Endoscopy International Open

Next-generation sequencing of pancreatic cyst wall specimens obtained using micro-forceps for improving diagnostic accuracy

Stuart Astbury, Aishwarya Baskar, Jane I Grove, Philip Kaye, Aloysious D Aravinthan, Martin James, Christopher Clarke, Guruprasad P Aithal, Suresh Vasan Venkatachalapathy.

Affiliations below.

Abstract:
Background and study aims: Pancreatic cysts are common incidental findings, with an estimated prevalence of 13-15% in imaging done for other reasons. Diagnosis often relies on collection of cyst fluid, but tissue sampling using micro-forceps may allow for a more reliable diagnosis and higher yield of DNA for next-generation sequencing (NGS). The primary aim was to assess the performance of NGS in identifying mucinous cyst. The secondary aims were to assess DNA yield between the cyst fluid and cyst wall tissue, complication rate and performance of conventional investigations.
Patients and methods: 24 patients referred for endoscopic ultrasound were recruited. Biopsies were taken using micro-forceps and the AmpliSeq Cancer Hotspot panel was used for NGS, a PCR assay targeting several hotspots within 50 genes, including GNAS, KRAS and VHL.
Results: The concentration of DNA extracted from 24 cyst wall samples was significantly higher than in the 9/24 available matched cyst fluid samples. The sensitivity, specificity and diagnostic accuracy of NGS for diagnosing mucinous cyst was 93%, 50% and 84%, standard of care -66.6%, 50% and 63.1% and for standard of care with NGS was 100%, 50% and 89.4% respectively. Cyst wall biopsy was able to diagnose 19/24 cysts (4 high risk, 7 intraductal papillary mucinous neoplasm, 4 cysts of mucinous origin and 4 benign).
Conclusions: NGS data correlates well with histology and may aid in diagnosis and risk stratification of pancreatic cysts. Cyst wall biopsy performs well in diagnosing cysts but was inadequate in 5/24 patients.

Corresponding Author:
MRCP Suresh Vasan Venkatachalapathy, Nottingham University Hospitals NHS Trust, Gastroenterology, Nottingham, United Kingdom of Great Britain and Northern Ireland, suresh.venkatachalapathy@nuh.nhs.uk

Affiliations:
Stuart Astbury, University of Nottingham, Nottingham Digestive Diseases Centre, Nottingham, United Kingdom of Great Britain and Northern Ireland
Aishwarya Baskar, University of Nottingham, Nottingham Digestive Diseases Centre, Nottingham, United Kingdom of Great Britain and Northern Ireland
Jane I Grove, University of Nottingham, Nottingham Digestive Diseases Centre, Nottingham, United Kingdom of Great Britain and

Thieme

Northern Ireland

[...]

Suresh Vasan Venkatachalapathy, Nottingham University Hospitals NHS Trust, Gastroenterology, Nottingham, United Kingdom of Great Britain and Northern Ireland

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Thieme

Introduction

Pancreatic cancer is the fourth leading cause of death with a median five-year survival of 4% [1], the overall median survival is 4.6 months from diagnosis [2]. Early diagnosis of pancreatic cancer remains a major challenge and is one of the root causes of the lack of improvement in outcomes of pancreatic cancer worldwide [3]. The Cancer of the Pancreas Screening Study (CAPS3) revealed that 10% of the population screened with a family history of pancreatic cancer had intraductal papillary mucinous neoplasms (IPMN) [4]. In a secondary screening programme involving the European Registry of Hereditary Pancreatitis and Familial Pancreatic Cancer (EUROPAC), cystic lesions were the most common finding of which more than half were IPMNs, although these were independent of genetic predisposition [5]. In this group of patients, earlier diagnosis of high-risk cysts may change their treatment and therefore outcome.

The incidence of pancreatic cysts is 2.5% in the general population but steadily increases to 10% in patients over the age of 70 [6, 7]. Approximately 15% of patients who undergo computed tomography (CT) or magnetic resonance imaging (MRI) for unrelated reasons will have a pancreatic cyst [8]. Common pancreatic cysts include serous cystadenoma (SCA), pancreatic pseudocyst, mucinous cyst neoplasm (MCN) and intraductal mucinous papillary neoplasm (IPMN). SCA and pancreatic pseudocysts are non-mucinous and do not carry malignant potential. However, MCN and IPMN have the potential for malignant transformation (35-50%) and the annual malignant transformation risk is 2%, necessitating regular follow-up [9, 10].

