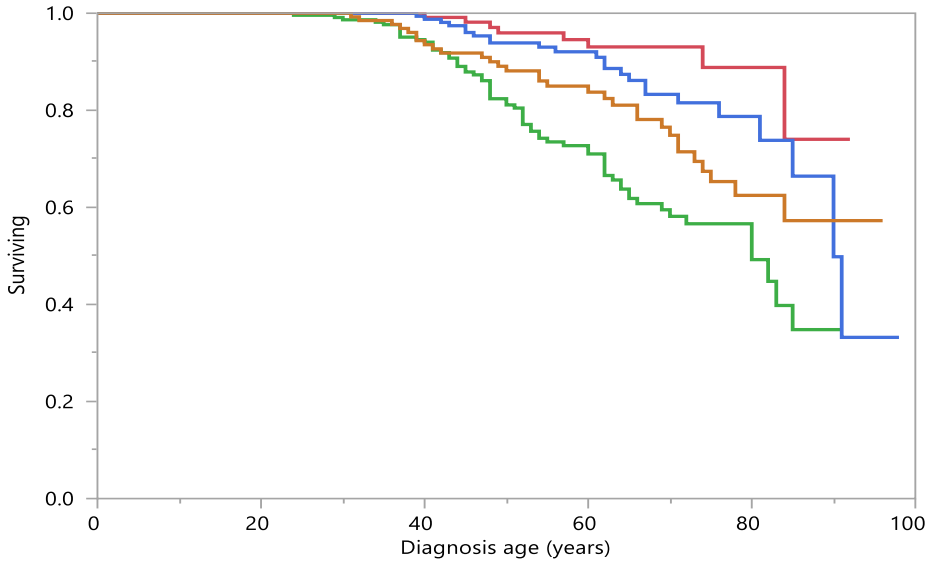
Figure 6.5: Kaplan-Meier plot for overall family A. Diagnosis of pancreatic cancer (p=0.0119) B. Diagnosis of breast cancer (p<0.001) (= PC family, HBOC family, 'neither' family, 'both' family) Table below shows at risk numbers for each time point.

A.



Time point (yrs)	0	20	40	60	80
PC	140	130	117	64	12
HBOC	291	254	189	91	26
Neither	237	207	160	84	18
Both	161	143	113	70	21

B.



Time point (yrs)	0	20	40	60	80
PC	140	129	116	63	11
HBOC	291	254	181	85	23
Neither	237	207	159	82	17
Both	161	143	112	68	19

6.5 Discussion

On investigation of the EUROPAC database 61 families were found to have pathogenic *BRCA2* mutations. These were then further subdivided into family groups using the HBOC and pancreatic criteria. Whilst the division of families in this way seems arbitrary, further analysis of the family groups suggested that the risk of pancreatic cancer and breast cancer was different between the groups. Individuals in PC families were found to have a higher risk of suffering with pancreatic cancer and individuals in 'Neither' families were found to have the lowest risk of suffering with pancreatic cancer. Most interestingly the 'Both' families had a significantly lower risk of PDAC than the PC families. Whilst this may seem like this difference is reflecting selection bias within the family groups if *BRCA2* individuals are modelled with a defined risk of breast cancer and PDAC then regardless of how the families are assigned, the risk of breast and pancreatic cancer is identical between the family groups.

There was not found to be any difference between the families for age of diagnosis of pancreatic cancer. This suggests that those individuals 'at risk' of pancreatic cancer suffer with the diagnosis at a similar age regardless of family risk however, individuals in certain *BRCA2* families have a different risk profile.

This is further reiterated by the analysis of breast cancer diagnoses within the family groups. Individuals in HBOC families were found to be more likely to suffer from breast cancer than those in other family groups. Pancreatic cancer families were least likely to suffer with breast cancer. There was found to be no difference in the age at diagnosis in each of the family groups. This highlights the different risk profile within the *BRCA2* family types.

There was no difference found with either prostate or ovarian cancer diagnoses or risk within the family types. This could reflect an accurate representation of the *BRCA2* families, but could also be due to the low numbers affected or reported with prostate and ovarian cancer in the families analysed.

Unfortunately, just over half of the *BRCA2* families had information available about what specific mutation the family carried. Further analysis was unable to link the mutation type with family group. There were only 11 families with mutations shared by more than one family. This suggested, again no link between mutation type and family group.

In summary there is evidence that there are different risk profiles between different families with *BRCA2* mutations and that the *BRCA2* mutation type does not adequately divide the family groups. This suggests that other factors may be linked to risk profile in families with pathogenic *BRCA2* mutations and these will be investigated further.

7. Results: Analysis of *BRCA2* family types

7.1 Introduction

It has been discussed previously that, at present, individuals with *BRCA2* mutations are expected to have the same increased risk of breast and pancreatic cancer. In the previous chapter it was noted that families with pathogenic *BRCA2* mutations can be grouped into those that fit criteria for HBOC, PC, both and Neither, and that there was a significant difference in the risk of developing breast and pancreatic cancer between these family groups. It was discussed that this required further analysis to evaluate the underlying reasons for this discrepancy. To this end the lifestyle and diabetes status of the individuals included with *BRCA2* mutations on the EUROPAC database will be further analysed to assess these possible confounding factors.

7.2 Lifestyle information availability

The 830 individuals previously explored for their cancer risk in *BRCA2* families were analysed further. Information about smoking status was available for 99 (11.93%) of the individuals and diabetes status was available for 471 (56.75%).

44 (44.44%) of the individuals with smoking data available were current or ex-smokers. Table 7.1 shows the data availability within each of the *BRCA2* family types. There was less smoking data availability for individuals in families which were categorized in the Neither group. There was no statistically significant difference found between the number of current or ex-smokers within the family groups (P=0.083)

Table 7.1: Smoking Data availability by *BRCA2* family type

Family type	**Total No Individuals**	**Individuals with data available (% of overall)**	**Current or Ex-smokers (% of data available)**
HBOC	291	42 (14.43)	13 (30.95)
PC	140	20 (14.29)	9 (45.0)
Both	154	24 (15.58)	14 (58.33)
Neither	245	13 (5.31)	8 (61.54)
Total	830	99 (11.93)	44 (44.44)

Twenty-two (4.67% of those with data available) individuals were found to have diabetes mellitus. There was no statistically significant difference in the number of individuals with diabetes within each family group (p=0.1188); HBOC 12/188 (6.38%), Neither 6/110 (5.45%), PC 4/82 (4.88%) and Both 0/91. There was again a reduction in the proportion of individuals with diabetes status recorded in the Neither group (44.90%) when compared with the PC (58.57%), Both (59.09%) and HBOC (64.60%) families.

7.3 Lifestyle information linked to cancer development

Smoking data was then linked to pancreatic and breast cancer development within the family groups. Table 7.2 shows the number of individuals who developed cancers within the *BRCA2* families compared to smoking status. There was no statistically significant difference between the development of pancreatic cancer or breast cancer by smoking status (P=0.40 pancreatic cancer, P=0.11 breast cancer)

Table 7.2: Cancer development compared to smoking status within the *BRCA2* families

Cancer type developed	Number of non-smokers (% of total)	Number of current/ ex-smokers (% of total)
None	38 (69.09)	20 (44.44)
Pancreatic	1 (1.82)	4 (8.89)
Breast	11 (20.0)	15 (33.33)
Ovarian	2 (3.64)	1 (2.22)
Prostate	1 (1.82)	1 (2.22)
Other	2 (3.64)	4 (8.89)
Total	55	45*

* Note one individual developed both pancreatic and breast cancer in the current/ex-smoker group

Diabetes data was also compared to cancer development in families with *BRCA2* mutations. Table 7.3 shows the number of individuals who developed cancers within the *BRCA2* families linked to diabetes status. There was no statistically significant difference between the development of pancreatic cancer or breast cancer by diabetes status (P=0.0771 pancreatic cancer, P=0.5359 breast cancer)

Table 7.3: Cancer development compared to diabetes status within the *BRCA2* families

Cancer type developed	Number of Individuals without DM (% of total)	Number of individuals with DM (% of total)
None	304 (66.52)	12 (52.17)
Pancreatic	45 (9.85)	4 (17.39)
Breast	65 (14.22)	4 (17.39)
Ovarian	7 (1.53)	0
Prostate	12 (2.63)	1 (4.35)
Other	24 (5.25)	2 (9.09)
Total	457*	23**

* Note one individual developed both pancreatic and ovarian cancer, five developed both pancreatic and breast cancer and two both breast and ovarian cancer in the no diabetes group

** Note one individual developed both pancreatic and breast cancer in the diabetes group

7.4 Survival analysis

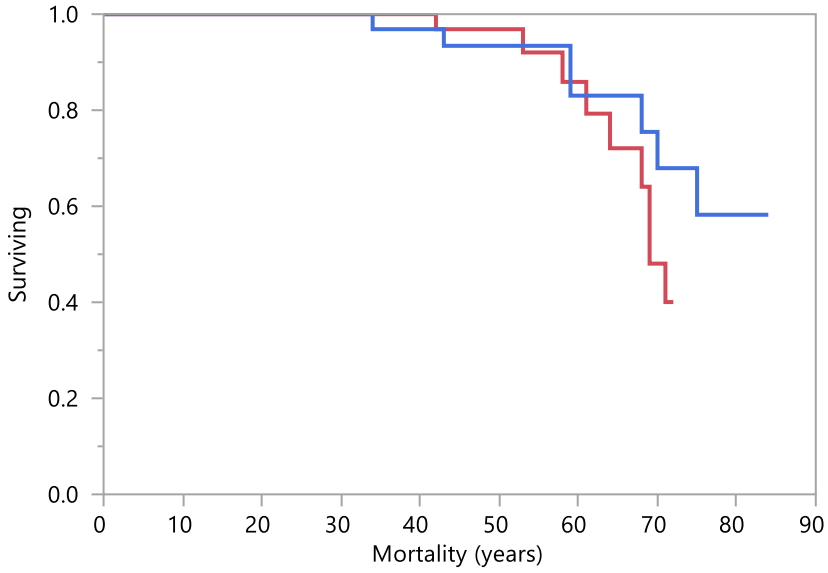
There was no difference in diagnosis of pancreatic cancer or development of breast cancer overall when censored for current or ex-smokers (see Figure 7.1).

There was no difference in overall diagnosis of pancreatic cancer or diagnosis of breast cancer between individuals with or without diabetes (see Figure 7.2).

The previous survival analysis by *BRCA2* family type and cancer type suggested that there was a significant difference in cancer risk between the family groups. This analysis was repeated with censoring for those individuals who were current or ex-smokers (see Figure 7.3). The difference between the family groups with diagnosis of pancreatic cancer was found to be more significant when censored for smoking history. The statistical significance between the *BRCA2* family types for development of breast cancer remained.

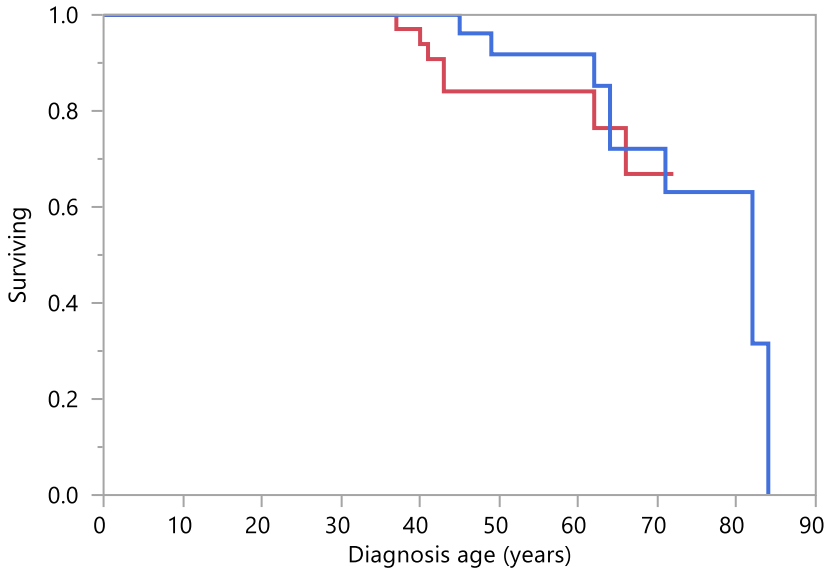
Figure 7.1: Survival analysis between non-smokers (red) and current/ex-smokers (blue) for A. Diagnosis of pancreatic cancer (P=0.24) B. Diagnosis of breast cancer (P=0.58). Tables show at risk numbers for each time point.

A.



Time point (yrs)	0	20	40	60	80
Non-smokers	55	47	33	15	2
Current/Ex-smokers	44	41	31	16	3

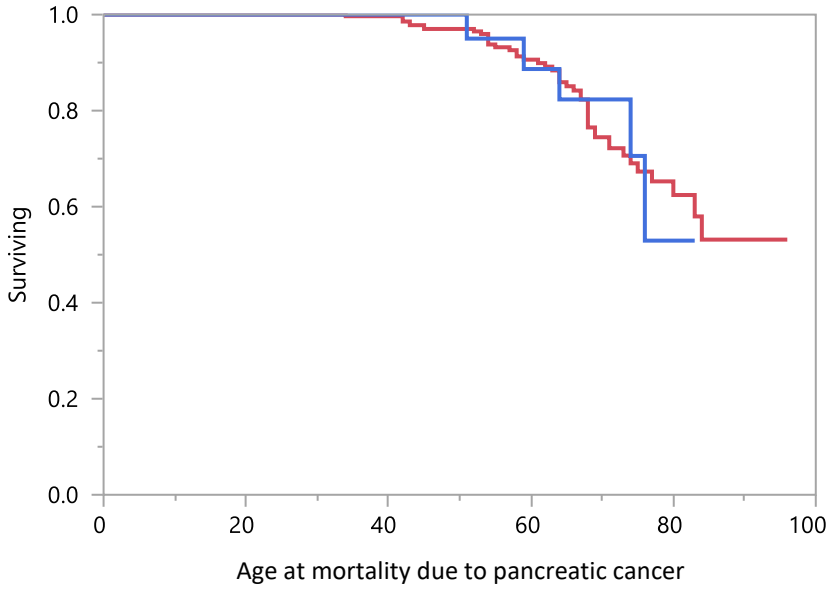
B.



Time point (yrs)	0	20	40	60	80
Non-smokers	55	47	31	13	2
Current/Ex-smokers	44	40	30	16	3

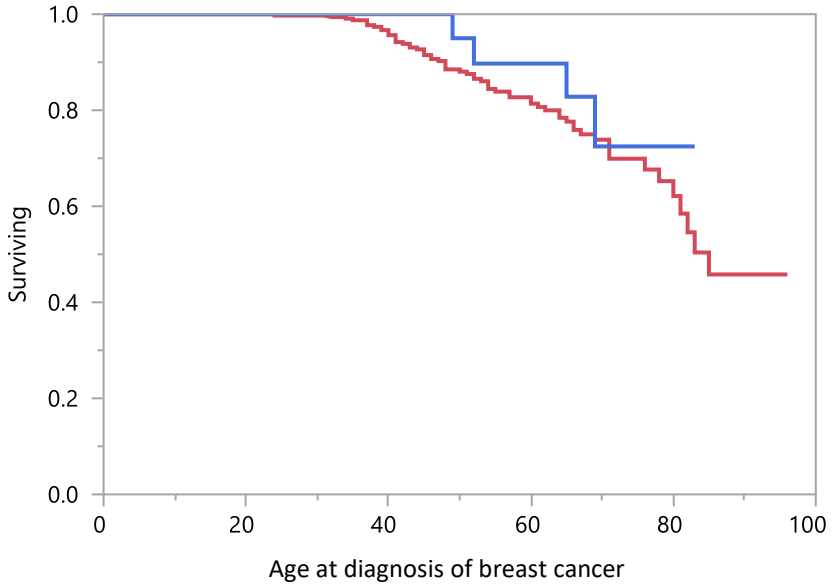
Figure 7.2: Survival analysis between non-diabetic individuals (red) and diabetics (blue) for A. Diagnosis of pancreatic cancer (P=0.8941) B. Diagnosis of breast cancer (P=0.4479)

A.



Time point (yrs)	0	20	40	60	80
No diabetes	449	379	284	129	23
Diabetes	22	22	22	15	2

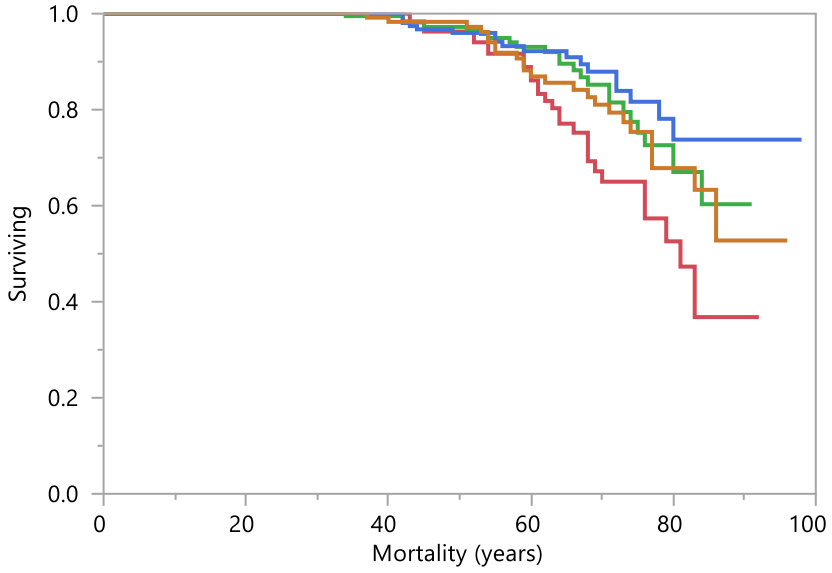
B.



Time point (yrs)	0	20	40	60	80
No diabetes	449	378	277	124	21
Diabetes	22	22	22	15	2

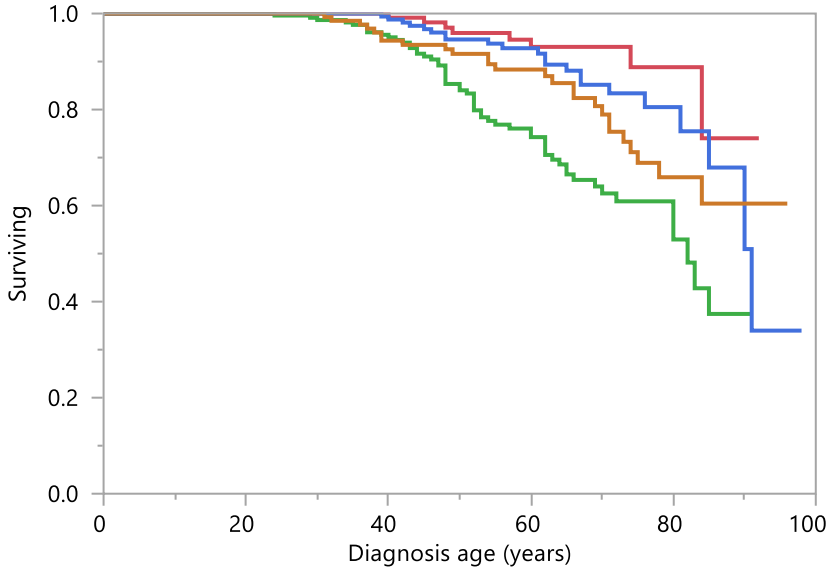
Figure 7.3: Survival analysis between the *BRCA2* family groups for current and ex-smokers A. Diagnosis of pancreatic cancer (p=0.0091) B. Diagnosis of breast cancer (p<0.0001) (= PC family, HBOC family, 'neither' family, 'both' family)

A.



Time point (yrs)	0	20	40	60	80
PC	140	130	117	64	12
HBOC	291	254	189	91	26
Neither	238	208	161	85	18
Both	161	143	113	70	21

B.



Time point (yrs)	0	20	40	60	80
PC	140	129	116	63	12
HBOC	291	254	181	85	23
Neither	238	208	160	83	17
Both	161	143	112	68	19

7.5 Discussion

A more in-depth analysis of the different *BRCA2* family groups has shown that the difference in cancer risk between the groups in the EUROPAC individuals cannot be explained by the difference in smoking history nor diabetes status of the individuals. This suggests that there may be other factors which alter an individuals' predisposition to cancer development within these families. One of the possible explanations for this could be a genetic haplotype coded by a group of SNVs which confer an altered inherited risk of cancer when coinciding with a *BRCA2* pathogenic mutation. This would explain the difference in risk of developing cancer within the *BRCA2* family groups and could aid in further risk stratification, prophylaxis and counselling of individuals with *BRCA2* mutations.

8. Results: Enrichment for BRCA2 chromosomal locus

8.1 Introduction

Previous chapters have highlighted the possibility of a haplotype altering the cancer predisposition of individuals with *BRCA2* mutations. In searching for the SNVs that make up such a haplotype it is important to know whether these are on the same strand of DNA as the mutation. Two possibilities exist for haplotypes defining different forms of *BRCA2* associated predisposition. It is possible that a linked haplotype exists which of necessity will segregate with the disease causing BRCA2 mutation, another possibility is that a *BRCA2* mutation has entered a kindred that has multiple unlinked risk modifying SNVs. In the former case we should see a haplotype associated with the *BRCA2* mutation that is different in the different family groups. In the latter SNV patterns associated with the different family groups will be associated with the non-mutant form of *BRCA2* and not the mutant form. Both possibilities can be tested by characterising the haplotype at the *BRCA2* locus, making this the most suitable locus to characterise. It seems most prudent to commence by searching the DNA around the *BRCA2* mutation itself and to be able to sequence these long DNA strands intact. Enrichment for the required genomic segment of roughly 935 kilobases (Kb) with the *BRCA2* gene centrally requires multiple stages. Firstly, formation of the cell-plugs allowed handling and enrichment of large fragment gDNA whilst reducing the sheering. Secondly treatment with CRISPR-CAS9 allowed targeted cleavage of the flanking gDNA segments. Lastly the fragments were separated using pulsed-field gel electrophoresis (PFGE). All of these separate steps required optimisation prior to use for the enrichment of the required fragment from individuals with pathogenic *BRCA2* mutations.

8.2 Optimising CRISPR-Cas9 cleavage for desired fragment

8.2.1 Designing the oligonucleotides

CRISPR-Cas9 cleavage to separate out the desired fragment was performed on agarose cell plugs, firstly using cultured R2797H (immortalised) fibroblasts for optimisation, then using PBMC samples from recruited individuals with *BRCA2* mutations. To optimise the cleavage process, target oligonucleotides were used which were specific to each end of the required fragment. Three oligonucleotides were tested for each end of the fragment, therefore six in total. These oligonucleotides were then tested in pairs (one for each end of the fragment) to assess for the highest yield of the desired fragment (9 pairs in total).

The specific oligonucleotides were chosen as the highest scoring for specificity and with the lowest number of off-target matches. Table 8.1 shows the chosen oligonucleotides with the protospacer adjacent motif (PAM) sequence, gene location and strand targeted.

Table 8.1: List of 6 chosen targeting oligonucleotides: PAM sequence, location and strand

	Sequence	**Gene location**	**PAM**	**Strand**
A	GGAGTCCTCTATCCACACAA	RXFP2 exon 18	TGG	-
B	ACTATCTACCAACCATGCTG	RXFP2 exon 18	AGG	+
C	TTTCAAAGTGGCTACTCCAC	RXFP2 exon 18	CGG	-
D	GGATCTAGGCTCTTATGCAG	PDS5B exon 21	AGG	-
E	GAGCCTAGATCCAAGCAACC	PDS5B exon 21	TGG	+
F	GTAGCTACCAAAGATTTCAA	PDS5B exon 21	AGG	-

8.2.2 Optimising PFGE settings for desired fragment size

Optimisation of PFGE separation was firstly performed with DNA ladders which include the required DNA size to monitor the degree of separation. The same percentage gel and type of agarose was used throughout. Optimisation of the PFGE CHEF DR II settings was performed initially using a lambda ladder (LL) and subsequently with a lambda and *S. Cerevisiae* (SC) DNA marker and required multiple attempts (see Table 8.2). The angle (120°) and cooling temperature (14°C) were consistent throughout PFGE optimisation. Gels were stained and imaged as mentioned previously.

Table 8.2: PFGE optimisation number with corresponding PFGE CHEF DR II settings

Number	Voltage	Start switch time (secs)	End switch time (secs)	Run time (hours)	Ladder(s)*	Corresponding figure(s)
1	6V/cm	60	120	24	LL & SC	
2	6V/cm	50	90	22	LL	1a, b
3	6V/cm	30	70	24	LL	2a-c
4	6V/cm	30	90	24	LL	3
5	6V/cm	30	100	28	LL	4
6	6V/cm	80	80	24	LL	5
7	6V/cm	80	80	24	LL & SC	6

* LL = lambda ladder, SC = *S.Cerevisiae* size marker

PFGE gel 1 showed evidence of speckling within the gel but no evidence of visible bands. This was thought initially to be due to the DNA ladders having run off the end of the gel however, there was also seen to be no visible staining in the wells, therefore this may have been due to poor staining. The settings were therefore adapted to allow a shorter run time and switch time and the next gel was stained during the run to ensure that the DNA was progressing through the gel.

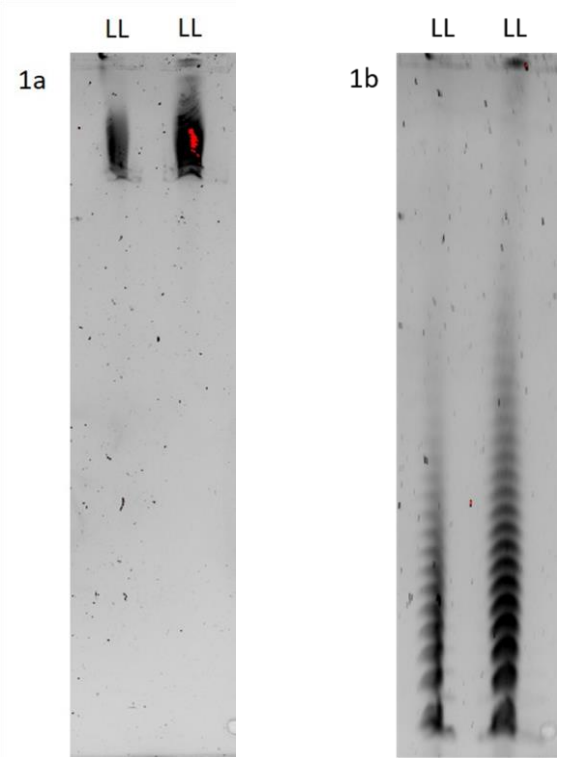
PFGE gel 2 was stopped at 3 hours to allow staining prior to completion (to ensure adequate staining). There was found to be visible smearing throughout the LL lanes (Fig. 8.1a). This gel was then run for a further 14 hours at which point 20 of the 21 bands were clearly visible (Fig. 8.1b). It was difficult to assess whether the 'missing' band was due to two bands being blurred at the top of the ladder or whether the lowest band had run-off the gel. Alongside this, whilst the ladder has separated reasonably, the higher bands are more difficult to define properly, which is where maximum separation is required for the desired fragment size.

PFGE gel 3 was stopped and stained at 4 hours, 18 hours and finally left to continue to 24 hours. Again, there was found to be visible smearing throughout the LL lanes (Fig. 8.2a). Re-staining at 18 hours showed the bands had not continued as far as gel 2 however, there appeared to be greater separation. It was also noted that there was more significant blurring of the bands than gel 2 (Fig. 8.2b). After re-staining at 24 hours there was found to be more significant separation between the lower bands than previously however, the upper bands were not separated and only 18 bands were visible.

PFGE gel 4 was left to run for the 24-hour programme. The smaller fragment bands were closer together than PFGE 3 however, the larger fragments showed much more separation and all 21 bands were visible (Fig. 8.3). There continued to be blurring of the bands however, this was improved from the previous gel, possibly due to the lack of intermediate staining. It seemed likely from this advancement that more time at higher switch time would provide better large fragment resolution.

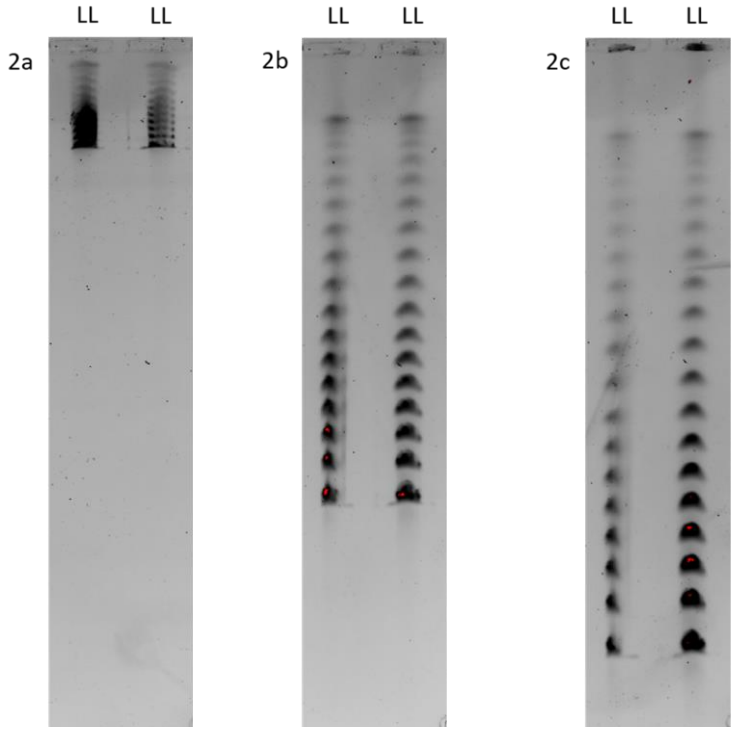
PFGE gel 5 was left to run for the 28-hour programme. There was similar separation to the previous gel but the heavier bands are more blurred (Fig. 8.4). It is likely that there is better separation of heavier bands but this was too difficult to see due to the poor resolution. The higher switch time provides better heavy-band resolution but the increased time at lower switch-times leads to significant blurring.

Figure 8.1: PFGE gel 2 settings with a.) staining at 3 hours and b.) staining at 17 hours (20/21 bands visible) (NOTE: Images shown are inverted to allow better visualisation of the DNA bands), LL = lambda ladder



PFGE settings:
Voltage: 6V/cm
Start switch time: 50 seconds
End switch time: 90 seconds
Run time: 22 hours

Figure 8.2: PFGE gel 3 settings with a.) staining at 4 hours; b.) staining at 18 hours and c.) staining at 24 hours (18/21 bands visible) (NOTE: Images shown are inverted to allow better visualisation of the DNA bands). LL = lambda ladder



PFGE settings:
Voltage: 6V/cm
Start switch time: 30 seconds
End switch time: 70 seconds
Run time: 24 hours

Figure 8.3: PFGE gel 4 settings with staining at 24 hours (21 visible bands)(NOTE: Images shown are inverted to allow better visualisation of the DNA bands)

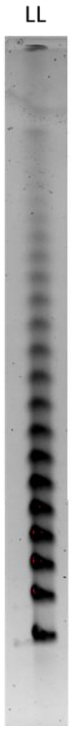


Figure 8.4: PFGE gel 5 settings with staining at 28 hours (19/21 visible bands)(NOTE: Images shown are inverted to allow better visualisation of the DNA bands)

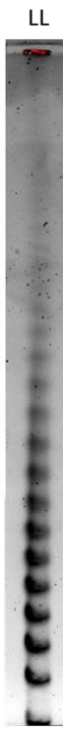
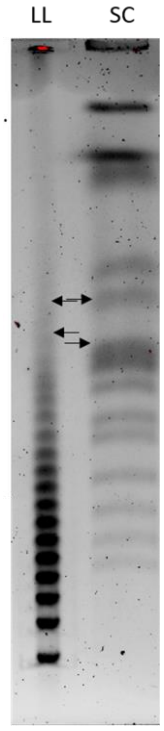


Figure 8.5: PFGE gel 6 settings with staining at 24 hours (21 visible bands)(NOTE: Images shown are inverted to allow better visualisation of the DNA bands)

Figure 8.6: PFGE gel 7 settings with staining at 24 hours. Arrows show the window for the desired weight fragment. (NOTE: Images shown are inverted to allow better visualisation of the DNA bands)



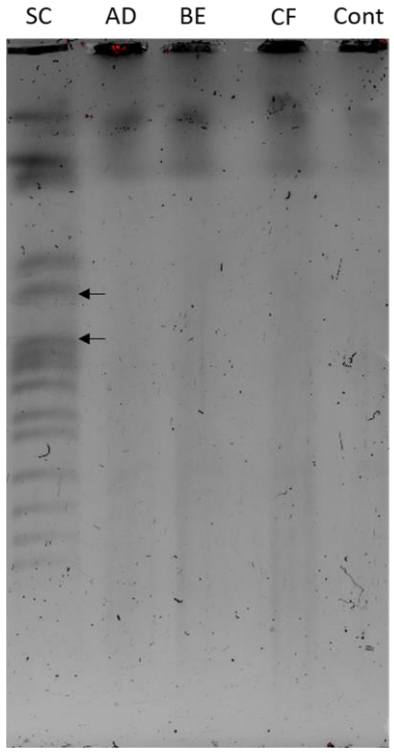
8.3 PFGE separation of fragment after CRISPR-Cas9 cleavage

Agarose cell plugs were treated to allow oligonucleotide-guided cleavage at the selected sites. These plugs were then processed using the PFGE settings optimised for the specific weight of fragment (see table 8.1 – settings 7). Figure 8.7 shows an example of a typical PFGE post-staining to visualise the ladder and size marker run after CRISPR-Cas9 cleavage. There was consistently found to be no clear band visible at the desired weight after cleavage, this was thought to be due to the relatively small amounts of large fragment gDNA. There was seen to be smearing throughout the lanes thought to be due to sheering. There was also found to be some visible bands at larger weight than the expected fragment (roughly 1.3-2.2Mb) which was found in all of the wells including the control (no CRISPR-Cas9 targeted cleavage), this was thought to possibly be due to mitochondrial DNA staining.

8.4 DNA Extraction from PFGE gel

The method for isopropanol DNA extraction from agarose gels is already well recognised and trusted, therefore did not require additional optimisation for large fragment DNA. There was a need to ensure that the correct portion of the gel underwent the extraction process – the aim being to minimise any impurities with the smallest gel fragment possible to extract as much DNA as possible. To this end, after cutting out one of the visible bands from a DNA size marker, a piece of the gel was cut in half from side-to-side. This was then re-imaged to assess whether the majority of the DNA was present in the top or bottom half of the gel. The majority of the stain remained in the top portion of the gel, with the bottom half containing little or no stain. To this end when further extraction for DNA was performed the bottom half of the gel was discarded.

Figure 8.7: PFGE example after CRISPR-Cas9 cleavage. No visible bands at desired weight (see arrows on SC size marker). SC = S. Cerevisiae size marker, AD, BE, CF = pairs of oligonucleotides used for CRISPR-Cas9 cleavage, Cont = control (no CRISPR-Cas9 treatment).



8.5 Optimisation of the oligonucleotide used for CRISPR-Cas9 cleavage

To ensure that the separated fragment was that required quantitative polymerase-chain reaction (qPCR) for *BRCA2* gDNA was performed using gDNA extracted with the required weight. The corresponding oligonucleotide visible band from the gel was also extracted and used as a control for each lane. It was decided this would be the most appropriate control for each qPCR as there was obviously visible DNA and this would account for any discrepancy in the DNA quantities used in each well.

Each oligonucleotide combination was tested alongside a positive and negative control. The oligonucleotide with the largest improvement in concentration of *BRCA2* (shown by a reduction in qPCR amplification threshold cycle) from the control fragment was chosen for the oligonucleotide for CRISPR-Cas9 treatment of patient samples (see Table 3). The positive and negative qPCR controls functioned as expected.

Table 8.3: Oligonucleotide guides for CRISPR-Cas9 cleavage amplification threshold for control versus expected fragment pieces of PFGE gel. The largest reduction in amplification threshold shows the best enrichment for *BRCA2*. (The oligonucleotides with improvement in the enrichment for BRCA2 are in green, those not enriched are in red)

Oligonucleotide pair	Control amplification threshold cycle	Fragment amplification threshold cycle	Difference
AD	26.15	24.26	-1.89
BE	24.21	22.73	-1.48
CF	27.46	22.17	-5.29
BD	26.57	25.68	-0.89
CE	29.97	44.94	14.97
AF	28.84	26.37	-2.47
CD	31.45	24.23	-7.22
AE	41.16	27.7	-13.46
BF	30.33	30.88	0.55

8.6 Discussion

This chapter has shown the success in optimising enrichment for a large gDNA fragment with the *BRCA2* gene at the centre. Firstly, PFGE was optimised using size markers to provide best separation at the size of DNA required. This included the assessment of multiple switch time changes. Throughout the procedure the other parameters were kept constant. It was noted that many of the gels were found to have speckling and unfortunately this was not rectified by renewing any buffers used, ensuring the gels were set in an environment that meant dust did not settle or by using a different batch of agarose. Despite this there did not seem to be any noted effect on the distance travelled between gels using the same PFGE settings.

Specific qPCR primers were designed to ensure that BRCA2 gDNA (rather than cDNA) was detected. There was a suggestion that post-staining of the gel with Sybr-safe to allow visualisation of the bands may lead to a change in qPCR result with Sybr-green however, this was found not to be the case.

Lastly, it was noted that PFGE after CRISPR-Cas9 separation of the fragment lead to smearing throughout the lanes and no visible band at the required fragment. This was largely due to the relatively small percentage of total DNA represented by the *BRCA2* locus, representing approximately 1 million out of 6 billion bases, so at best we would have expected a band more than 1000 times less bright than the total plug DNA, alongside this we have also noted fragmentation of the DNA. It is also likely that, even with treatment of cells within the agarose plugs and then adding them straight to PFGE, there would be an element of DNA shearing seen. These findings are consistent with similar experiments using large fragment DNA and will be discussed further in the main discussion section.

9. Results: Amplification, purification and sequencing of gDNA

9.1 Introduction

Prior to purification and sequencing of the desired gDNA fragment amplification of the segment was required. Due to low numbers of gDNA ends with the same amount of larger DNA strands, amplification was required prior to sequencing to ensure an adequate number of reads for the desired gDNA fragment. This is due to the sequencing device commencing from tethers attached to the end of double-stranded DNA. The amplification steps were required to be optimised and adapted for large fragment DNA.

Nanopore sequencing was chosen as a method to obtain long-reads intact so that the whole of the extracted segment of gDNA could be sequenced together, thus detecting whether any SNVs detected were on the same DNA strand as each other (allowing for formation of haplotypes) and the *BRCA2* mutation. Prior to sequencing, purification is important for any sequencing technique to reduce the number of spurious reads and, specifically for sequencing with nanopores, to reduce clogging and maintain function of the pores. Purification and sequencing preparation steps were adapted from the Oxford Nanopore Technology MinION protocols to ensure they were along with manufacturers guidance.

9.2 Optimising Rolling Circle Amplification of DNA fragment

Rolling circle amplification (RCA) is a well-known method of DNA amplification however, use with larger DNA fragments (>500Kb) is not well described. To optimise RCA for larger fragments, different amounts of substrates necessary for amplification were used, along with an increased time period for which the reaction was run. RCA requires ligation of the DNA ends (to form circular DNA), treatment

with exonucleases (to destroy non-circularised DNA) and amplification (using phi29 DNA polymerase). Throughout the optimisation steps the ligation and use of exonucleases were the same, it was the amplification steps that required adaptation.

9.2.1 Reaction 1

Firstly, the amplification steps were run according to a local, already used protocol. This involved adding 1µL phi29, 4µL deoxyribonucleotide triphosphate mix (dNTP), 0.2µL bovine serum albumin (BSA) and 2µL phi29 reaction buffer after ligation and exonuclease treatment. The reaction was left overnight at 30°C and then heat inactivated. Quantitative PCR (qPCR) was then used to assess whether there was amplification of the desired gDNA fragment. This showed that post-amplification samples had an increase in amplification threshold cycle for BRCA2 (therefore had a decrease in the desired product), when compared with their control samples (non-RCA samples) (See Table 9.1). Positive and negative controls performed as expected.

Table 9.1: qPCR results of reaction 1 in duplicate with pre-RCA sample of the same DNA used for control

	Sample 1	**Sample 2**
Pre-amplification amplification threshold cycle	25.89	25.69
Post-amplification amplification threshold cycle	29.83	28.64
Difference	3.94	2.95

It was thought that the increase in amplification threshold cycle (therefore the decrease in DNA) was due to dilution of the samples during RCA alongside the RCA not having the desired effect. This was thought to be due to lack of substrates required for RCA of such large DNA fragments along with more time required for amplification.