At present, there is no single investigation that will accurately differentiate between high-risk cysts that need intervention, intermediate cysts that require surveillance and low risk cysts, which do not require further surveillance. The current standard of care for patients with pancreatic cysts is to have either CT or MRI with or without endoscopic ultrasound (EUS) and cyst fluid assessment for carcinoembryonic antigen (CEA), amylase and cytology when a sufficient amount of cyst fluid can be aspirated [11].

Next-generation sequencing (NGS) of cyst fluid has shown promise in improving the diagnostic accuracy of differentiating pancreatic cysts [12]. Singhi *et al.* extracted DNA from cyst fluid and used a deep sequencing panel, targeting mutations in KRAS/GNAS and TP52/PIK3CA/PTEN to identify IPMN and advanced neoplasia, respectively. Other studies have demonstrated the utility of NGS for stratifying pancreatic cysts, potentially biasing results towards cysts that more readily shed DNA [13-15]. Additionally, aspiration of adequate fluid through fine needle aspiration (FNA) needle for cytology and DNA extraction is more challenging as the viscosity of cyst fluid increases [16].

1

The micro-forceps are biopsy forceps that can be passed through a 19-gauge fine needle aspiration (FNA) needle, allowing for sampling of pancreatic cyst wall tissue [17]. We conducted a prospective observational study to test the effectiveness of the Moray micro-forceps at increasing the amount of DNA available for NGS relative to that of cyst fluid, and whether this could provide an aid to standard histological and radiological diagnosis. We evaluated the use of cyst wall biopsies obtained with Moray micro-forceps compared to cyst fluid sampling for earlier detection of somatic variants indicating potential for cancer development, and whether this could aid in the diagnosis of progressive cyst types.

Materials and methods

This was a single centre, prospective, observational cohort study. Patients were recruited from Nottingham University Hospitals NHS Trust. Ethical approval was obtained from the Nottingham Health Science Biobank access committee (Reference: 15/NW/0685, approval ACP000282). All adult patients with pancreatic cyst(s) $\geq$ 1.5 cm diameter who were referred for EUS assessment from the regional hepato-pancreato-biliary multi-disciplinary meeting between 15 January 2019 and 15 June 2020 were eligible for inclusion. Patients with a cyst size of less than 1.5cm, or who were clinically diagnosed with pseudocysts were excluded from the study. Patient demographics, cytology, histopathology, imaging findings, CEA analysis, fluid amylase, and EUS findings were documented. The study is outlined as a flow diagram in Figure 1.

Primary and secondary objectives

The primary objective of the study was to identify whether a larger quantity of DNA suitable and sufficient for NGS can be obtained from the cyst wall compared to cyst fluid. The secondary objectives were to assess (1) if NGS alone or in combination with standard of care (CT/MRI + EUS assessment $\pm$ fluid CEA $\pm$ fluid amylase) improves the sensitivity, specificity and diagnostic accuracy of mucinous cyst identification, (2) adequacy of samples for histological assessment, and (3) the complications associated with the cyst wall biopsy using Moray micro-forceps.

EUS procedure and cyst wall biopsy

The procedure was carried out under conscious sedation using a combination of fentanyl and midazolam. EUS was performed by independent endo-sonographers who had received formal training in the fine needle aspiration (FNA) technique. Using EUS, the cyst was identified, evaluated for any high-risk features (>3cm, dilated pancreatic duct >9mm, enhancing wall, mural nodule) and the most accessible portion of the cyst adjacent to the mucosa on EUS was selected. A 19-gauge FNA needle was then introduced into the cyst with 5-10ml dry suction. Fluid was aspirated where possible, allowing the cyst to partially collapse and was sent for amylase, carcinoembryonic antigen (CEA) and the surplus sample was snap-frozen in liquid nitrogen for DNA extraction. Through the FNA needle, Moray micro forceps were introduced into the cyst, opened under ultrasound guidance, withdrawn close to the needle tip, needle tip advanced to the opposite wall, biopsy forceps closed, needle gently withdrawn until there was denting of the cyst wall and then the biopsy forceps were removed in a single motion. This was repeated 3-5 times and intravenous antibiotics were given at the end of the procedure. Samples were placed in formalin for routine histology and surplus tissue was immediately snap-frozen in liquid nitrogen and stored at -80°C.