9.2.2 Reaction 2

Amplification was adapted to include multiple cycles. The first cycle was the same as for reaction 1. After this, two further overnight cycles were run with 1.2 µL phi29, 5 µL dNTP and 2.4 µL phi29 buffer added each time. This ensured that over a total of three cycles 3.4 µL phi29, 14 µL dNTP and 6.8 µL phi29 buffer was added. qPCR was again performed using the pre-amplification sample as a control (Table 9.2). Samples were run in duplicate. To ensure that the ligation and exonuclease treatment had been effective, samples without ligation but with exonuclease treatment were also run through the same RCA cycles and compared with qPCR.

Table 9.2: qPCR results of reaction 2 in duplicate with pre-RCA sample of the same DNA used for control, alongside a non-ligase treated sample

	Sample 1	**Sample 2**
Pre-RCA amplification threshold cycle	32.51	32.49
Post-RCA amplification threshold cycle	26.92	25.76
Difference	-5.59	-6.73
No ligase sample amplification threshold cycle	NA	NA

NA = No amplification

This shows a decrease in amplification threshold on qPCR by a mean of 6.16 cycles, meaning that there is considerably more *BRCA2* in the sample post-RCA compared with pre-RCA. Given a decrease in 1 cycle of amplification threshold corresponds to a doubling of the amount of the target this suggests that there was between a 32-fold and 64-fold increase in product after reaction 2. It was also noted that the no ligase sample was found to be negative for *BRCA2*, suggesting that as expected without the ligase treatment the endonuclease digest the DNA (as it is not circularised) and therefore there is no DNA present for the RCA cycles.

9.3 Sequencing using the MinION

Multiple sequencing runs were completed using the MinION (see Table 9.3 for summary).

Table 9.3: Summary of sequencing runs using the MinION

Sequencing run number	**1***	**2***	**3**	**4**	**5**	**6**	**7**
Run time (hours)	1.5	2	4	2.5	3.6	1.3	1
Read number	760	160	843	2199	2204	978	1471
Total bases (est. Mb)	3.81	0.54	2.21	5.91	2.72	0.81	0.98
Passed reads (%)	87.89	60	82.56	86.90	29.49	0	7.27
Median passed read length (kB)	3.45	2.97	3.30	3.45	3.21	-	1.37
Max passed read length (kB)	48.6	43.45	8.77	10.16	8.34	-	3.48

* Indicates samples that were used without RCA

9.3.1 Sequencing run 1

Using the preparation and sequencing protocol as per Oxford Nanopore Technology MinION steps (SQK-LSK109, protocol updated 23/05/18) for a non-RCA sample. The protocol included DNA repair, end-preparation (dA-tailing), adapter ligation, clean-up, flow cell priming and loading steps. The original gDNA sample was nano-dropped to assess rough quantity and quality and suggested 8.8ng/µL was present with a 260/280 of 1.86. After the initial DNA repair and end-preparation steps this increased to 17.1ng/µL with a 260/280 of 3.29. 760 sequencing reads were performed with an estimated 3.81Mb total read. 668 (87.89%) of the reads were passed with an overall median read length of 3.39kB (IQR 1.69kB-5.44kB). This was increased in passed reads to 3.45kB (IQR 2.03kB-5.80kB). The maximum passed read length was 48.59kB. The sequences were checked for matches to the sequence against the human genome using Blast and showed no evidence of human DNA.

9.3.2 Sequencing run 2

Run as per sequencing run 1 but for a total of 2 hours for a non-RCA sample. The original sample nano-dropped as 4.3 ng/µL (260/280 of 2.59), which increased to 20

ng/µL (260/280 8.53) after DNA repair and end-preparation steps were performed. 160 sequencing reads were performed, 96 (60%) of which passed with an estimate of 0.54Mb total read. The overall median read length was 2.28kB (IQR 0.60kB-3.88kB) which increased to 2.97kB (IQR 1.23kB-4.78kB) when only passed reads were included. Again, the sequences were checked for matches to the sequence against the human genome using Blast and showed no evidence of human DNA. It was thought that this could be an issue with DNA loss through the purification steps, therefore a further run was completed with the same preparation steps but with an assessment by qPCR throughout to assess where DNA was being lost.

9.3.3 Sequencing run 3

The run was stopped after 4 hours to review the data acquired. There was found to be 843 sequencing reads with an estimated 2.21Mb total read. 696 (82.56%) of the reads were passed reads. Overall, the median read length was 3.26Kb (IQR 1.67kB - 3.35kB) which increased to 3.30kB (IQR 2.96kB-3.36kB) for passed reads. The maximum passed read length was 8.77kB. The passed sequences were blasted and showed no evidence of human DNA. Waste from the AMPure XP bead washouts was collected through both clean-up steps as well as keeping the AMPure XP beads used. The waste as well as washout with NFW from the AMPure XP beads and the original sample tubes were run through qPCR for *BRCA2* to assess where the product was being lost (see Figure 9.1 for clarification). After qPCR the waste tubes were not noted to have any amplification for *BRCA2* however, both of the AMPure XP bead washouts (from both pellet 1 and 2) showed evidence of amplification. The original sample tube and the first AMPure XP washout had similar qPCR amplification cycles (31.99 and 31.18 respectively). The last Eppendorf used before sequencing was also washed and included in the qPCR with this showing no amplification for *BRCA2* (see Table 9.4 for amplification thresholds for qPCR throughout the sequencing runs).

The qPCR results suggested that a significant amount of DNA remained adherent to the AMPure XP beads used for purification. The bead manufacturer (Beckman

Coulter) was contacted to obtain advice about large fragment elution from the beads. They suggested:

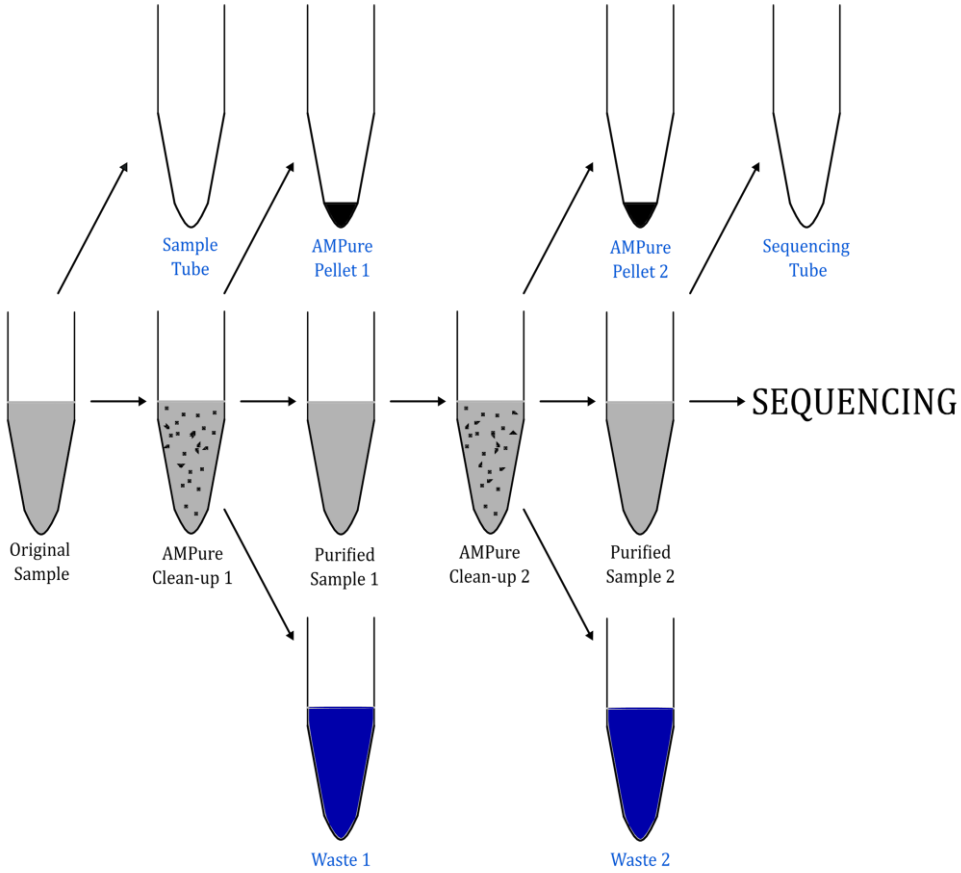
1. Making sure the beads and the elution buffer were well mixed before pelleting
2. Allowing additional incubation with elution buffer before pelleting
3. Heated elution (between 37 and 55°C)
4. Making sure all ethanol removed prior to elution
5. Eluting in NFW or diluting the elution buffer with NFW

9.3.4 Sequencing run 4

This run was adapted to include some of the changes above. During the elution steps the elution buffer was kept as normal but it was ensured that all of the ethanol was pipetted off the bead pellet prior to elution. The elution buffer and beads were well mixed prior to pelleting on a magnet. The elution steps were changed from 2 minutes for the first and 10 minutes for the second both at room temperature to 15 minutes and 30 minutes at 37°C. The steps above were followed throughout to ensure that qPCR could be used if no product was found to assess where this had been lost (Figure 9.1). The run was stopped after 2.5 hours with a total of 2199 reads and an estimated 5.91Mb read. 1911 (86.90%) of the reads were passed with an overall median read length of 3.43kB (IQR 1.51kB-3.49kB). This increased to a median of 3.45kB (IQR 2.01kB-3.50kB) when only passed reads were included. The sequences were checked again for matches to the sequence against the human genome using Blast with 8 human reads found; none were from the desired fragment.

qPCR was performed as previously which again showed that, whilst the waste products remained negative for *BRCA2*, both AMPure XP bead washes continued to show evidence of amplification for *BRCA2* (see Table 9.4).

Figure 9.1: Annotated steps for purification using AMPure XP beads. Black text signifies normal purification steps. Blue text signified the tubes/waste and beads used for qPCR.



9.3.5 Sequencing run 5

The elution changes were kept as per sequencing run 4 however, the elution temperature was increased to 55°C. Again, the steps were followed to ensure qPCR of both the waste and a further bead wash could be performed. The run was stopped after 3 hours and 40 minutes with a total of 2204 reads and an estimated total read length of 2.72Mb. 650 (29.49%) of the reads were passed with an overall median read length of 0.43kB (IQR 0.28kB-1.64kB). It was suspected that the low percentage of passed reads and shortened read length was due to DNA damage caused by the increased temperature. The median read length increased considerably to 3.21kB (IQR 1.29kB-3.43kB) when only passed reads were included. The sequences were checked again for matches to the sequence against the human genome using Blast and 7 matched human DNA reads with none from the desired fragment.

qPCR was performed as previously which showed that the AMPure XP bead washes continued to remain positive on amplification for *BRCA2*, with both of the bead pellet washes continuing to strongly amplify for *BRCA2* (see Table 9.4).

9.3.6 Sequencing run 6

The elution changes were kept as per runs 4 and 5 however, the elution temperature was increased to 80°C. Again, the steps were followed to ensure qPCR of both the waste and a further bead wash could be performed. The run was stopped after 1 hour and 20 minutes with a total of 978 runs however, all of these were failed. This was thought to be due to significant DNA damage at a higher temperature. The median read length was 0.57kB (IQR 0.39kB-0.72kB) and again it was thought that this much shorter read length was due to DNA damage and fragmentation. qPCR was performed as previously and the bead washout from the AMPure XP beads continued to have amplification for *BRCA2*, though this was noted to be less than previous attempts (see Table 9.4).

Table 9.4: qPCR amplification thresholds for comparison throughout each sequencing run from AMPure XP bead washes (the positive control used throughout is the same therefore has been included as a comparator) (AT = amplification threshold)

Sequencing run number	**3**	**4**	**5**	**6**
Positive control AT	24.74	24.65	24.72	25.06
Waste 1 AT	-	-	-	-
Waste 2 AT	-	-	-	-
Sample tube washout AT	31.99	-	32.46	36.06
AMPure pellet 1 washout AT	31.18	38.80	29.97	33.49
AMPure pellet 2 washout AT	37.14	-	33.90	37.64
Sequencing tube washout AT	-	-	45	-

9.3.7 Sequencing run 7

Discussion with other teams using AMPure XP beads suggested increasing the elution time further, plus decreasing the proportion of beads to DNA may alter the elution success. To this end the ratio of sample to AMPure XP beads was increased from 1:1 to 1:0.6, and any elution off the beads was left overnight at room temperature. The sequencing run was run for 1 hour and had a total of 1471 reads of which only 107 (7.27%) were passed reads. The median read length was 0.48kB (IQR 0.33kB-0.66kB) which increased to 1.37kB (IQR 0.99kB-1.77kB) when only passed reads were included. The low percentage of passed reads and the low read length was again suspected to be due to increased DNA damage when the elution time was increased to overnight. There was again found to be some human gDNA after checking for matches to the sequence against the human genome using Blast however, none of this was from the desired fragment. Given there continued to be DNA fragment adherent to the AMPure XP beads it was decided to assess the difference between elution solutions for detaching from the beads alongside the ratio of beads to original DNA solution.

9.4 AMPure XP bead elution

Assessment of elution by both Nanodrop and qPCR for *BRCA2* were used for the same DNA containing sample but with four different treatments. The four samples used were:

1. AMPure XP bead to solution ratio 1:1 with elution in NFW
2. AMPure XP bead to solution ratio 1:1 with elution in Tris-Cl
3. AMPure XP bead to solution ratio 0.6:1 with elution in NFW
4. AMPure XP bead to solution ratio 0.6:1 with elution in Tris-Cl

Performed as previously the ethanol used to wash the beads was removed with pipetting and the pellet was left to completely dry. The eluting solution and the beads were mixed well and incubated at room temperature for 10 minutes for the elution.

Nanodropping was used to assess DNA purity and amount, with ideal 260/280 values at 1.8-2 and 260/230 between 2 and 2.2. qPCR was performed as per previous. Table 9.5 shows the results of both the Nanodrop assessment and qPCR.

Table 9.5: Nanodrop and qPCR amplification threshold assessment for elution from AMPure XP beads using various solutions (better values shown in green, worst in red)

AMPure to solution ratio	Elution solution	260/280	260/230	Nanodrop DNA amount (ng/µL)	qPCR amplification threshold
1:1	NFW	1.57	2.09	25.8	33.97
1:1	TrisCl	1.65	1.87	39.6	36.93
0.6:1	NFW	1.71	2.36	30.2	36.84
0.6:1	TrisCl	1.56	1.8	23.4	38.86

It was noted that both the Nanodrop and qPCR assessment suggested that the elution 0.6:1 using Tris-Cl performed the worst. There appeared to be more DNA detected by Nanodrop in the 1:1 Tris-Cl however, the most amplification for *BRCA2*

was seen in the 1:1 NFW sample. The 260/280 ratio was low for all of the samples, suggesting a high level of contamination however, this would likely be reduced with further AMPure XP bead treatment. Given the 1:1 NFW treatment appeared to have the best result for the desired fragment it was decided to attempt further optimisation using AMPure XP bead elutions.

It was also suggested that another measure of DNA amount could be used. Qubit fluorometric quantitation analysis was performed on the above samples alongside the original post-RCA sample. This measured 1.11ng/ml in the post-RCA sample but all other samples registered <0.50ng/ml and were therefore too low to accurately measure DNA quantity.

In an attempt to elute the majority of the sample off the AMPure XP beads after RCA a sample was treated with AMPure XP bead clean-up preparation at a ratio of 1:1 beads: sample and eluted four times (washed with NFW and re-pelleted) in 10µL NFW (see Figure 9.2 for simplification).

qPCR for *BRCA2* showed no evidence of amplification in the waste products, as previously. There continued to be amplification for *BRCA2* on qPCR throughout each washout and this did not seem to diminish with each successive wash. This suggested an ongoing considerable attachment between the large DNA fragment and the AMPure XP beads.

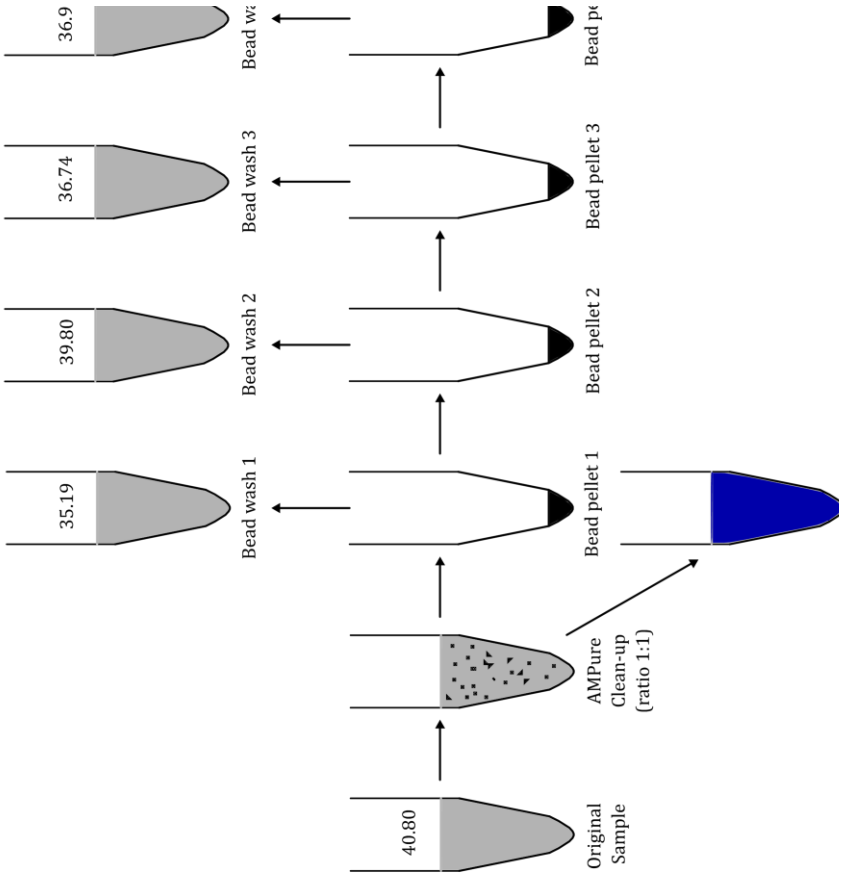


Figure 9.2: Diagram of multiple AMPure bead washes with NFW prior to qPCR. Numbers shown are the amplification threshold for qPCR for *BRCA2*

9.4 Discussion

The RCA process appeared to have performed well, although the difficulties with larger fragment DNA made assessment of the exact amount of DNA impossible. The positive control used for qPCR was diluted from a nano-dropped sample and contained 3.75ng/µL human DNA. The positive control for reaction 2 reported an amplification threshold of 26.85. This is very close to that of the RCA samples. It is highly likely that the actual amount of DNA in the RCA sample is significantly less than the positive control, given that the RCA sample has been selected for *BRCA2*. Unfortunately, it is very difficult to assess exactly how much of the sample is the expected fragment due to the size of the DNA fragment. It may have been beneficial to perform RCA for further cycles to ensure as much as possible of the desired fragment was present. This may also have benefitted from further PFGE or southern blotting to ensure that adequate RCA cycles had been performed to allow a visible band at the desired fragment size.

The MinION is a reasonably recently created sequencing platform and has continuous protocol updates. The preparation protocol used was the most up-to-date at the time. Although following the original protocol (Sequencing run Number 1) provided the maximum read length, there were more reads with an increased elution time and temperature.

Unfortunately, even after multiple elutions, the fragment remained adherent to the AMPure XP beads. Whilst increasing the temperature did improve this, it did so at the cost of sheering and a reduced number of passed reads.

The reported problems with sequencing are likely due to an issue of input DNA amount, and it may be that, even with a significant amount of DNA remaining on the beads, if larger DNA input amounts were used the sequencing may be improved. There could also be an issue of input DNA quality however, it was noted that the percentage of passed reads reached up to 87.89% with non-RCA samples and 86.90% with initial RCA samples. The drop in read length from the non-RCA samples to the RCA samples could indicate either incomplete replication during the

RCA cycles or an increase in shearing due to the extra procedural steps. Further work with increased input DNA may have achieved desired reads. It may be possible to achieve this in the future with further RCA cycles however, as discussed above this may come at the expense of DNA quality. Sheering of the large fragment DNA has been used in other similar projects with improved read results and DNA quality however, this would not then allow for intact sequencing of single long strands. This would lose information about sequences relative to each other and make forming haplotypes impossible. Clonal populations of the required DNA strand could also have been produced by using Bacterial Artificial Chromosomes, although this would be a highly complex and labour-intensive approach.

10. Results: IMMPACT study

10.1 Introduction

The IMMPACT study aimed to recruit individuals at increased risk of pancreatic cancer. This includes those with known genetic predisposition as well as those with a significant family history. The recruited individuals have been discussed previously. This study was created aiming to increase the yield of pancreatic cancer found during secondary screening by stratifying individuals. To this end, samples collected from recruited individuals were used to assess the efficacy of a panel of blood biomarkers, the IMMRAY platform (IMMUNOVIA). In this chapter we will discuss the outcome of the individuals recruited to the IMMPACT study along with the outcome of the IMMRAY platform.

10.2 IMMPACT individuals

At the time of writing, 85 individuals were recruited to the IMMPACT study, with a total of 203 blood samples collected. To date one individual withdrew from the study due to individual preference. Recruitment began in April 2017.

During the study period 14 (16.47%) individuals were found to have abnormal scans. Four of these individuals were known to have scan abnormalities prior to recruitment into the study. These individuals were all found to have branch-duct IPMNs requiring ongoing screening. During the study period one of these individuals developed a further main duct IPMN which required further investigation. The cytology was found to be benign and after MDT and patient discussion was planned for further monitoring but no further treatment.

Ten individuals were found to have abnormal screening findings as a result of the screening process. Table 10.1 shows the number and type of abnormalities found.

Table 10.1: Number and type of abnormalities found during screening along with their outcome

Screening Abnormality	Outcome	Number
IPMN - Body	Referred to local pancreatic centre	1
IPMN – branch-duct	Screening continued	5
IPMN – main-duct	Aspiration found benign; screening continued	1
Dilated pancreatic duct	Total pancreatectomy	1
Branch duct ectasia	Screening continued	1
Hepatic focal nodule hyperplasia	Referred to liver centre	1

Median secondary screening follow up was 34 months (IQR 27-37 months). This follow-up period included consecutive CT or EUS secondary screening regardless of whether individuals continued to attend for further blood sample collection.

The individual who required a pancreatectomy was found to have a mixed-duct type IPMN with high grade dysplasia but no invasive malignancy.

10.3 Normal Vs abnormal screening findings

This comparison does not include the individuals who were noted to have abnormal scan findings prior to recruitment.

Two (20%) of the individuals with abnormal scans had known genetic mutations (one *PRSS1* and one *BRCA2*) compared with 12 (16.9%) of the individuals with normal scans (2 *PRSS1*, 1 *BRCA1*, 6 *BRCA2*, 1HNPCC, 1 PJS and 1FAMMM).

7 (70%) of those individuals with abnormal scans compared with 48 (67.6%) of those without reported themselves to be female. There was no significant difference between the groups (p=0.73).

The median age of those with abnormal scans is 60.5 years (IQR 57.25-62.75) and 56 years (IQR 49.5-64.75) for those without. There was no significant difference between the groups (p=0.39).

There was also no significant difference in number of blood samples between the two groups (p=0.8956).

10.4 Immunovia studies

The IMMPACT study was created, in part, to allow collaboration with Immunovia (Lund, Sweden), aiming to validate their biomarker IMMray© platform for detection of early pancreatic cancer in individuals with higher than population risk (panFAM-1 study). The panFAM-1 study halted recruitment in October 2020, with over 3000 samples collected from 1265 subjects from 23 sites (Press release 15/2/21). The final samples are due to be collected by the end of April 2021 and the platform tested during the second half of 2021.

The IMMray© platform has already been tested for use in individuals with early-stage pancreatic cancer when compared with healthy controls (ROC AUC 0.950). There is ongoing testing for the platform in individuals with non-specific but concerning symptoms (PanSYM-1 study). Results of this study so far are promising for detecting early pancreatic cancer (specificity 92%, sensitivity 80%) however, there is a suggestion that the combination of the validation samples plus new samples were used to assess the platform thus far. The IMMray© platform is also being assessed for detection of early pancreatic cancer in individuals with new onset type 2 diabetes mellitus over the age of 50 years, with the PanDIA-1 study currently recruiting across multiple sites.

10.5 Discussion

The PanFam-1 study completed recruitment in October 2020 and completed blood sample collection in April 2021. The results of the PanFam-1 study are due to be released later in 2021. 16.47% of the individuals recruited to the IMMPACT study were found to have abnormal scans, the majority as expected being IPMNs. There were no findings of PDAC during the study period. It is likely that the PanFam-1 study, given the multi-centre, international, large-scale recruitment will provide a better assessment of PDAC radiological screening, alongside epidemiological data and assessment of the multi-biomarker platform (IMMray©).

11. Discussion

11.1 Introduction

During this project the population recruited to the EUROPAC study has been evaluated to allow for further adaptation. Alongside this there has been recruitment to the IMMPACT and the PanFam-1 studies which aim to further evaluate the population receiving screening for PDAC along with validation of multiple biomarkers to precipitate early detection of PDAC. Evaluation of these platforms has allowed in-depth analysis of a population with pathogenic *BRCA2* mutations and their risk of PDAC. Multiple confounding factors for PDAC risk in conjunction with *BRCA2* mutations have been compared.

During the IMMPACT study period multiple abnormalities were found on radiological screening however, none of these were PDAC. As previously discussed, it is likely that the larger, international PanFam-1 study will have a higher number of PDAC cases found and therefore be further able to evaluate cancer risk and early diagnosis biomarkers.

11.2 EUROPAC and IMMPACT studies

11.2.1 Who is deemed high risk for pancreatic cancer?

Screening and risk of pancreatic cancer continues to be an ongoing debate within the scientific and clinical communities. Discussion about inclusion criteria for screening is critical to understanding the underlying EUROPAC population studied. The ideal aim of pancreatic cancer screening is for early diagnosis when treatment is more likely to be curative. The lifetime risk of pancreatic cancer in the general population remains relatively low (~1%) and to this end screening of the general population would not be feasible or viable due to the low actual numbers of pancreatic cancer within the population. Screening then needs to be offered to individuals who are more at risk than the general population to take into account

the risk of false positives alongside screening cost. In 2013 The International Cancer of the Pancreas Screening Consortium (CAPS) agreed that individuals >5% lifetime risk of pancreatic cancer should be offered image screening for early detection (71).

The majority (98.89%) of the 'high risk' pancreatic cancer group within the FPC group in EUROPAC were individuals with a strong family history of pancreatic cancer (those with no known pathogenic mutations causing an increase in risk). It is suspected that there is an autosomal dominant genetic predisposition in many of these families leading to an increased risk of pancreatic cancer (19). It is possible in this population that some (or all) of the individuals affected by pancreatic cancer within a family could be sporadic, and that there is therefore no increased risk to other members of the family. The relative risk amongst individuals in FPC families vary from 6 to 120- fold depending on the family composition and clearly those at the lower risk end of the spectrum are less likely to benefit from pancreatic screening (120,121). The difficulty continues to be separating the lower risk families from those with an actual high-risk predisposition to pancreatic cancer.

Coupled with this, if an autosomal dominant unknown mutation is responsible for the majority of the pancreatic cancer risk in the majority of true FPC families then not all of the individuals within the family are at risk, given not all will carry the mutation. Should the affected mutation truly follow autosomal dominant characteristics then only 50% of the next generation should carry the mutation. To further complicate this the actual incidence of pancreatic cancer within this group will depend on the gene penetrance, plus the additional possibility of truly sporadic pancreatic cancer within these families in an individual without a mutation.

There is also evidence that some families with a higher incidence of pancreatic cancer may have an underlying mutation known to predispose to a cancer syndrome but not have any other familial features of that syndrome to warrant testing, for example a family with a *BRCA2* mutation but no evidence of breast cancer outside of normal population parameters. This could amount to as much as 12% of the FPC population and some groups have suggested that testing for

mutations known to increase pancreatic cancer should be routine in all new presentations (122–124).

Risk of developing pancreatic cancer is also known to be variable in individuals with hereditary pancreatitis depending on both the underlying mutation status along with health and social factors (41,79).

Regardless of the underlying genetic predisposition there are also other complicating factors which may alter an individuals' risk of developing pancreatic cancer and thus creating difficulties for screening programmes. Most of these risks, whilst well known, are not easy to quantify against genetic risk and therefore have not been added into screening programmes and criteria.

Diabetes mellitus is now accepted to be bidirectionally linked with pancreatic cancer however, this link remains complex (31,125). Efforts to select a screening cohort from those individuals with new onset diabetes mellitus who are at increased risk of developing pancreatic cancer are ongoing (126). Screening within this group will have to account for the large numbers of the general population who develop diabetes mellitus but are not at an increased risk of pancreatic cancer. Smoking tobacco is also known to increase risk of pancreatic cancer and it is accepted that this risk is even more significant in those individuals with hereditary pancreatitis (80). Body mass index is also inversely related to risk of pancreatic cancer even when adjusted for other lifestyle factors, such as smoking (127). These health and social factors are well investigated and reported however, at present adjusting an individuals' risk of pancreatic cancer using these markers alongside familial predisposition is not possible. Again, this potentially means that individuals are included as high risk when perhaps they would not be, but also that some individuals who are at higher risk are not included for secondary screening. Practically speaking, at present, the inclusion of health and lifestyle factors into screening criteria is not possible.

Increased age has been shown to increase risk of pancreatic cancer both in sporadic and hereditary cases (72,128).

There continues to be disagreement about precisely the age at which the risk of developing pancreatic cancer in those at risk becomes high enough to justify secondary screening in those families deemed high risk. In 2013 CAPS suggested that screening for this population should commence at 50 years old however, there was only 51% agreement. There was also a suggestion that individuals with hereditary pancreatitis should begin screening at 40 years old and that smokers should potentially be offered screening earlier however, there was no consensus (71). In reality many European screening programmes, including EUROPAC, commence secondary screening at 40 years old or within 10 years of the youngest affected family member (129,130). There also remains no consensus as to when secondary screening should cease in high risk individuals (71). At present the EUROPAC group stop screening based on clinical rather than age grounds, meaning that individuals who would not be clinical candidates for a pancreatic resection would generally be encouraged not to have further secondary screening. The upper age of screening is important due to the possibility of frailty making surgery inappropriate and because there is more risk of an older individual not being a carrier of a mutation (known or unknown) that increases their risk of PDAC.

Individuals are offered secondary screening based on any syndrome known to increase the risk of pancreatic cancer however, these different syndromes themselves do not confer the same risk. For example, an individual with a CDKN2A mutation (Familial atypical multiple mole melanoma syndrome (FAMMM)) with a family member affected with pancreatic cancer would be offered the same annual secondary screening as an individual with a MLH1 mutation (Lynch syndrome) with a family history although the two syndromes confer a significantly different risk of pancreatic cancer (FAMMM cumulative risk 17% at 75 years, Lynch cumulative risk 3.7% at 70 years) (81). Some cancer syndromes confer an increased risk of pancreatic cancer but is deemed insufficient to warrant screening in the UK; for example *BRCA1* mutations leading to HBOC, which is reported as carrying a 2.2-fold increase in risk of pancreatic cancer (131). Alongside this there is also the risk that when offering screening to families with a pathogenic mutation and a family history

of pancreatic cancer, that the individual affected was a sporadic rather than hereditary case of pancreatic cancer.

To further complicate this matter individuals with the same 'at risk' syndrome are screened as if they have identical risk patterns. This is known to be untrue, with individuals with the same clinical mutation having wildly different cancer risk profiles (99,132–135). This will be discussed further within the context of *BRCA2* mutations later. Further stratifying these risk profiles amongst individuals with a pathogenic mutation may help to improve cancer detection rates and reduce unnecessary screening.

11.2.2 Comparison of screened and non-screened individuals

Firstly, it important to note the large amount of missing data, especially regarding non-screened individuals. Whilst every effort was used to make sure that the data was as complete as possible by reviewing the paper notes available to EUROPAC, there was still found to be between 11.0% and 54.5% data completeness for the non-screened and between 60.1% and 99.7% for the screened individuals. There is the possibility that this could mean the non-screened population is not representative of that population as a whole and that the data available holds significant selection bias. There is also the issue with release of the NICE 2018 guidelines highlighting the importance for pancreatic screening for high risk individuals, this may mean that some individuals who would previously been referred to EUROPAC may be screened locally by their GP and therefore actual numbers of screened individuals could be increased (116). Coupled with this, there is no direct comparison between individuals who were offered secondary screening and declined, and individuals who have not been approached or are not aware they are of increased risk. It is likely that there are differences between these groups however, without approaching individuals it is difficult to assess their knowledge of

their pancreatic cancer risk awareness to further analyse the impact. Within this study these two groups have been combined as the not screened group.

Overall, only a small proportion of individuals deemed suitable for secondary screening attended (6%). This is comparable with some selective screening programmes for other cancers both in the UK and in the USA (136,137). Unfortunately to date no information is available to compare this with attendance rates for pancreatic cancer screening in the same population in other countries. It is difficult to compare with the figures available for the breast and cervical cancer screening programmes, both of which report roughly 70% uptake, within the UK as they report a success as one screening episode within the screening age range, rather than regular annual attendance (138,139). A more comparable screening programme in the UK would be that for Barrett's oesophagus, as this requires a gastroscopic investigation and studies have suggested roughly 40% of individuals regularly attend (140). Whilst this is a comparable investigation to endoscopic ultrasound (EUS) and more targeted screening, there is a significant difference in the risk of oesophageal cancer with Barrett's oesophagus (lifetime risk 1 in 8 to 1 in 14) than pancreatic cancer with FPC (141).

There continues to be a wide range of reasons why individuals may not attend for or accept cancer screening. Many psychosocial reasons have previously been raised with other screening programmes such as individuals finding tests unpleasant or simply not being aware of their risk (137). This could be the case with pancreatic cancer screening given the main annual examination is with EUS which is an invasive test. Interestingly there may also be difficulties with general awareness of pancreatic cancer risk within the general population. Triplette and colleagues highlighted the issue of public health campaigns to increase awareness of certain cancer risks if they do not reflect a cancer that is common in the overall population, instead they recommended targeting at risk populations and their healthcare professionals to increase uptake (136). This does represent a challenge for the EUROPAC group given there is ethical agreement that the team cannot contact individuals who have not been referred, either by self-referral or a healthcare

professional, but have been highlighted by other family members during family history screening. This is crucial to stop passing of confidential information, by making individuals aware that another family member has contacted EUROAPAC, but it does mean that individuals who would benefit from screening yet are not aware of their risk are missed. Whilst the 2018 NICE guidelines: Pancreatic cancer in adults (116) highlights the importance of screening in individuals at higher risk it may be that General Practitioners and other healthcare givers are unaware. It is important to note that whilst the registered EUROPAC population may only encompass a small percentage of the at-risk group within the United Kingdom population, there also remains a large gap between those eligible for screening and those attending.

It was also noted that attendance for secondary screening was significantly less (2.5% versus 7%) for the HP than the FPC group. There are likely to be multiple reasons for this difference. Firstly, and most predominantly, multiple individuals are referred to EUROPAC for HP and have no discernible family history. These individuals are more likely to suffer from idiopathic pancreatitis and therefore require more considerable medical investigations prior to screening to rule out any obvious treatable causes of chronic pancreatitis. Alongside this, the HP and FPC populations have different profiles, with the HP population being more likely to be male, in a lower sociodemographic group, live further from screening centres, currently smoke tobacco, but less likely to drink alcohol. The difference between the FPC and HP groups in lifestyle habits, such as smoking tobacco and drinking alcohol is interesting as the lack of alcohol in the HP group may suggest a linkage with concern over worsening of pancreatitis symptoms and complications however, the increased tobacco smoking, goes against this hypothesis. This may reflect the knowledge of the general public and some health professionals who may not be aware of the increased risk of pancreatitis with smoking tobacco. It was noted that considerably more of the HP individuals with diabetes mellitus attended for secondary screening and this could reflect a marker of pancreatitis severity, given

that individuals with severe or multiple episodes are more likely to suffer with diabetes as a result.

11.2.2.1 Self-reported gender and screening attendance

It was reported that 64.3% of the screened and 54.2% of the non-screened individuals reported themselves as female. This reflects a female preponderance both in the individuals registered with EUROPAC and those attending for secondary screening. Overall, only 7.07% of the females and 4.77% of the males that were suitable for screening attended. This is particularly interesting when considering that both pancreatic cancer and pancreatitis are more common in males, although there is emerging evidence that the risk of pancreatic cancer in those with FPC is similar regardless of sex (128,142). In England the only national cancer screening programme that invites both men and women, the Bowel Cancer Screening Programme (BCSP), reported in 2012 that although males were at higher risk of colorectal cancer, only 49.65% men compared with 54.4% women returned their tests (143). The discrepancy in the UK between the sexes' acceptance of screening may reflect the fact that women are invited to multiple other cancer screening programmes from a younger age (i.e. cervical and breast screening) and therefore may be more understanding and accepting of screening compared with men. Alongside this, Vrinten and colleagues report that stigma against cancer diagnosis is more likely in men, for example they more commonly hold the belief that cancer is a career-ending diagnosis and that it ruins personal relationships (144). The disparity between male and female uptake is however, not consistent with other cancer screening programmes internationally, where multiple studies have shown comparable uptake for both sexes (136,137).

11.2.2.2 Age and screening attendance

The reported median age for individuals attending for screening was marginally younger than the non-screened group (57 years compared with 59 years). The available evidence reports an inconsistent relationship between age and screening

attendance with some studies reporting no effect and others suggesting that younger individuals are less likely to attend for screening (137,145). There was also noted to be a significant difference between the age distributions between the groups however, it is possible that this could reflect a considerable difference in the sample sizes. Interestingly, the violin plot of the age distribution within the non-screened population suggests the possibility of two age peaks, one roughly at 47 years and the other at 72 years. This may represent the different risk populations in the participants registered for EUROPAC, with the early large peak representing a mix of actual high risk and assumed high risk individuals with the wasting of the graph due to early deaths. The smaller later peak could then represent those individuals who were assumed high risk but are not. This two-peak distribution could possibly also reflect gene penetrance of a high-risk pancreatic cancer gene.

Alongside this it appears that older and younger individuals do not attend for screening. This may reflect a low perceived risk in younger individuals and older individuals may differ in their understanding of the value of screening when other health parameters may take priority.

11.2.2.3 Distance from screening centre and screening attendance

It was noted that the majority of individuals both screened and not screened had either Liverpool or London as their closest centre. It is highly likely that this is more a statement of the distribution of centres than individuals choosing to live in proximity to one of these centres. It was noted that the median distance from the nearest screening centre was reduced in the non-screened compared with the screened population. Alongside this the violin plots of the distance from centre appear to show that the distance distribution between the two groups is different, with the non-screened groups appearing to have multiple peaks (likely at distances of heavily populated areas, i.e. cities), whereas the screened distribution appears much smoother. It is likely that this represents the difference in population sizes between the two groups. It is difficult to compare these results with the literature given that the majority of other screening programmes are run from many more,

and therefore likely closer centres. Maheswaran and colleagues provide a good example of this, with women being more likely to attend for breast cancer screening if they were within 5Km of the screening centre, clearly much closer than distances in our population (146). The distance difference between the screened and non-screened population within the EUROPAC group could more reflect ease of travel than distance travelled and this has been reflected previously in the literature. Wang found a considerable positive link in the UK population between attendance for cancer screening and car ownership, as well as access to public transport and attendance for screening in younger groups (147). The studies described are for population screening programmes, to date there is no comparable data for any targeted cancer screening.

11.2.2.4 Family history of pancreatic cancer and screening attendance

Firstly, it is important to note that the number of individuals with family history of pancreatic cancer will be skewed given that the group are referred due to an increased risk of pancreatic cancer. The inclusion criteria for the EUROPAC database require either a pathogenic mutation and usually at least one relative affected or at least two first-degree relatives affected, hence the majority of EUROPAC individuals have at least two relatives affected. There was found to be no difference in the distribution of the number of family members affected with pancreatic cancer between the screened and non-screened individuals. This suggests that the presence of more family members who have suffered with pancreatic cancer does not provide any further inclination to attend for secondary screening aiming to catch this at a potentially treatable stage. This may reflect fear behaviours which makes some individuals more aware of the risks of cancer (and the paths it may take) but less likely to attend for screening (148). The result of higher cancer fear leading to reduction in screening attendance is reflective of individuals being worried of 'finding something wrong' and 'preferring not to know' however, this reaction is not consistent (149). Consedine and colleagues did suggest a model for predicting the 'fear' relationship with attendance for screening with those

individuals concerned specifically about the screening process being less likely to attend, but those with a non-specific fear of cancer being more likely to attend (149). They also suggested that behaviour of individuals worried about their possible screening outcome was difficult to predict. Unfortunately, due to lack of consistency and predictability in indiviudals' behaviour and reasoning when deciding to attend for screening, this is incredibly difficult to target on a population level. This would be an interesting and worth-while topic for further study however, it would be difficult to assess given there is likely also a correlation between individuals who are more likely to fill in a survey assessing their cancer awareness and perceptions, and those who are more likely to attend for cancer screening.