Histopathological and radiological analysis

The CT/MRI images were reported by an expert independent radiologist blinded to clinical information, previous radiology reports and laboratory results. Similarly, histology was reported by an independent histopathologist blinded to the clinical, radiological and laboratory biochemical information. The slides were prepared at Nottingham University Hospitals NHS Trust, stained for haematoxylin & eosin and sent to the histopathologist with an anonymised numeric code.

DNA extraction and NGS of cyst wall biopsies and cyst fluid

DNA was extracted from cyst wall samples using the Qiagen QIAmp DNA Micro kit and quantified using a Qubit Fluorometer. PCR amplicons were generated using the AmpliSeq

for Illumina Cancer Hotspot Panel v2 (primer sequences in Supplementary table 1) and sequenced on an Illumina MiSeq. Demultiplexed paired end reads were filtered for erroneously short and long sequences, low quality scores (<Phred 30) and adapter sequences using Trimmomatic. Alignment to the human genome reference hg38 was carried out using bwa-mem [18]. The remaining processing was carried out following GATK best practices [19], alignments were filtered using BaseRecalibrator, and somatic variants were called using Mutect2 and FilterMutectCalls [20]. As healthy cyst-adjacent pancreatic tissue from each participant was not available as a control, a panel of normal samples was constructed using publicly available data from the 1000 Genomes project [21] to filter alignment artifacts, and the gnomAD database [22] was used to filter known germline variants. Finally, the filtered list of variants for each sample was annotated using SNPeff [23] to identify coding and non-coding variants. The limit of detection used was a mutant allele frequency (MAF) of 1%, with a minimum sequencing depth of 7000x. Due to the small sample size the list of targets used for analysis was limited to genes with hotspots previously linked to pancreatic malignancies according to the Catalogue of Somatic Mutations in Cancer (COSMIC) database [24].

Confirmatory NGS was carried out on cyst fluid where samples were available. DNA was extracted using the Qiagen ccfDNA extraction kit and primers from the AmpliSeq panel were used to amplify relevant amplicons (these sequences are highlighted in supplementary table 1). Amplified DNA was cleaned using the Qiagen PCR cleanup kit and normalised to 20ng/μl. Samples were sequenced by Azenta/GENEWIZ (Essex, UK). Raw sequence data was processed using an identical workflow to that described above for the cyst wall samples.

Case definition

Four patients in the study had surgical resection. The cyst wall biopsy specimens were compared against the resected surgical specimen. The cyst wall biopsy correlated well with the surgical specimen. Hence, cyst wall tissue histopathology was used as the standard and all parameters were compared against histopathological diagnosis. Standard of care (SOC) is defined as patients having CT/MRI ± EUS± Cytology ± Fluid CEA/amylase to identify if they have high-risk cysts.

Statistical analysis

Categorical variables were presented as numbers with percentages; continuous variables were presented as median and Interquartile range (IQR). Statistical analyses were performed using Prism 9 (GraphPad Inc., San Diego, CA) and SPSS Statistics for Windows, Version 28.0 (IBM, Armonk, NY). A p value <0.05 was considered to indicate statistical significance. Diagnostic accuracy for individual diagnostic test/parameters was calculated

using sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio, and negative likelihood ratio.

Results

Clinical characterisation of patients
Twenty-four patients (54% female) were prospectively recruited to the study (Table 1); the median age was 71 years (IQR 59 – 78). Sixteen patients (67%) were of European descent, 1 (4%) was Asian and for 7 (32%) the ethnicity was not known. After the procedure, patients were followed up for a median of 491 days (IQR 126 – 674).

Histological analysis
The histopathologist was able to give a definite diagnosis based on H&E staining in 15/24 (62.5%) cases, probable diagnosis in 4/24 (16.6%) cases, and in 5/24 (20.8%) of cases the sample was inadequate for diagnosis; although the sample was macroscopically visible in all patients at the time of acquisition. The histological outcomes are shown in Table 1. Cyst wall biopsy specimens were compared against resected surgical specimens in four available cases.

Radiological analysis:
22/24 (91.7%) patients had either a CT or MRI scan. In 2 patients the CT/ MRI was performed in referring hospitals and the images could not be retrieved for independent evaluation. Communication with the pancreatic duct (PD) was established only in 6 patients (27.7%). The main PD was visualised in 14/22 patients (63.6%) and the median diameter was 3mm (IQR 1-2). Mural nodules and septations were identified in 4 (18.1%) and 3 (13.6%) patients, respectively. 15/22 (68.2%) were reported as side branch IPMN (SB-IPMN), 5 as MCN, 1 SCN and 1 adenocarcinoma.