11.2.2.5 Sociodemographic score and screening attendance

A larger proportion of the screened individuals were found to be in the two most affluent sociodemographic groups when compared to the non-screened. Multiple studies in developed countries show a similar pattern for cancer screening uptake (150–153). Interestingly many of these studies are performed in areas where health services are more privatised and therefore a significant improvement has been seen in cancer screening uptake with access to free or subsidised screening (150,151). Previous research has also suggested that individuals from lower sociodemographic groups are more likely to present with later stage disease than their more affluent counterparts (154,155). It is difficult to completely assess the underlying reasons for this disparity however, there is some evidence that this may reflect a lack of perceived benefit and an increase in perception of the risks of cancer screening in this group alongside reduced health awareness (156). Targeting this disparity remains an issue for healthcare as a whole, rather than just cancer services and multiple efforts are ongoing to close the perceived 'post-code lottery' for healthcare provision within the UK.

The significant difference in sociodemographic score persisted even when the FPC and HP groups were separated. It was noted that there was a significant difference in the sociodemographic score between the HP and FPC groups, with the HP group

having a larger proportion in the less affluent groups. It may be that this difference relates to a disparity in predisposing lifestyle factors between the more and less affluent. To date there is no available research describing the link between hereditary pancreatitis and socioeconomic factors, nor the difference between HP and FPC. Han and colleagues found that smokers with chronic pancreatitis were more likely to be from low-income households and were also more likely to require enzymatic replacement, linking to more severe disease (157). It is likely that this link persists in individuals with hereditary pancreatitis. Further discussion about the disparities in lifestyle factors in the screened and non-screened group can be found below.

There continued to be a significant amount of data not available for the non-screened individuals for the postcode, with only 16% of the data present. Whilst there is no obvious reason for those missing to bias the sociodemographic score it cannot be ruled out that individuals from more deprived backgrounds may be less likely to present to medical care. This absence of data may actually be masking a more pronounced difference between the screened and non-screened individuals.

11.2.2.6 Lifestyle factors and screening attendance

There was a statistically significant difference in the proportion of individuals who smoked tobacco in the non-screened compared with the screened group (19.7% Vs. 7.3%). This could indirectly reflect the sociodemographic data above but could also be a marker for the more health-conscious individuals, with those individuals who are more health conscious being more likely to attend for screening and not smoke. There also continued to be a disparity between the HP and FPC groups, with the HP individuals being more likely to currently smoke, again this could link smoking to pancreatitis episodes and severity in the HP group. Multiple studies have echoed that individuals who did not attend for screening were more likely to smoke tobacco but also be from a more deprived background (158,159). Again, it is difficult to separate whether this correlation is due to background or smoking status or some combination of both.

Interestingly there was found to be a significant difference in those reporting currently drinking alcohol, with more individuals in the screened than the non-screened group reporting regular alcohol intake. This discrepancy between alcohol intake and smoking could reflect current social norms, with drinking alcohol being considered more socially acceptable than smoking tobacco (160,161). There was also a significant difference in the alcohol intake reported between the HP and FPC groups, with the HP group having significantly fewer individuals reporting a regular alcohol intake. This difference is likely due to the individuals who are predisposed to pancreatitis being keen to avoid any precipitating factors, such as alcohol however, it is interesting to note that this is not reflected in the smoking behaviour of the HP group. Richard and colleagues reported no relationship between alcohol intake and attendance for cervical cancer screening (159). Lagerlund and colleagues however, found that women who did not attended for breast cancer screening more commonly did not report any alcohol intake in the last year (162). It could be that this difference in alcohol intake and smoking tobacco represents the concern of the individual, for example individuals concerned about their cancer risk may be less likely to smoke and more likely to attend screening, whereas individuals concerned with pain (such as individuals with HP) may be less likely to drink alcohol, as this precipitates attacks. There is little further research assessing any relationship between alcohol intake and attendance for cancer screening.

A further possible link between screening and health consciousness of individuals is found with BMI. There was found to be an increase in individuals classed as overweight, obese and morbidly obese in the non-screened group however, this was not statistically significant. It has been suggested that individuals in the obese and overweight categories are less likely to attend for cancer screening (159) and alongside this several studies have reported a link between a more sedentary lifestyle and cancer screening non-attendance however, this was not assessed in our study group (159,162).

There remains very little published evidence assessing both sexes and these lifestyle factors due to most population cancer screening being for cervical, breast or

prostate cancers. Alongside this most evidence appears to assess population-based screening programmes rather than more directed screening. There appears to be no evidence available for the comparison of lifestyle factors between individuals attending and not attending for pancreatic cancer screening. Clearly, there are difficulties with assessing lifestyle data given that the majority is self-reported and therefore subject to multiple biases. Furthermore, there remains significant missing data from the non-screened group which could further skew the analysis.

11.2.2.7 Presence of pathogenic mutations and screening attendance

There was noted to be significantly more individuals with pathogenic mutations in the screened compared with the non-screened group. Overall, 12.43% of the screened and 4.56% of the non-screened populations had known pathogenic mutations. Multiple studies have assessed the rate of germline mutation amongst individuals with presumed sporadic pancreatic cancer and found between 4 and 25% (27,163,164). It is likely that the rate of mutations in the EUROPAC group would be higher than reported given that current guidelines do not suggest that every individual with pancreatic cancer should be tested for a mutation, therefore only those who meet the criteria for cancer syndromes are tested. These criteria are much stricter in the UK than in the USA. Multiple studies have suggested that the rates of pathogenic germline mutations in both individuals with assumed pancreatic cancer and those with FPC are higher than expected even without a 'classical' family history (27,163–166).

BRCA2 was the most common mutation in the FPC group with 2.03% of the screened and 0.33% of the non-screened having a confirmed pathogenic mutation. The rate of *BRCA2* mutations within the general population is suspected to be 1 in 195 (167). Studies have confirmed the rate of *BRCA2* mutation amongst individuals with pancreatic cancer to be 0.5-2.5% (27,163,164).

CDKN2A mutations were found in 1.74% of the screened and 0.44% of the non-screened FPC groups. Little information is available about the suspected prevalence

of *CDKN2A* mutations within the general population however, multiple studies have suggested between 0.12- 2.5% in individuals with apparently sporadic pancreatic cancer (27,163,165). A 2015 study assessing the rates of *CDKN2A* mutations in individuals related to those with sporadic pancreatic cancer found 2.5% had a pathogenic mutation (166).

It was also noted that individuals with confirmed pathogenic mutations were more likely to attend for secondary screening. This may be linked to an increase in awareness of cancer outcomes as well as having witnessed other relatives with malignancies in the past. There are no current published studies assessing the link between presence of a pathogenic mutation and attendance for secondary screening however, this is unsurprising given that in many other cancers an increase in screening is not offered to those without evidence of a predisposing mutation. This again raises the issue of screening individuals with pathogenic mutations as not all individuals with a cancer syndrome confer the same cancer risk. There is the potential that some individuals attend cancer screening and have significant anxiety about developing cancer but carry very little (or perhaps no) increased risk.

11.2.2.8 Presence of diabetes mellitus and screening attendance

There was found to be no overall difference in the number of individuals diagnosed with diabetes between the screened and non-screened group. This remained the case with just the FPC group however, diabetes was significantly more common in the screened HP group compared with the non-screened (57.38% Vs 13.11%). It is likely that this reflects disease severity, with more severe chronic pancreatitis being more likely to lead to diabetes. This greater impact on the individuals' health may make them more likely to attend secondary screening. There is a possibility that diabetes could be under-reported within the non-screened group. Multiple studies have assessed the natural history and development of diabetes in individuals with chronic pancreatitis, with rates ranging from 25-80% (41,80,168–170).

No data is currently available assessing the relationship between current diabetes status and attendance for pancreatic cancer screening. Much of the current published information about diabetes and cancer screening relates again to population screening and predominantly female cancers, where there appears to be no relationship between diabetes and screening attendance (159,171). It is important to note that pancreatic cancer correlates significantly with diabetes mellitus and to this end multiple studies are currently recruiting to further assess whether screening individuals with new onset type 2 diabetes leads to earlier diagnosis of pancreatic cancers (36,84).

11.2.3 Comparison of individuals recruited to the IMMPACT study

The comparison of the group recruited to the IMMPACT study with those attending EUROPAC for secondary screening shows no obvious differences, aside from being more likely to attend secondary screening in Liverpool. This suggests that the IMMPACT group is broadly representative of those individuals attending for secondary screening. It is important to note that this group was selected with more stringent criteria than those attending for secondary screening therefore it was expected that the two groups would be similar. Alongside this, the groups' attendance at secondary screening already suggests a highly motivated, health-conscious group who are more likely to attend health appointments.

The small number of HP individuals within the IMMPACT group is likely due to the change in inclusion criteria in November 2018 to include individuals with PRSS1 mutations and a history of chronic pancreatitis rather than any difficulties with recruitment of such individuals.

11.2.4 IMMPACT study recruitment and PanFAM1 study

The IMMPACT study overall recruited 85 individuals, 14 (16.47%) of whom were found to have abnormal scans (10 of which were new findings). At the time of writing no PDACs were found during the study period.

These individuals were recruited to provide samples and contribute to a larger multinational study, PanFAM1 (Immunovia). This study completed recruitment in October 2020 and completed sample collection in April 2021. To date there has been collection of over 3000 blood samples from 1265 high risk individuals. So far no information is available on the overall cancer detection rates or the performance of the IMMray platform however, this is expected towards the end of 2021 (85).

11.3 *BRCA2* families

61 families were found to have pathogenic *BRCA2* mutations consisting of 830 individuals. Obviously within this group there will be individuals without *BRCA2* mutations however, there is no reason to suspect that their number would differ between the families. It was noted that only 35% of the families had mutation data available despite efforts made to contact genetics services and general practitioners to gain this information. Whilst this does mean that analysis of the mutation type is missing for this group, it is still likely they have a pathogenic mutation given that referral to EUROPAC requires a clinician confirmation of this prior to screening. As previously discussed, the rate of pathogenic *BRCA2* mutation amongst the EUROPAC group appears to sit within the range found in published literature within HPC populations.

It was noted that only 5 of the mutations were common to more than one family. There was no correlation between the site of the mutation and the family group. This suggests that specific sites of *BRCA2* mutations do not define specific risk for particular cancer types however, there are too few shared mutations to be fully confident in this conclusion. Multiple publications suggest that *BRCA2* mutation location does not completely explain the heterogeneity in cancers found within families with pathogenic mutations (99,132,133,135,172).

11.3.1 Difficulties with population testing for BRCA2

It is likely that more of the population within EUROPAC have a *BRCA2* mutation than have been tested. The strict testing criteria for HBOC in the UK advised by the National Institute for Clinical Excellence (NICE) at present do not involve any guidance on when individuals with pancreatic or prostate cancer should be included for consideration of testing (107). Multiple studies have found that pathogenic mutations causing cancer syndromes (including *BRCA2* mutations) can be present without families fitting the strict testing criteria and many of these go on to suggest that all individuals with pancreatic cancer should be tested regardless of family history (72,81,122,123). This however, has not led to a change in practice within the UK. Clearly additional genetic testing within the NHS would carry additional cost, as well as an increase in anxiety in individuals and their families. It is also likely that these individuals would have to consider and possibly undergo significant prophylactic procedures to reduce their cancer risk in the future, when perhaps their risk may not be significant. Multiple studies have discussed that *BRCA2* mutations appears to have significant variation in actual cancer risk (99,132,134,135).

Alongside this are difficulties with barriers to genetic testing amongst those individuals deemed to be at high risk of carrying a pathogenic mutation. Foster and colleagues reported that cancer worry, concern about not receiving screening if result negative, age and travel distance were all barriers to at risk women attending for *BRCA1/2* screening (173). It may be that such families only have a single family member affected with pancreatic cancer and therefore are not eligible for EUROPAC screening without confirmation of a mutation.

11.3.2 Grouping of BRCA2 families

To our knowledge this is the first work describing and comparing families with *BRCA2* mutations by cancer type in this manner. Similar work by Thompson and Easton found that mutations in the central portion of the *BRCA2* gene was linked

with a higher ratio of ovarian to breast cancers however, they did not assess pancreatic cancer prevalence within the families due to small numbers (100). Agreement between the clinicians was reached easily about the *BRCA2* family grouping into HBOC, PC, Neither and Both families. It is expected that the presence of pancreatic cancer within the EUROPAC *BRCA2* group would be higher than that of the general *BRCA2* population in the UK, given the referral criteria for EUROPAC. This means that PC families are likely to be over-represented compared with HBOC and Neither families within the EUROPAC *BRCA2* population.

It was noted that mutation data was available for more families fitting either the PC or Both criteria. Alongside this there were only a relatively small number of families with mutation data available. This could potentially skew the results by over-representing mutation data for these groups. There was noted to be an increase in frameshift mutations in the HBOC and PC groups however, this was not statistically significant. There is no available published evidence collating *BRCA2* mutation type with family cancer type.

The low numbers of ovarian and prostate cancers within the *BRCA2* families meant that further evaluation of these cancers was not possible. It is difficult to draw comparisons between the numbers of prostate and ovarian cancers to known expected numbers contained within published literature. It is likely that prostate cancers within the group are under-reported and, given the spectrum of disease symptoms, may even be undiagnosed (174). The number of prostate cancers reported was 21 (5.38% of male individuals), this is significantly lower than would be expected given the lifetime risk within the general population, which is reported to be 13% (175). The reported lifetime risk of ovarian cancer within the general female population is 2% and this increases to 6-18% in individuals with a *BRCA2* mutation (172,176). It was noted that in the individuals within this study only 15 had a reported diagnosis of ovarian cancer (3.41% of female individuals), this may reflect a lack of reporting to EUROPAC or may reflect the vague symptoms and difficulties with diagnosis in early disease (177). It is important to note that not all of the individuals within these families will have a *BRCA2* mutation (and not all have

been tested), alongside this, as previously discussed, there is known to be variation in cancer risk even amongst those with a pathogenic mutation.

11.3.3 BRCA2 family survival analysis

Firstly, it was noted that there was no significant difference in the ages of those affected by breast or pancreatic cancer between the family types. This suggests that those individuals who are 'at risk' of developing such cancers within these families are affected at the same age regardless of risk. It is already well established that mutation-associated cancers in individuals with *BRCA2* mutations occur at a younger age than in the general population (75,132,178) however, there is no current available work comparing ages of those deemed 'at risk' within *BRCA2* families.

It was expected due to the criteria used to group the families that there would be an increase in pancreatic cancer within the PC families and breast cancer within the HBOC families. There was also noted to be a difference in pancreatic cancer and breast cancer diagnosis. Pancreatic cancer diagnosis was significantly different between the family types with an increase in PC families and reduction in the Neither families. Alongside this breast cancer risk was increased in the HBOC families and decreased in the PC families. This clear separation between the risks of these cancers between the *BCRA2* family types suggests that there are different risk profiles between these distinct groups. This has the potential to have a large impact for individuals with *BRCA2* mutations for the future given at present they are assumed to have similar risks for all mutation-associated cancers. Multiple groups have suggested that other genetic markers may play a role in the variation in cancer risk within individuals and families with *BRCA2* mutations (99,132–135). Further evaluation of these markers may help to more accurately stratify cancer risk within families with *BRCA2* mutations and therefore have far-reaching implications for both cancer screening and prophylaxis, as well as allowing more individualised strategies. In the short term my results have been used to confirm EUROPAC's

current policy of only screening *BRCA2* mutation carriers if there is a pancreatic cancer case in the family.

11.3.4 Factors contributing to survival in BRCA2 families

Data availability for tobacco smoking status and the presence of diabetes mellitus overall was poor (11.93% and 56.75% respectively). This missing data was noted to be in a higher proportion of the Neither individuals than the other families. The lack of data likely represents a lack of reporting from individuals within these families who do not attend for secondary screening with EUROPAC. This is likely to be increased in the Neither families as there may be a reduced perception of cancer risk within these families (due to reduced cancer numbers). Lack of data proportionally in the Neither group could lead to confounding of the results however, this would not explain the disparity in the *BRCA2* groups' survival in the other family types. The family groups were then further analysed to assess whether other factors could account for the survival disparity between the groups.

11.3.4.1 Tobacco smoking

There was no significant difference in the numbers of individuals who smoked tobacco between the *BRCA2* family groups. Alongside this, there was also no significant difference in overall survival from pancreatic or breast cancer between smokers and non-smokers in individuals with *BRCA2* mutations. There continues to be a divide in published literature as to the risk of breast cancer with smoking in individuals with *BRCA2* mutations. Two case-control studies have suggested a moderate increase in risk of breast cancer, especially in individuals under 50 years of age (179,180). However, two similar studies found no difference between the risk of breast cancer in individuals with *BRCA2* mutations who smoke tobacco (181,182). All of the involved studies were case-controlled and therefore subject to significant confounding and recall bias. Interestingly the two studies that matched the case and control population by age and mutation position were those which found no

increase in risk with smoking. There is currently no available published work linking tobacco smoking and *BRCA2* mutation with pancreatic cancer risk.

Lastly, it was noted that when censored for individuals who smoked, the statistically significant difference in survival from both pancreatic and breast cancers between the family types increased. This suggests that differences in tobacco smoking between the family groups does not account for the difference in survival between the four *BRCA2* family types.

11.3.4.2 Diabetes Mellitus

There was no difference in the numbers of individuals with diabetes mellitus (DM) between the family groups and no difference in the overall survival from pancreatic or breast cancer between individuals with and without a diagnosis of DM. There is no current published research assessing the impact of DM alongside a *BRCA2* mutation on development of pancreatic cancer. Bordeleau and colleagues found that there was no increased diagnosis of breast cancer in individuals with *BRCA2* mutations and DM but did find that once a diagnosis of breast cancer was made there was a 2-fold increase in diagnosis of DM within the next 15-year period (183). There were multiple confounding factors, including other causative factors for metabolic syndrome, most noticeably body mass index.

11.4 DNA cleavage & enrichment

Overall CRISPR-Cas9 was chosen as the method of DNA cleavage due to being cost effective and able to specifically target cleavage sites within the genome (184). It also allowed targeted enrichment of specific sites rather than sequencing the whole genome and filtering the data, which would be wasteful and expensive, as well as time consuming (117). Alongside this without DNA cleavage sequencing for this purpose could only be achieved with massively parallel sequencing which would not give us the long haplotype data we need. This method has been used for targeted

enrichment for both large and small segments of DNA, and has recently been used for SNV detection in and around the *BRCA1* gene (117,185). Indeed Bennet-Baker and Mueller found over 100-fold enrichment for their large target DNA segment (185).

Other enrichment methods have been researched. Firstly, Rec-A mediated cleavage has been used in multiple previous studies, although the majority of its use appears to have been replaced by CRISPR-Cas9. Groups have compared CRISPR-Cas9 and Rec-A but only with regards to gene editing (where CRISPR-Cas9 was deemed superior), no comparison of cleavage alone has been performed to date (186). Lauer, Schneider, and Gnirke used Rec-A to select ~1Mb fragments of yeast DNA however, they found that for human DNA the method had a high background and low yield, concluding it should not be used for preparation of DNA (187). They also suggested that the technique was complex and there were often reported difficulties with complete methylation of DNA within agarose plugs.

More recently Stevens and colleagues enriched lambda gDNA with a CRISPR-Cas9 negative enrichment technique (184). This involved attaching CRISPR-Cas9 as previously but then adding exonucleases to digest the unwanted DNA. They found the technique simple and easy to use however, there were concerns about contamination. This technique could be optimised in the future by using PFGE to separate the isolated fragment from the contaminants after exonuclease treatment however, it is likely that there would continue to be difficulties with DNA quantities.

11.4.1 Optimisation of CRISPR-Cas9 oligonucleotides

Firstly, the combinations of oligonucleotides were tested using cultured human fibroblast cell plugs. This cell line was chosen due to being reasonably easy to culture and available within the unit. Unfortunately, due to cell size, it was noted that fewer cells (1×10^6 per plug fibroblasts, 1.5×10^6 per plug PBMCs) than previously thought were able to be placed into agarose cell plugs before the plugs degraded. Initially it was planned to have a cell concentration of 2×10^6 cells per plug. There remains no consensus on the precise cell number required for adequate

DNA quantities, Gabrieli and colleagues used 1 x 10^6 cells per plug of PBMCs for detection of SNVs within the *BRCA1* gene however, their fragment size was considerably smaller (~500kB) (117). Lauer and colleagues used a concentration of 1.5x10^7 lymphoblastoid cells per plug aiming to detect a fragment of up to 2.37Mb on PFGE (187). Lastly, the longest current Nanopore read is 2.27Mb which required 5x10^7 mammalian cells (188). Optimisation of the CRISPR-Cas9 process required DNA be detected by qPCR which is highly sensitive. It is likely that the optimisation process was not affected by the quantity of DNA present however, this could account for some of the difficulties that occurred when sequencing, as discussed later.

The oligonucleotides chosen had the highest efficiency and specificity scores and the lowest off-target scores using The Human Genome Browser (118). This browser has been used previously for CRISPR-Cas9 targeting multiple times and has been sited >37,000 times in previous genomic study publications. The nine pairs of oligonucleotides were then tested. Due to time, cost and cell quantity constraints multiple repeats were not performed.

11.4.2 Optimisation of PFGE settings

Optimisation of PFGE settings was performed using size markers to ensure adequate separation at the required fragment size. In retrospect, it is likely that the first PFGE run had an issue with post-staining or insufficient DNA within the wells given that no visible bands were found. It was noted that the settings used varied slightly from those used in other publications and in the literature with the size markers (185,187). The settings chosen provided sufficient separation at the desired DNA fragment size however, it may have been possible to further optimise these settings given further time and funding. It is important to note that the ladders were compared for each PFGE run once the settings were deemed optimised and there was little difference between the spread of the ladders. There is no evidence that the problems encountered later in the process could be resolved by further

optimisation. Alongside this the size markers were always compared on the specific gel to ascertain the desired size fragment.

There was noted to be speckling which was present in all the PFGE gels. As previously mentioned, multiple techniques were attempted to rectify this including renewing the buffer, changing the agarose and ensuring the gel covered whilst drying. None of these methods resolved this speckling. It was noted that despite this all the PFGE gels appeared to run in a comparable fashion with no effect on the running of the gels, as previously mentioned the size markers were compared for each PFGE experiment. It may be that this speckling would add to contamination of the extracted DNA however, this is difficult to assess.

11.4.3 PFGE after CRISPR-Cas9 treatment

It was noted that there was no visible DNA band at the desired fragment size. This is likely due to small DNA quantities and is consistent with large fragment staining in other published studies. Bennett-Baker and Mueller found no visible band staining for their required fragments (~230Kb and 2.35 Mb) but southern blot revealed evidence of a visible band at the desired size (185). It may be helpful to consider southern blotting in the future after treatment with the CRISPR-Cas9 to fully ensure capture of all the desired DNA. Unfortunately, this would occur at the expense of being able to reuse the DNA for sequencing, therefore would have to be proof that the targeting functions fully, rather than allowing visualisation of the DNA band for each PFGE.

There was also smearing throughout the lanes after treatment with CRISPR-Cas9. Smearing was consistent throughout all PFGE experiments with human DNA in both cell types after treatment with CRISPR-Cas9 and the control (with no CRISPR-Cas9 treatment). This is again consistent with findings from previous publications using animal DNA after treatment to enrich for specific targets (185,187) and did not appear to affect results in previous studies.

Bands at roughly 1.3-2.2Mb were noted in all the PFGE experiments run with human DNA. There is no evidence of published comparable experiments showing similar bands (117,185,187). This was thought to possibly correspond to circular mitochondrial DNA within the gel. It is also possible that these could represent large fragments of genomic DNA however, it seems less likely that there would be significant fragmentation at a single site to cause a visible DNA band in the same position using all the CRISPR-Cas9 oligonucleotides and with the control. The bands seen are at too small a size to be whole human chromosomes. Human mitochondrial DNA is roughly 16.5Kb in size and circular (189). Multiple studies have reported the unpredictable properties of circular DNA during PFGE and how generally progression through the gel seems to be much slower than its linear counterpart (190,191). Alongside this several publications have suggested that larger circular DNA (>100Kb) does not run into the gel and remains in the wells (190,192).

11.4.4 Assessing CRISPR-Cas9 oligonucleotides

Assessment of the oligonucleotide pairs was performed using quantitative polymerase chain reaction (qPCR). This method was used as it allows comparison of quantities of specific DNA sequences, in this case sequences containing part of *BRCA2* gDNA. The qPCR primers were designed specifically to detect *BRCA2* gDNA (rather than cDNA). Multiple repeats of qPCR on the extracted DNA after CRISPR-Cas9 and PFGE were not used as there was limited resources (i.e. DNA quantities).

The visible DNA band within the PFGE was used as the control for the qPCR for each lane. This was to control for differing quantities of DNA from cells placed into each plug. The visible band was used as there was clear, visible evidence of DNA present in this area of the PFGE gel. The majority of the oligonucleotide pairs had some level of enrichment compared with their lane control. The best performing oligonucleotide pair had an amplification threshold reduction of 13.46 (~10,000-fold difference in concentration).

11.5 DNA amplification and sequencing

11.5.1 Rolling-circle amplification

Rolling-circle amplification (RCA) was adapted and optimised for use with the long DNA sequence. Initially it was noted that there was no amplification of the desired segment however, after adding further reagents and longer incubation to counteract the extended section for amplification, there appeared to be at least a 32-fold increase in the product. RCA appeared to function well and the no ligase controls remained negative. These experiments were performed in duplicate and results were similar for repeats. Li and colleagues used a similar technique to amplify DNA prior to nanopore sequencing however, they worked with a much smaller DNA sequence and fragmented their DNA prior to sequencing (average read length 730, maximum read length ~40Kb) (193). They did also find an issue with chimeric reads however, it is extremely unlikely that the same issue was present in this study due to the size of the DNA molecules. Multiple other issues have been suggested with amplification of DNA such as the potential for allelic bias, difficulties with more complex DNA and the possibility of copying contaminating sequences (184). Whilst these issues cannot be ruled out within this study, there remains difficulty with the lack of input DNA which requires amplification before sequencing. Indeed, it remains likely that the failure of sequencing remains due to the lack of input DNA.

It may have been prudent to run further RCA cycles and compare DNA quantities using qPCR and a further PFGE, or other DNA quantification methods. It is unlikely that there would be sufficient DNA of the required length to see a visible band on post-staining of the PFGE gel. It is likely that the quantity of large fragment DNA required to see a visible PFGE band would lead to significant shearing and may indeed be counter-productive however, this may be an avenue for further optimisation in the future. Another possible method of assessment of the integrity and amplification of the DNA sequence would be southern blotting. Southern blot

after PFGE of a gel containing lanes with DNA after different numbers of RCA cycles may allow further optimisation of the amplification process and allow visualisation of not only the required DNA segment but also assess the shearing due to the process (185). Southern blot was not done as the gel used for blotting could not be used for DNA extraction and then sequencing.

11.5.2 DNA quantification

Throughout these experiments there were difficulties with monitoring DNA quantities given the size of the DNA segment required. qPCR was used to ensure the specific fragment was detected within the sample and to assess amplification. As previously discussed, the diluted human DNA sample used as a positive control was diluted to 3.75ng/µL and gave a similar amplification threshold to that of the post-RCA DNA samples. Given only a small proportion of this human sample (<1%) will have been *BRCA2* it is highly likely that the actual quantity of *BRCA2* fragment is several orders of magnitude less than 3.75ng/µL. qPCR does provide some information about the relative quantity of DNA but not specific quantities, therefore the precise amount of the desired fragment both pre and post RCA is unknown.

A UV-Vis spectrophotometer (Nanodrop) was used to assess the samples prior to sequencing and after AMPure XP bead purification. This method is quick, requires low input volumes and provides graphs to assess DNA quality. However, the method is non-specific, open to contamination and gives a poor resolution of DNA quantity below 2ng/µL (194). It is also possible that increased fragmentation within the sample would give an increase in the Nanodrop reading for DNA quantity without any increase in the desired fragment quantity. The NFW 1:1 eluted sample had a Nanodrop detected quantity of DNA of 25.8ng/µL.

All the samples assessed gave a reasonable 260 (DNA) peak. The assessment of bead elution methods using Tris-Cl and water to compare showed all scores below 1.8 for 260/280 score. This suggests presence of contamination with phenol, protein or other DNA extraction substrate (194). 260/230 ratio was >2 for both NFW

samples suggesting a higher quantity of pure DNA. Interestingly qPCR analysis suggested a larger amount of the required fragment in the 1:1 NFW eluted sample, whereas the Nanodrop suggested higher DNA quantity in the 1:1 Tris-Cl eluted sample. Due to the qPCR analysis 1:1 NFW elution was deemed to be the most efficient for elution of the required DNA fragment.

Fluorometric assessment (Qubit) of the DNA was used once the equipment was available. This showed 1.11ng/mL in the post RCA sample, but less than 0.5ng/mL for the other samples. This is clearly several orders of magnitude less than that detected by the Nanodrop method and highlights the clear difficulties of low quantity of target DNA. Qubit analysis is known to be more accurate than UV-Vis DNA detection methods at lower DNA concentrations and less prone to contamination. This method is not specific for the DNA required and has a longer preparation time than the UV-Vis method, it also can be less accurate with lower salt concentrations within the sample (195).

Other DNA quantification methods are available; however, many involve more expensive equipment and increased preparation times. Southern blotting has already been discussed as a possible method to visualise the required DNA band. This method is known to be highly accurate, although does not give an absolute quantified DNA amount (185). Alongside this, using the gel for southern blot would then not allow use of the gel extracted DNA for further experiments and sequencing. This method would provide visual evidence of the DNA target and detect whether RCA leads to a detectable increase in this target fragment. Capillary electrophoresis is a well-known, accurate DNA quantification method however, as yet even when membranes are adapted, there is no evidence of separation for DNA over 600 bases (196).

Overall, it may have been beneficial to complete further RCA cycles as discussed previously with assessment by qPCR, southern blot, Qubit and Nanodrop analysis until the combination of quantification methods showed sufficient of the required DNA fragments. Whilst this would likely take a while to amplify sufficiently, this

would hopefully avoid difficulties with removal from AMPure XP beads and provide sufficient DNA fragments for sequencing.

11.5.3 DNA purification and sequencing

There continued to be ongoing difficulties with purification of the DNA sample with AMPure XP beads even after discussion with the manufacturer. There appeared to be difficulties with elution of the DNA from the beads to then allow sequencing. It is possible, as previously discussed, that this issue may have been less significant with a larger quantity of input DNA however, it is also possible that the beads affinity for longer DNA molecules could have continued to be an issue even with larger DNA quantities. It was noted that after assessment with qPCR, none of the waste steps contained the *BRCA2* DNA fragment, meaning that the beads affinity for the long DNA fragment was ensuring none was washed away.

Contact with the manufacturer led to suggestion of increasing the temperature of the elution, increasing incubation time of the elution, eluting with NFW, reducing the bead: DNA ratio and ensuring all the ethanol was removed completely prior to elution. All of these steps were carried out. Firstly, the temperature was increased, initially from room temperature to 37°C, 55°C then 80°C. Whilst the initial increase to 37°C did show an increase in the read number, further increase in temperature lead to fewer passed reads (to 0 at 80°C) likely due to increased DNA shearing. It was therefore suggested that 37°C is the optimum temperature for bead elution for this fragment whilst maintaining DNA integrity as far as possible. Multiple other studies have used AMPure XP beads for DNA purification with long-read DNA sequences however, usually these have then sheared their DNA, have significantly shorter sequences to start or have used large input DNA quantities (117,197,198).

Elution with NFW did appear to lead to more *BRCA2* sequence found on qPCR after elution and decreasing the ratio of beads to solution to 0.6:1 appeared to not have a significant beneficial effect on elution. Overall, it was suggested that elution with 1:1 NFW was beneficial. There were also concerns that significantly increasing the

elution time to overnight led to significant shearing, and therefore reduced passed reads when sequenced.

Lastly, the two samples which did not use RCA for amplification had a significantly higher maximum read length. This was thought to be due to increased shearing with the RCA steps however, it was noted that none of the sequences when checked for matches to the sequence against the human genome using Blast were human DNA. It is possible that RCA amplified contaminants within the DNA, or that the amplification led to complicated DNA structures due to attempted amplification of the long sequence. Again, further PFGE with southern blotting would help assess the product after RCA.

The longest Nanopore recorded read to date is 2.27Mb however, the protocol used very large quantities of input DNA from e.coli cultures or $5x10^7$ mammalian cells (188). The cells obtained are PBMCs from individual blood samples it would be difficult to obtain larger DNA amounts without significantly increasing the volume of blood taken from each individual. Given an average of $2x10^6$ PBMCs per ml of blood this would require 25ml assuming 100% efficiency with extraction. Alongside this, this does not account for samples collected for contribution to other studies, this would solely be 25ml whole blood collected and processed purely for *BRCA2* testing.

There was noted to be an increased in quantity of DNA on Nanodrop testing after end-repair and dA-tailing with an associated increase in 260/280. Given there is no actual change in DNA quantity it is possible that this increase could be the A-tail addition to the required DNA segment. Without progression to sequencing it is difficult to know if this is the case. The Nanopore 1D ligation kit suggests 1000ng ds DNA is required for sequencing (SQK-LSK109, version: GDE_9063_v109_revD_23May2019) however, this quantity is irrespective of DNA length. It is likely that the actual quantity requirement for longer DNA will be increased due to the MinION sequencing requiring a certain number of DNA strand ends.

11.6 Aim and objectives

The original objectives were:

1. Recruiting individuals at higher than population risk of pancreatic cancer to the IMMPACT study along with blood sample collection from these individuals
2. Co-ordinating the IMMPACT study (including submissions to Health Research Authority (HRA) and ethics committee)
3. To evaluate the population recruited to the EUROPAC study, assessing individuals who are likely to attend for screening and those who are not
4. To create a set of clinical criteria to subdivide families with *BRCA2* mutations to allow stratified analysis of cancer risk
5. To fully analyse whether *BRCA2* mutation families carry different risk profiles for pancreatic and breast cancer
6. To assess possible confounding factors in the relationship between *BRCA2* families and cancer risk
7. To optimise and test a method for enrichment of required genome sequence prior to sequencing
8. To assess the presence of genome SNVs local to *BRCA2* and their presence to further define *BRCA2* groups

The first two objectives were completed with successful recruitment to the IMMPACT study with associated involvement in the multi-national PanFAM-1 study. Co-ordinating the study included contacting and consenting individuals, collecting samples, reviewing scan results and contacting ethics committees for amendments.

The EUROPAC recruited individuals have been evaluated further to assess factors which may impact attendance for secondary screening.

BRCA2 individuals and families were evaluated against a newly developed criteria showing a clear divide in risks of breast and pancreatic cancer between the family

groups. Known other factors were evaluated to assess for possible confounding and it was shown that these other factors were not responsible for the differences in cancer risks between the *BRCA2* family types.

To further evaluate the impact of haplotypes on the cancer risks of *BRCA2* families a method for enrichment was developed for a specific DNA segment including the *BRCA2* gene and surrounding sequences.

Lastly the enriched DNA was sequenced however, this was unsuccessful, suspected due to DNA quantity and difficulties with purification. This would require continuing adaptation and optimisation for the future.

The overall aim was to facilitate an increase in yield of pancreatic cancer cases from screening by evaluating the screened population. This has been accomplished by analysis of the EUROPAC population, recruitment to further studies and further evaluation of the *BRCA2* population recruited to EUROPAC.

12. Conclusion

I have shown in this thesis that targeting of screening on the basis of mutation type is not an efficient way to improve yield from pancreatic cancer screening, although awareness of mutations may increase screening uptake. In conclusion future screening programmes should place greater emphasis on family history as a context for interpreting any identified mutations. They should also include assessing and informing screening participants of a more personalised relative risk.

13. References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394–424.

2. Roch AM, Schneider J, Carr RA, Lancaster WP, House MG, Zyromski NJ, et al. Are BRCA1 and BRCA2 gene mutation patients underscreened for pancreatic adenocarcinoma? J Surg Oncol. 2019;1–7.

3. Yurgelun MB, Chittenden AB, Morales-Oyarvide V, Rubinson DA, Dunne RF, Kozak MM, et al. Germline cancer susceptibility gene variants, somatic second hits, and survival outcomes in patients with resected pancreatic cancer. Genet Med. 2019;21(1):213–23.

4. Walsh N, Zhang H, Hyland PL, Yang Q, Mocci E, Zhang M, et al. Agnostic Pathway/Gene Set Analysis of Genome-Wide Association Data Identifies Associations for Pancreatic Cancer. JNCI J Natl Cancer Inst. 2018;111(6):557–67.

5. Office for National Statistics. Mortality statistics - underlying cause, sex and age [Internet]. Nomis. 2019 [cited 2021 May 14]. Available from: https://www.nomisweb.co.uk/query/construct/submit.asp?forward=yes&menuopt=201&subcomp=

6. Chiaravalli M, Reni M, O'Reilly EM. Pancreatic ductal adenocarcinoma: State-of-the-art 2017 and new therapeutic strategies. Cancer Treat Rev. 2017;60:32–43.

7. Lingala SM, Ghany MGMMhs. Common variation at 2p13.3, 3q29, 7p13 and 17q25.1 associated with susceptibility to pancreatic cancer. Nat Genet. 2016;25(3):289–313.

8. CRUK. Pancreatic cancer survival statistics [Internet]. CRUK. 2017 [cited 2021 May 15]. Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/pancreatic-cancer/survival

9. Neoptolemos J, Kleeff J, Michl J, Costello E, Greenhalf W, Palmer D. Therapeutic developments in pancreatic cancer: current and future perspectives. Nat Rev Gastroenterol Hepatol. 2018;15:333–48.

10. Poruk K, Firpo M, Adler D, Milvihill S. Screening for pancreatic cancer: why, how, and who? Ann Surg. 2013;257:17–26.

11. Peters M, Eckel A, Mueller P, Tramontano A, Weaver D, Lietz A, et al. Progression to pancreatic ductal adenocarcinoma from pancreatic intraepithelial neoplasia: Results of a simulation model. Pancreatology. 2018;18(8):928–34.

12. Tanaka M, Chari S, Adsay V, Fernandez-del Castillo C, Falconi M, Shimizu M, et al. International consensus guidelines for management of intraductal papillary mucinous neoplasms and mucinous cystic neoplasms of the pancreas. Pancreatology. 2006;6(1–2):17–32.

13. Crippa S, Salvia R, Warshaw A, Domínguez I, Bassi C, Falconi M, et al. Mucinous cystic neoplasm of the pancreas is not an aggressive entity: lessons from 163 resected patients. Ann Surg. 2008;247(4):571–9.

14. Distler, M., Aust, D., Weitz, J., Pilarsky, C., & Grützmann R. Precursor lesions for sporadic pancreatic cancer: PanIN, IPMN, and MCN. Biomed Res Int. 2014;2014:474905.

15. Zaccari P, Cardinale V, Severi C, Pedica F, Carpino G, Gaudio E, et al. Common features between neoplastic and preneoplastic lesions of the biliary tract and the pancreas. World J Gastroenterol. 2019;25(31):4343–59.

16. Tanaka M, Fernández-del Castillo C, Adsay V, Chari S, Falconi M, Jang JY, et al. International consensus guidelines 2012 for the management of IPMN and MCN of the pancreas. Pancreatology. 2012;12(3):183–197.

17. McGuigan A, Kelly P, Turkington RC, Jones C, Coleman HG, McCain RS. Pancreatic cancer: A review of clinical diagnosis, epidemiology, treatment and outcomes. World J Gastroenterol. 2018;24(43):4846–61.

18. Wood HE, Gupta S, Kang JY, Quinn MJ, Maxwell JD, Mudan S, et al. Pancreatic cancer in England and Wales 1975-2000: patterns and trends in incidence, survival and mortality. Aliment Pharmacol Ther. 2006;23(8):1205–14.

19. Llach J, Carballal S, Moreira L. Familial Pancreatic Cancer: Current Perspectives. Cancer Manag Res. 2020;12:743–58.

20. Hruban RH, Canto MI, Goggins M, Schulick R, Klein AP. Update on familial pancreatic cancer. Adv Surg. 2010;44:293–311.

21. Wahi MM, Shah N, Schrock CE, Rosemurgy, A. S., 2nd Goldin SB. Reproductive factors and risk of pancreatic cancer in women: a review of the literature. Ann Epidemiol. 2009;19(2):103–11.

22. Sterling J, Garcia MM. Cancer screening in the transgender population: a review of current guidelines, best practices, and a proposed care model. Transl Androl Urol. 2020;9(6):2771–85.

23. Wolpin B, Kraft P, Gross M, Helzlsouer K, Bueno-de-Mesquita H, Steplowski E, et al. Pancreatic cancer risk and ABO blood group alleles: results from the pancreatic cancer cohort consortium. Cancer Res. 2010;70(3):1015–23.

24. Risch HA, Yu H, Lu L, Kidd MS. ABO blood group, Helicobacter pylori seropositivity, and risk of pancreatic cancer: a case-control study. JNCI J Natl Cancer Inst. 2010;102(7):502–5.

25. Arnold LD, Patel A V., Yan Y, Jacobs EJ, Thun MJ, Calle EE, et al. Are racial

disparities in pancreatic cancer explained by smoking and overweight/obesity? Cancer Epidemiol biomarkers Prev. 2009;18(9):2397–405.

26. Pernick NL, Sarkar FH, Philip PA, Arlauskas P, Shields AF, Vaitkevicius VK, et al. Clinicopathologic analysis of pancreatic adenocarcinoma in African Americans and Caucasians. Pancreas. 2003;26(1):28–32.

27. Shindo K, Yu J, Suenaga M, Fesharakizadeh S, Cho C, Macgregor-Das A, et al. Deleterious germline mutations in patients with apparently sporadic pancreatic adenocarcinoma. J Clin Oncol. 2017;35(30):3382–90.

28. Brand RE, Lerch MM, Rubinstein WS, Neoptolemos JP, Whitcomb DC, Hruban RH, et al. Advances in counselling and surveillance of patients at risk for pancreatic cancer. Gut. 2007;56(10):1460–9.

29. Holter S, Borgida A, Dodd A, Grant R, Semotiuk K, Hedley D, et al. Germline BRCA mutations in a large clinic-based cohort of patients with pancreatic adenocarcinoma. J Clin Oncol. 2015;33(28):3124–9.

30. Ben Q, Xu M, Ning X, Liu J, Hong S, Huang W, et al. Diabetes mellitus and risk of pancreatic cancer: A meta-analysis of cohort studies. Eur J Cancer. 2011;47(13):1928–37.

31. Sharma A, Chari ST. Pancreatic cancer and Diabetes Mellitus. Curr Treat Options Gastroenterol. 2018;16(4):466–78.

32. Andersen DK, Korc M, Petersen GM, Eibl G, Li D, Rickels MR, et al. Diabetes, Pancreatogenic Diabetes, and Pancreatic Cancer. Diabetes. 2017;66(5):1103–10.

33. Singhi AD, Koay EJ, Chan ST, Maitra A. Early detection of pancreatic cancer: Opportunities and challenges. Gastroenterology. 2019;156:2024–40.

34. Diabetes.co.uk. Diabetes prevalence [Internet]. Diabetes.co.uk. 2018 [cited 2021 Jun 16]. p. 106–7. Available from: https://www.diabetes.co.uk/diabetes-prevalence.html

35. LCTC. UK Early Detection Initiative for Pancreatic Cancer [Internet]. LCTC. 2020 [cited 2021 Jun 16]. Available from: https://www.lctc.org.uk/research/uk-edi/

36. Immunovia. PanDIA1 [Internet]. 2020 [cited 2021 Apr 25]. Available from: https://immunovia.com/immray-pancan-d/prospective-studies-and-next-step/pandia-1/

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

38. Iodice S, Gandini S, Maisonneuve P, Lowenfels AB. Tobacco and the risk of pancreatic cancer: a review and meta-analysis. Langenbeck's Arch Surg.

2008;393(4):535–45.

39. Yadav D, Lowenfels AB. The epidemiology of pancreatitis and pancreatic cancer. Gastroenterology. 2013;144(6):1252–61.

40. Weissman S, Takakura K, Eibl G, Pandol SJ, Saruta M. The Diverse Involvement of Cigarette Smoking in Pancreatic Cancer Development and Prognosis. Pancreas. 2020;49(5):612–20.

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

42. NICE. Stop smoking interventions and services: NICE guideline [NG92] [Internet]. NICE guidelines. 2018 [cited 2021 May 27]. Available from: https://www.nice.org.uk/guidance/ng92/chapter/Recommendations#ftn.footnote_4

43. McNeill A, Brose L, Calder R, Bauld L, Robson D. Evidence review of e-cigarettes and heated tobacco products 2018: executive summary. A report commissioned by Public Health England. 2018.

44. Rohrmann S, Linseisen J, Vrieling A, Boffetta P, Stolzenberg-Solomon RZ, Lowenfels AB, et al. Ethanol intake and the risk of pancreatic cancer in the European Prospective Investigation into Cancer and Nutrition (EPIC). Cancer Causes Control. 2009;20(5):785–94.

45. Coughlin SS, Calle EE, Patel A V., Thun MJ. Predictors of pancreatic cancer mortality among a large cohort of United States adults. Cancer Causes Control. 2000;11(10):915–23.

46. Heuch I, Kvåle G, Jacobsen BK, Bjelke E. Use of alcohol, tobacco and coffee, and risk of pancreatic cancer. Br J Cancer. 1983;48(5):637–43.

47. Genkinger JM, Spiegelman D, Anderson KE, Bergkvist L, Bernstein L, van den Brandt PA, et al. Alcohol intake and pancreatic cancer risk: a pooled analysis of fourteen cohort studies. Cancer Epidemiol biomarkers Prev. 2009;18(3):765–76.

48. Aune D, Greenwood DC, Chan DS, Vieira R, Vieira AR, Navarro Rosenblatt DA, et al. Body mass index, abdominal fatness and pancreatic cancer risk: a systematic review and non-linear dose-response meta-analysis of prospective studies. Ann Oncol. 2012;23(4):843–52.

49. Zhou B, Wu D, Liu H, Du LT, Wang YS, Xu JW, et al. Obesity and pancreatic cancer: An update of epidemiological evidence and molecular mechanisms. Pancreatology. 2019;19(7):941–50.

50. Anderson K, Sinha R, Kulldorff M, Gross M, Lang N, Barber C, et al. Meat intake and cooking techniques: associations with pancreatic cancer. Mutat Res. 2002;506:225–31.

51. Anderson K, Kadlubar F, Kulldorff M, Harnack L, Gross M, Lang N, et al. Dietary intake of heterocyclic amines and benzo(a)pyrene: associations with pancreatic cancer. Cancer Epidemiol biomarkers Prev. 2005;14(9):2261–5.

52. Li D, Day R, Bondy M, Sinha R, Nguyen N, Evans D, et al. Dietary mutagen exposure and risk of pancreatic cancer. Cancer Epidemiol Biomarkers Prev. 2007;16(4):655–61.

53. Stolzenberg-Solomon R, Cross A, Silverman D, Schairer C, Thompson F, Kipnis V, et al. Meat and meat-mutagen intake and pancreatic cancer risk in the NIH-AARP cohort. Cancer Epidemiol Biomarkers Prev. 2007;16:2664–75.

54. Larsson SC, Bergkvist L, Wolk A. Consumption of sugar and sugar-sweetened foods and the risk of pancreatic cancer in a prospective study. Am J Clin Nutr. 2006;84(5):1171–6.

55. Mueller NT, Odegaard A, Anderson K, Yuan JM, Gross M, Koh WP, et al. Soft drink and juice consumption and risk of pancreatic cancer: the Singapore Chinese Health Study. Cancer Epidemiol Biomarkers Prev. 2010;19(2):447–55.

56. Schernhammer ES, Hu FB, Giovannucci E, Michaud DS, Colditz GA, Stampfer MJ, et al. Sugar-sweetened soft drink consumption and risk of pancreatic cancer in two prospective cohorts. Cancer Epidemiol Biomarkers Prev. 2005;14(9):2098–105.

57. Risch HA. Etiology of pancreatic cancer, with a hypothesis concerning the role of N-nitroso compounds and excess gastric acidity. J Natl Cancer Inst. 2003;95(13):948–60.

58. Chen XZ, Schöttker B, Castro FA, Chen H, Zhang Y, Holleczek B, et al. Association of helicobacter pylori infection and chronic atrophic gastritis with risk of colonic, pancreatic and gastric cancer: A ten-year follow-up of the ESTHER cohort study. Oncotarget. 2016;7(13):17182–93.

59. Vincent A, Herman J, Schulick R, Hruban RH, Goggins M. Pancreatic cancer. Lancet. 2011;378(9791):607–20.

60. Keane MG, Horsfall L, Rait G, Pereira SP. A case-control study comparing the incidence of early symptoms in pancreatic and biliary tract cancer. BMJ Open. 2014;4(11):e005720.

61. Cheng Y, Briarava M, Lai M, Wang X, Tu B, Cheng N, et al. Pancreaticojejunostomy versus pancreaticogastrostomy reconstruction for the prevention of postoperative pancreatic fistula following pancreaticoduodenectomy. Cochrane Database Syst Rev. 2017;9:CD012257.

62. Pericleous S, Middleton N, McKay S, Bowers K, Hutchins R. Systematic review and meta-analysis of case-matched studies comparing open and laparoscopic distal pancreatectomy: is it a safe procedure? Pancreas. 2012;41:993–1000.

63. Mollberg N, Rahbari N, Koch M, Hartwig W, Hoeger Y, Büchler M, et al.

Arterial resection during pancreatectomy for pancreatic cancer. Ann Surg. 2011;254:882–93.

64. Roth MT, Cardin DB, Berlin JD. Recent advances in the treatment of pancreatic cancer. F1000Research. 2020;9(F1000 Faculty Rev):131.

65. Conroy T, Hammel P, Hebbar M, Ben Abdelghani M, Wei AC, Raoul JL, et al. FOLFIRINOX or Gemcitabine as Adjuvant Therapy for Pancreatic Cancer. N Engl J Med. 2018;379(25):2395–406.

66. Ghosn M, Kourie HR, El Rassy E, Haddad FG, Hanna C, El Karak F, et al. Where does chemotherapy stands in the treatment of ampullary carcinoma? A review of literature. World J Gastrointest Oncol. 2016;8(10):745–50.

67. Casolino R, Braconi C, Malleo G, Paiella S, Bassi C, Milella M, et al. Reshaping preoperative treatment of pancreatic cancer in the era of precision medicine. Ann Oncol. 2021;32(2):183–96.

68. Zhan HX, Xu JW, Wu D, Wu ZY, Wang L, Hu SY, et al. Neoadjuvant therapy in pancreatic cancer: a systematic review and meta-analysis of prospective studies. Cancer Med. 2017;6(6):1201–19.

69. Janssen QP, Dam JL van, Bonsing BA, Bos H, Bosscha KP, Coene PPLO, et al. Total neoadjuvant FOLFIRINOX versus neoadjuvant gemcitabine-based chemoradiotherapy and adjuvant gemcitabine for resectable and borderline resectable pancreatic cancer (PREOPANC-2 trial): study protocol for a nationwide multicenter randomized controlled trial. BMC Cancer. 2021;21:300.

70. Bilimoria KY, Bentrem DJ, Ko CY, Ritchey J, Stewart AK, Winchester DP, et al. Validation of the 6th edition AJCC pancreatic cancer staging system. Cancer. 2007;110(4):738–44.

71. Canto MI, Harinck F, Hruban RH, Offerhaus GJ, Poley JW, Kamel I, et al. International cancer of the pancreas screening (CAPS) consortium summit on the management of patients with increased risk for familial pancreatic cancer. Gut. 2013;62(3):339–47.

72. Hu C, Hart SN, Polley EC, Gnanaolivu R, Shimelis H, Lee KY, et al. Association between inherited germline mutations in cancer predisposition genes and risk of pancreatic cancer. JAMA - J Am Med Assoc. 2018;319(23):2401–9.

73. Bartsch DK, Gress TM, Langer P. Familial pancreatic cancer-current knowledge. Nat Rev Gastroenterol Hepatol. 2012;9(8):445–53.

74. Goggins M, Overbeek KA, Brand R, Syngal S, Del Chiaro M, Bartsch DK, et al. Management of patients with increased risk for familial pancreatic cancer: updated recommendations from the International Cancer of the Pancreas Screening (CAPS) Consortium. Gut. 2020;69(1):7–17.

75. Toss A, De Matteis E, Cortesi L, Venturelli M, Tagliafico E, Barbieri E, et al.

Hereditary Pancreatic Cancer: A Retrospective Single-Center Study of 5143 Italian Families with History of BRCA-Related Malignancies. Cancers (Basel). 2019;11(2):193.

76. Hearle N, Schumacher V, Menko FH, Olschwang S, Boardman LA, Gille JJ, et al. Frequency and spectrum of cancers in the Peutz-Jeghers syndrome. Clin cancer Res. 2006;12(10):3209–15.

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

78. van Lier, M. G., Wagner, A., Mathus-Vliegen, E. M., Kuipers, E. J., Steyerberg, E. W., & van Leerdam ME. High cancer risk in Peutz-Jeghers syndrome: a systematic review and surveillance recommendations. Am J Gastroenterol. 2010;105(6):1258–65.

79. Rebours V, Levy P, Ruszniewski P. An overview of hereditary pancreatitis. Dig liver Dis . 2012;44:8–15.

80. Rebours V, Boutron-Ruault MC, Schnee M, Férec C, Le Maréchal C, Hentic O, et al. The natural history of hereditary pancreatitis: a national series. Gut. 2009;58(1):97–103.

81. Ohmoto A, Yachida S, Morizane C. Genomic features and clinical management of patients with hereditary pancreatic cancer syndromes and familial pancreatic cancer. Int J Mol Sci. 2019;20:561.

82. Antwi SO, Fagan SE, Chaffee KG, Bamlet WR, Hu C, Polley EC, et al. Risk of Different Cancers Among First-degree Relatives of Pancreatic Cancer Patients: Influence of Probands' Susceptibility Gene Mutation Status. JNCI J Natl Cancer Inst. 2019;111(3):1–8.

83. Lu C, Xu CF, Wan XY, Zhu HT, Yu CH, Li YM. Screening for pancreatic cancer in familial high-risk individuals: A systematic review. World J Gastroenterol. 2015;21(28):8678–86.

84. Pereira SP, Oldfield L, Ney A, Hart PA, Keane MG, Pandol SJ, et al. Early detection of pancreatic cancer. lancet Gastroenterol Hepatol. 2020;5(7):698–710.

85. Immunovia. Immunovia's PanFAM-1 Prospective Study Gathers Over 3000 Familial Hereditary Pancreatic Cancer Risk Samples And Will Be Analysed In The Second Half Of 2021 [Internet]. 2021 [cited 2021 Apr 25]. Available from: https://immunovia.com/investorfeed/immunovias-panfam-1-prospective-study-gathers-over-3000-familial-hereditary-pancreatic-cancer-risk-samples-and-will-be-analysed-in-the-second-half-of-2021/

86. Vasen H, Ibrahim I, Ponce CG, Slater EP, Matthäi E, Carrato A, et al. Benefit of Surveillance for Pancreatic Cancer in High-Risk Individuals: Outcome of Long-Term Prospective Follow-Up Studies From Three European Expert Centers. J

Clin Oncol. 2016;34(17):2010–9.

87. Canto MI, Almario JA, Schulick RD, Yeo CJ, Klein A, Blackford A, et al. Risk of neoplastic progression in individuals at high risk for pancreatic cancer undergoing long-term surveillance. Gastroenterology. 2018;155(3):740–51.

88. Corral, J. E., Mareth, K. F., Riegert-Johnson, D. L., Das, A., & Wallace MB. Diagnostic Yield From Screening Asymptomatic Individuals at High Risk for Pancreatic Cancer: A Meta-analysis of Cohort Studies. Clin Gastroenterol Hepatol. 2019;17(1):41–53.

89. Sheel ARG, Harrison S, Sarantitis I, Nicholson JA, Hanna T, Grocock C, et al. Identification of Cystic Lesions by Secondary Screening of Familial Pancreatic Cancer (FPC) Kindreds Is Not Associated With the Stratified Risk of Cancer. Am J Gastroenterol. 2019;114(1):155–64.

90. EUROPAC. About EUROPAC [Internet]. 1997 [cited 2021 Jul 13]. Available from: https://www.lctu.org.uk/LCTU_NET/Frontend/Default.aspx?Data=W1tRMjl1ZEdWdWRFbEVdXVtNakF4XVtbYkc5allXeGxdXVtPUT09XQ%3D%3D

91. Wooster R, Bignell G, Lancaster J, Swift S, Seal S, Mangion J, et al. Identification of the breast cancer susceptibility gene BRCA2. Nature. 1995;378:789–92.

92. Gayther SA, Mangion J, Russell P, Seal S, Barfoot R, Ponder BAJ, et al. Variation of risks of breast and ovarian cancer associated with different germline mutations of the BRCA2 gene. Nat Genet. 1997;15(1):103–5.

93. Fradet-Turcotte A, Sitz J, Grapton D, Orthwein A. BRCA2 functions: From DNA repair to replication fork stabilization. Endocr Relat Cancer. 2016;23(10):T1–17.

94. Hoang LN, Gilks BC. Hereditary Breast and Ovarian Cancer Syndrome: Moving beyond BRCA1 and BRCA2. Adv Anat Pathol. 2018;25(2):85–95.

95. Schlacher K, Christ N, Siaud N, Egashira A, Wu H, Jasin M. Double-Strand Break Repair Independent Role For BRCA2 In Blocking Stalled Replication Fork Degredation By MRE11. Cell. 2011;145(4):529–42.

96. Mao Z, Bozzella M, Seluanov A, Gorbunova V. Comparison of nonhomologous end joining and homologous recombination in human cells. DNA Repair (Amst). 2008;7(10):1765–71.

97. Roy R, Chun J, Powell SN. BRCA1 and BRCA2: Different roles in a common pathway of genome protection. Nat Rev Cancer. 2016;12(1):68–78.

98. Mavaddat N, Peock S, Frost D, Ellis S, Platte R, Fineberg E, et al. Cancer risks for BRCA1 and BRCA2 mutation carriers: Results from prospective analysis of EMBRACE. J Natl cancer Inst. 2013;105(11):812–22.

99. Barnes DR, Antoniou AC. Unravelling modifiers of breast and ovarian cancer

risk for BRCA1 and BRCA2 mutation carriers: Update on genetic modifiers. J Intern Med. 2012;271(4):331–43.

100. Thompson D, Easton D. Variation in cancer risks, by mutation position, in BRCA2 mutation carriers. Am J Hum Genet. 2001;68:410–9.

101. Michailidou K, Beesley J, Lindstrom S, Canisius S, Dennis J, Lush M, et al. Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer. Nat Genet. 2015;47(4):373–80.

102. Kuchenbaecker KB, Neuhausen SL, Robson M, Barrowdale D, McGuffog L, Mulligan AM, et al. Associations of common breast cancer susceptibility alleles with risk of breast cancer subtypes in BRCA1 and BRCA2 mutation carriers. Breast Cancer Res. 2014;16:3416.

103. Kuchenbaecker KB, Ramus SJ, Tyrer J, Lee A, Shen HC, Beesley J, et al. Identification of six new susceptibility loci for invasive epithelial ovarian cancer. Nat Genet. 2015;47(2):164–71.

104. Antoniou AC, Beesley J, Mcguffog L, Sinilnikova OM, Neuhausen SL, Ding YC, et al. Common breast cancer susceptibility alleles and the risk of breast cancer for BRCA1 and BRCA2 mutation carriers : implications for risk prediction. Cancer Res. 2010;70(23):9742–54.

105. Ryan DP, Mamon H. Treatment for potentially resectable exocrine pancreatic cancer [Internet]. Up-to-Date. 2018 [cited 2021 Jul 15]. Available from: https://www.uptodate.com/contents/treatment-for-potentially-resectable-exocrine-pancreatic-cancer

106. Golan T, Hammel P, Reni M, Van Cutsem E, Macarulla T, Hall MJ, et al. Maintenance Olaparib for Germline BRCA-Mutated Metastatic Pancreatic Cancer. N Engl J Med. 2019;381(4):317–27.

107. NICE. Familial breast cancer : classification , care and managing breast cancer and related risks in people with a family history of breast cancer (NGC164): Last updated 2019. Natl institue Heal Care Excell UK [Internet]. 2013;(June):48. Available from: https://www.nice.org.uk/guidance/cg164/resources/familial-breast-cancer-classification-care-and-managing-breast-cancer-and-related-risks-in-people-with-a-family-history-of-breast-cancer-35109691767493

108. Chowdhury S, Dent T, Pashayan N, Hall A, Lyratzopoulos G, Hallowell N, et al. Incorporating genomics into breast and prostate cancer screening: Assessing the implications. Genet Med. 2013;15(6):423–32.

109. Serrano R, Brunet J, Lastra E, Llort G, Chirivella I, Graña B, et al. SEOM clinical guidelines in Hereditary Breast and ovarian cancer. Clin Transl Oncol. 2015;17(12):956–61.

110. Kanzi AM, San JE, Chimukangara B, Wilkinson E, Fish M, Ramsuran V, et al. Next Generation Sequencing and Bioinformatics Analysis of Family Genetic

Inheritance. Front Genet. 2020;11:544162.

111. Heather JM, Chain B. The sequence of sequencers: The history of sequencing DNA. Genomics. 2016;107(1):1–8.

112. Watson M, Warr A. Errors in long-read assemblies can critically affect protein prediction. Nat Biotechnol. 2019;37:124–126.

113. Salk JJ, Schmitt MW, Loeb LA. Enhancing the accuracy of next-generation sequencing for detecting rare and subclonal mutations. Nat Rev Genet. 2018;19(5):269–85.

114. Weirather JL, Cesare M de, Wang Y, Piazza P, Sebastiano V, Wang X-J, et al. Comprehensive comparison of Pacific Biosciences and Oxford Nanopore Technologies and their applications to transcriptome analysis. F1000Research. 2017;6(1):100.

115. Koren S, Phillippy AM, Simpson JT, Loman NJ, Loose M. Reply to 'Errors in long-read assemblies can critically affect protein prediction.' Nat Biotechnol. 2019;37:127–128.

116. National Institute for Health and Care Excellence. Pancreatic cancer in adults: diagnosis and management [Internet]. 2018 [cited 2020 Jun 15]. Available from: https://www.nice.org.uk/guidance/ng85/chapter/Recommendations#people-with-inherited-high-risk-of-pancreatic-cancer

117. Gabrieli T, Sharim H, Fridman D, Arbib N, Michaeli Y, Ebenstein Y. Selective nanopore sequencing of human BRCA1 by Cas9-assisted targeting of chromosome segments (CATCH). Nucleic Acids Res. 2018;46(14):e87.

118. Kent W, Sugnet C, Furey T, Roskin K, Pringle T, Zahler A, et al. The human genome browser at UCSC. Genom Res. 2002. p. 996-1006e.

119. NIH. BLAST [Internet]. US National Library of Medicine. 2005 [cited 2019 Apr 16]. Available from: https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome

120. Grocock C, Vitone L, Harcus M, Neoptolemos J, Raraty M, Greenhalf W. Familial pancreatic cancer: a review and latest advances. Adv Med Sci. 2007;52:37–49.

121. Klein A. Identifying people at a high risk of developing pancreatic cancer. Nat Rev Cancer. 2013;13:66–74.

122. Zhen D, Rabe K, Gallinger S, Syngal S, Schwartz AG, Goggins MG, et al. BRCA1, BRCA2, PALB2 and CDKN2A mutations in familial pancreatic cancer: a PACGENE study. Genet Med. 2015;17(7):569–77.

123. Chaffee KG, Oberg AL, McWilliams RR, Majithia N, Allen BA, Kidd J, et al. Prevalence of germ-line mutations in cancer genes among pancreatic cancer

patients with a positive family history. Genet Med. 2018;20(1):119–27.

124. Hu C, Hart SN, Bamlet WR, Moore RM, Nandakumar K, Eckloff BW, et al. Prevalence of Pathogenic Mutations in Cancer Predisposition Genes among Pancreatic Cancer Patients. Cancer Epidemiol Biomarkers Prev. 2016;25(1):207–11.

125. Li Y, Bian X, Wei S, He M, Yang Y. The relationship between pancreatic cancer and type 2 diabetes: cause and consequence. Cancer Manag Res. 2019;11:8257–68.

126. Maitra A, Sharma A, Brand RE, Eeden SK Van Den, Fisher WE, Hart PA, et al. A Prospective Study to Establish a New-Onset Diabetes (NOD) Cohort: From the Consortium for the Study of Chronic Pancreatitis, Diabetes, and Pancreatic Cancer. Pancreas. 2018;47(10):1244–8.

127. Naudin S, Viallon V, Hashim D, Freisling H, Jenab M, Weiderpass E, et al. Healthy Lifestyle and the Risk of Pancreatic Cancer in the EPIC Study. Eur J Epidemiol. 2019;Online ahead of print.

128. Ilic M, Ilic I. Epidemiology of pancreatic cancer. World J Gastroenterol. 2016;22(44):9694–705.

129. Schneider R, Slater E, Sina M, Habbe N, Fendrich V, Matthäi E, et al. German national case collection for familial pancreatic cancer (FaPaCa): ten years experience. Fam Cancer. 2011;10:323–30.

130. Mocci E, Guillen-Ponce C, Earl J, Marquez M, Solera J, Salazar-López M, et al. PanGen-Fam: Spanish registry of hereditary pancreatic cancer. Eur J Cancer. 2015;51:1911–1917.

131. Iqbal J, Ragone A, Lubinski J, Lynch HT, Moller P, Ghadirian P, et al. The incidence of pancreatic cancer in BRCA1 and BRCA2 mutation carriers. Br J Cancer. 2012;107(12):2005–9.

132. Begg CB, Haile RW, Borg A, Malone KE, Concannon P, Thomas DC, et al. Variation of Breast Cancer Risk Among BRCA1/2 Carriers. JAMA - J Am Med Assoc. 2008;299(2):194–201.

133. Antoniou A, Pharoah P, Narod S, Al. E. Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case Series unselected for family history: a combined analysis of 22 studies. Am J Hum Genet. 2003;72(5):1117–30.

134. Mitchell R, Buckingham L, Cobleigh M, Rotmensch J, Burgess K, Usha L. Can chimerism explain breast/ovarian cancers in BRCA non-carriers from BRCA-positive families? PLoS One. 2018;13(4):e0195497.

135. Kuchenbaecker KB, McGuffog L, Barrowdale D, Lee A, Soucy P, Dennis J, et al. Evaluation of polygenic risk scores for breast and ovarian cancer risk prediction in BRCA1 and BRCA2 mutation carriers. J Natl Cancer Inst.

2017;109(7):1–15.

136. Triplette M, Thayer JH, Pipavath SN, Crothers K. Poor Uptake of Lung Cancer Screening: Opportunities for Improvement. J Am Coll Radiol. 2019;16:446–50.

137. Symonds EL, Cock C, Meng R, Cole SR, Fraser RJL, Young GP. Uptake of a colorectal cancer screening blood test in people with elevated risk for cancer who cannot or will not complete a faecal occult blood test. Eur J cancer Prev. 2018;27(425–432).

138. England PH. Cervical screening standards data report 1 April 2018 to 31 March 2019 [Internet]. 2020. Available from: https://www.gov.uk/government/publications/cervical-screening-standards-report-2018-to-2019

139. NHS Digital. Breast screening programme, England 2018-9 [Internet]. 2020 [cited 2020 Jun 15]. Available from: https://digital.nhs.uk/data-and-information/publications/statistical/breast-screening-programme/england---2018-19

140. Eckardt V, Kanzler G, Barnhard G. Investigation, Life expectancy and cancer risk in patients with Barrett's oesophagus: a ptospective controlled. Am J Med. 2001;111:33–7.

141. Gatenby P, Caygill C, Wall C, Bhatacharjee S, Ramus J, Watson A, et al. Lifetime risk of esophageal adenocarcinoma in patients with Barrett's oesophagus. World J Gastroenterol. 2014;20(28):9611–7.

142. 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.

143. Logan RFA, Patnick J, Nickerson C, Coleman L, Rutter MD, von Wagner C. Outcomes of the Bowel Cancer Screening Programme (BCSP) in England after the first 1 million tests. Gut. 2012;61(10):1439–46.

144. Vrinten C, Gallagher A, Waller J, Marlow L. Cancer Stigma and Cancer Screening Attendance: A Population Based Survey in England. BMC Cancer. 2019;19(1):566.

145. Blanks RG, Benson VS, Alison R, Brown A, Reeves GK, Beral V, et al. Nationwide bowel cancer screening programme in England: cohort study of lifestyle factors affecting participation and outcomes in women. Br J Cancer. 2015;112(9):1562–1567.

146. Maheswaran R, Pearson T, Jordan H, Black D. Socioeconomic deprivation, travel distance, location of service, and uptake of breast cancer screening in North Derbyshire, UK. J Epidemiol Community Heal. 2006;60(3):208–12.

147. Wang C. The impact of car ownership and public transport usage on cancer screening coverage: Empirical evidence using a spatial analysis in England. J

Transp Geogr. 2016;56:15–22.

148. Vrinten C, Waller J, Wagner C von, Wardle J. Cancer fear: facilitator and deterrent to participation in colorectal cancer screening. Cancer Epidemiol Biomarkers Prev. 2015;24(2):400–5.

149. Consedine NS, Magai C, Krivoshekova YS, Ryzewicz L, Neugut AI. Fear, Anxiety, Worry, and Breast Cancer Screening Behavior: A Critical Review. Cancer Epidemiol Biomarkers Prev. 2004;13(4):501–10.

150. Andersen SW, Blot WJ, Lipworth L, Steinwandel M, Murff HJ, Zheng W. Association of Race and Socioeconomic Status With Colorectal Cancer Screening, Colorectal Cancer Risk, and Mortality in Southern US Adults. JAMA Netw Open. 2019;2(12):e1917995.

151. Cooper GS, Kou TD, Dor A, Koroukian SM, Schluchter MD. Cancer preventive services, socioeconomic status, and the Affordable Care Act. Cancer. 2017;123(9):1585–9.

152. Watanabe G, Ishida T. Increased centrosome number in BRCA -related breast cancer specimens determined by immunofluorescence analysis. 2018;(March):2027–35.

153. Maheswaran R, Pearson T, Jordan H, Black D. Socioeconomic deprivation, travel distance, location of service, and uptake of breast cancer screening in North Derbyshire, UK. J Epidemiol Community Heal. 2006;60:208–12.

154. Karpinskyj C, Burnell M, Gonzalez-Izquierdo A, Ryan A, Kalsi J, Jacobs I, et al. Socioeconomic Status and Ovarian Cancer Stage at Diagnosis: A Study Nested Within UKCTOCS. Diagnostics. 2020;10(2):89.

155. Thobie A, Mulliri A, Dolet N, Eid Y, Bouvie V, Launoy G, et al. Socioeconomic status impacts survival and access to resection in pancreatic adenocarcinoma: A high-resolution population-based cancer registry study. Surg Oncol. 2018;27(4):759–66.

156. Whitaker KL, Good A, Miles A, Robb K, Wardle J, Wagner C von. Socioeconomic inequalities in colorectal cancer screening uptake: does time perspective play a role? Heal Psychol. 2011;30(6):702–9.

157. Han S, Pate B, Min M, Bocelli L, Kheder J, Wachholtz A, et al. Quality of life comparison between smokers and non-smokers with chronic pancreatitis. Pancreatology. 2018;18(3):269–74.

158. Malmqvist J, Siersma V, Thorsen H, Heleno B, Rasmussen JF, Brodersen J. Did psychosocial status, sociodemographics and smoking status affect non-attendance in control participants in the Danish Lung Cancer Screening Trial? A nested observational study. BMJ Open. 2020;10(2):e030871.

159. Richard A, Rohrmann S, Schmid S, Tirri B, Huang D, Güth U, et al. Lifestyle and health-related predictors of cervical cancer screening attendance in a Swiss

population-based study. Cancer Epidemiol. 2015;39(6):870–6.

160. Office for National Statistics. Adult smoking habits in the UK: 2019 [Internet]. 2020 [cited 2021 Apr 24]. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthandlifeexpectancies/bulletins/adultsmokinghabitsingreatbritain/2019#adult-smoking-habits-in-the-uk-data

161. NHS Digital. Health Survey for England 2018: Adult's health-related behaviours [Internet]. NHS digital. 2019 [cited 2021 Apr 24]. Available from: https://files.digital.nhs.uk/A3/2F03E4/HSE18-Adult-Health-Related-Behaviours-rep-v2.pdf

162. Lagerlund M, Drake I, Wirfält E, Sontrop J, Zackrisson S. Health-related lifestyle factors and mammography screening attendance in a Swedish cohort study. Eur J Cancer Prev. 2015;24(1):44–50.

163. Brand R, Borazanci E, Speare V, Dudley B, Karloski E, Peters MLB, et al. Prospective study of germline genetic testing in incident cases of pancreatic adenocarcinoma. Cancer. 2018;124(17):3520–7.

164. Grant R, Selander I, Connor A, Selvarajah S, Borgida A, Briollais L, et al. Prevalence of Germline Mutations in Cancer Predisposition Genes in Patients with Pancreatic Cancer. Gastroenterology. 2015;148(3):556–64.

165. McWilliams RR, Wieben ED, Rabe KG, Pedersen KS, Wu Y, Sicotte H, et al. Prevalence of CDKN2A mutations in pancreatic cancer patients: implications for genetic counseling. Eur J Hum Genet. 2011;19:472–8.

166. Zhen DB, Rabe KF, Gallinger S, Syngal S, Schwartz AG, Goggins MG, et al. BRCA1, BRCA2, PALB2, and CDKN2A Mutations in Familial Pancreatic Cancer (FPC): A PACGENE Study. Genet Med. 2015;17(7):569–77.

167. Jervis S, Song H, Lee A, Dicks E, Harrington P, Baynes C, et al. A risk prediction algorithm for ovarian cancer incorporating BRCA1, BRCA2, common alleles and other familial effects. J Med Genet. 2015;52(7):465–75.

168. Hart PA, Bellin MD, Andersen DK, Bradley D, Cruz-Monserrate Z, Forsmark CE, et al. Type 3c (pancreatogenic) diabetes mellitus secondary to chronic pancreatitis and pancreatic cancer. lancet Gastroenterol Hepatol. 2016;1(3):226–237.

169. Malka D, Hammel P, Sauvanet A, Rufat P, O'Toole D, Bardet P, et al. Risk factors for diabetes mellitus in chronic pancreatitis. Gastroenterology. 2000;119(5):1324–32.

170. Wang W, Guo Y, Liao Z, Zou DW, Jin ZD, Zou DJ, et al. Occurrence of and risk factors for diabetes mellitus in Chinese patients with chronic pancreatitis. Pancreas. 2011;40(2):206–12.

171. Banks E, Beral V, Cameron R, Hogg A, Langley N, Barnes I, et al. Comparison

of various characteristics of women who do and do not attend for breast cancer screening. Breast cancer Res. 2002;4(1):R1.

172. Teixeira N, Hout A Van Der, Oosterwijk JC, Vos JR, Devilee HP, Engelen K Van, et al. The association between cancer family history and ovarian cancer risk in BRCA1/2 mutation carriers : can it be explained by the mutation position? Eur J Hum Genet. 2018;848–57.

173. Foster C, Evans DG, Eeles R, Eccles D, Ashley S, Brooks L, et al. Non-uptake of predictive genetic testing for BRCA1/2 among relatives of known carriers: attributes, cancer worry, and barriers to testing in a multicenter clinical cohort. Genet Test. 2004;8(1):23–9.

174. Litwin MS, Tan HJ. The Diagnosis and Treatment of Prostate Cancer: A Review. JAMA. 2017;317(24):2532–42.

175. US Preventive Services Task Force, Grossman DC, Curry SJ, Owens DK, Bibbins-Domingo, K., Caughey AB, Davidson KW, et al. Screening for Prostate Cancer: US Preventive Services Task Force Recommendation Statement. JAMA. 2018;319(18):1901–13.

176. Gupta KK, Gupta VK, Naumann RW. Ovarian cancer: screening and future directions. Int J Gynecol cancer. 2019;29(1):195–200.

177. Stewart C, Ralyea C, Lockwood S. Ovarian Cancer: An Integrated Review. Semin Oncol Nurs. 2019;35(2):151–6.

178. Kim D, Crawford B, Ziegler J, Beattie M. Prevalence and characteristics of pancreatic cancer in families with BRCA1 and BRCA2 mutations. Fam Cancer. 2009;8(2):153–8.

179. Ko KP, Kim SJ, Huzarski T, Gronwald J, Lubinski J, Lynch HT, et al. The association between smoking and cancer incidence in BRCA1 and BRCA2 mutation carriers. Int J Cancer. 2018;142(11):2263–72.

180. Breast Cancer FR, (Australasia) KCC for R into FBC, (Canada) OCGN. Smoking and risk of breast cancer in carriers of mutations in BRCA1 or BRCA2 aged less than 50 years. Breast Cancer Res Treat. 2008;109(1):67–75.

181. Ginsburg O, Ghadirian P, Lubinski J, Cybulski C, Lynch H, Neuhausen S, et al. Smoking and the risk of breast cancer in BRCA1 and BRCA2 carriers: an update. Breast Cancer Res Treat. 2009;114(1):127–35.

182. Ghadirian P, Lubinski J, Lynch H, Neuhausen SL, Weber B, Isaacs C, et al. Smoking and the risk of breast cancer among carriers of BRCA mutations. Int J cancer. 2004;110(3):413–6.

183. Bordeleau L, Lipscombe L, Lubinski J, Ghadirian P, Foulkes WD, Neuhausen S, et al. Diabetes and breast cancer among women with BRCA1 and BRCA2 mutations. Cancer. 2011;117(9):1812–8.

184. Stevens C, Steele L, Glover WR, Sanchez-Garcia JF, Simpson SD, O'Rourke D,

et al. A novel CRISPR/Cas9 associated technology for sequence-specific nucleic acid enrichment. PLoS One. 2019;14(4):1–12.

185. Bennett-Baker PE, Mueller JL. CRISPR-mediated isolation of specific megabase segments of genomic DNA. Nucleic Acids Res. 2017;45(19):e165.

186. Tian P, Wang J, Shen X, Rey JF, Yuan Q, Yan Y. Fundamental CRISPR-Cas9 tools and current applications in microbial systems. Synth Syst Biotechnol. 2017;2(3):219–25.

187. Lauer P, Schneider SS, Gnirke A. Construction and validation of yeast artificial chromosome contig maps by RecA-assisted restriction endonuclease cleavage. Proc Natl Acad Sci U S A. 1998;95(19):11318–23.

188. Payne A, Holmes N, Rakyan V, Loose M. Whale watching with BulkVis: A graphical viewer for Oxford Nanopore bulk fast5 files. bioRxiv. 2018 Jan 1;312256.

189. Chial H, Craig J. mtDNA and mitochondrial diseases. Nat Educ. 2008;1:217.

190. Wang M, Lai E. Pulsed field separation of large supercoiled and open-circular DNAs and its application to bacterial artificial chromosome cloning. Electrophoresis. 1995;16(1):1–7.

191. Levene SD, Zimm BH. Separations of open-circular DNA using pulsed-field electrophoresis. Proc Nati Acad Sci USA. 1987;84:4054–7.

192. Khan SR, Kuzminov A. Pulsed-field gel electrophoresis does not break E. coli chromosome undergoing excision repair after UV irradiation. Anal Biochem. 2017;526:66–8.

193. Li C, Chng KR, Boey EJH, Ng AHQ, Wilm A, Nagarajan N. INC-Seq: Accurate single molecule reads using nanopore sequencing. Gigascience. 2016;5(34).

194. Bali L El, Diman A, Bernard A, Roosens NHC, Keersmaecker SCJ De. Comparative Study of Seven Commercial Kits for Human DNA Extraction from Urine Samples Suitable for DNA Biomarker-Based Public Health Studies. J Biomol Tech. 2014;25(4):96–110.

195. Nakayama Y, Yamaguchi H, Einaga N, Esumi M. Pitfalls of DNA Quantification Using DNA-Binding Fluorescent Dyes and Suggested Solutions. PLoS One. 2016;11(3):e0150528.

196. Durney BC, Crihfield CL, Holland LA. Capillary electrophoresis applied to DNA: determining and harnessing sequence and structure to advance bioanalyses (2009-2014). Anal Bioanal Chem. 2015;407(23):6923–38.

197. Schalamun M, Nagar R, Kainer D, Beavan E, Eccles D, Rathjen JP, et al. Harnessing the MinION: An example of how to establish long-read sequencing in a laboratory using challenging plant tissue from Eucalyptus pauciflora. Mol Ecol Resour. 2019;19(1):77–89.

198. Jain M, Koren S, Miga KH, Quick J, Rand AC, Sansani TA, et al. Nanopore sequencing and assembly of a human genome with ultra-long reads. Nat Biotechnol. 2018;36(4):338–45.