EUS and fluid cytology
The median diameter of the PC was 25mm (IQR 16 – 33). The median PD diameter was 2mm (IQR 2 – 3.5) and communication of the cyst with PD was defined in 23 patients (95.8%). Mural nodules and septations were identified in 4/24 (16.6%) and 7/24 (29.1%), respectively. Obtaining fluid for cytology from the cyst was feasible in 16/24 (66.6%) of patients, but definite diagnosis from fluid cytology was possible in 4/16 (25%) of these patients. The low aspiration rate was due to increased viscosity of the fluid. The remainder had either bland epithelium or the sample was acellular. Aspirated cyst fluid was sent for CEA and amylase analysis. CEA analysis was possible in 7/24 (29.1%) and amylase in 12/24

(50%) of patients. This was because the volume of fluid aspirated was only sufficient for the analysis of amylase.

Follow-up and serious adverse events

One patient developed acute severe pancreatitis requiring prolonged hospitalisation. There were no other complications secondary to the procedure and there were no deaths within 30-days of the procedure. Follow up information was available for all patients. The median follow up period was 491 days (IQR 125 – 674). Four patients had surgical resection, 2 patients had successful radiofrequency ablation of the cyst, 1 had progression of the cyst, 7 had stable cysts, 4 were discharged with benign cysts, 3 were discharged after intervention and 3 died (2 within 12 months and 1 within 2 years). The patients who died had adenocarcinoma (126 days - progression of cancer), IPMN with low-grade dysplasia (74 days-sepsis) and IPMN with high grade dysplasia (514 days - progression to cancer). The patient who had high grade dysplasia progressed to cancer and died because of locally advanced pancreatic cancer.

Evaluation of DNA yield from samples

Of the 16 cyst fluid samples collected, sufficient DNA for NGS was obtained from 9 samples; whereas sufficient DNA was extracted from 24/24 (100%) of cyst wall biopsy samples. The median concentration of DNA from cyst fluid and cyst wall was 1.03 ng/μl (IQR 0.28 – 5.79) and 7.71 ng/μl (IQR 4.74 - 27.5), respectively (Figure 2). The difference between the DNA quantity obtained from the cyst wall tissue and the cyst fluid was significant (p=0.003, Mann-Whitney U test).

Next-generation sequencing of cyst wall DNA

Mutations in KRAS or GNAS were found in all 7 IPMN samples, and in 2/4 samples graded as probable mucinous cyst. The remaining 2 probable mucinous cysts displayed mutations in TP53. Of the 5 samples with an unknown diagnosis based on histology, 2 carried mutations in KRAS or GNAS and another contained a mutation in TP53. Of the 4 benign cysts, one sample (a pseudocyst) contained a KRAS and VHL mutation. Mutations are summarised in Table 2.

Confirmatory NGS in cyst fluid samples

Confirmatory sequencing was carried out in DNA extracted from cyst fluid on hits in the top 4 genes linked to pancreatic cancer or IPMNs (KRAS, GNAS, TP53, VHL). Variants were confirmed in 4/11 of hits found in cyst wall DNA. A further 2 variants found in the cyst wall DNA were detected in the cyst fluid but failed to pass the Mutect2 quality filter (Table 3). No

6

additional variants were detected in cyst fluid DNA that were not detected in the corresponding cyst wall DNA.

Evaluation of the diagnostic accuracy of NGS

The sensitivity, specificity and diagnostic accuracy of standard of care for diagnosing a cyst that needs surveillance or treatment was 66.6%, 50% and 63.1%, respectively. NGS of the cyst wall sample performed better than standard of care with sensitivity of 93.3%, specificity 50%, and diagnostic accuracy of 84.2%. Combining standard of care with NGS improved the above to 100%, 50% and 89.2%, respectively (Table 4). The sensitivity, specificity and diagnostic accuracy of EUS + fluid cytology for identifying mucinous cysts were 40%, 100% and 52%, respectively (Table 2). The sensitivity, specificity and diagnostic accuracy of cytology was 12.5%, 100% and 30%, respectively. It was 100%, 25% and 82% for CT/MRI, respectively.

Discussion

This is the first study to demonstrates that the yield of extracted DNA is significantly higher from cyst wall tissue samples compared to cyst fluid. The micro-forceps yielded sufficient cyst wall tissue for histological analysis in 79% (19/24) of patients. In contrast, DNA adequate to perform NGS was obtained from all cyst wall samples. The diagnostic accuracy for current standard of care is 63% but the addition of molecular analysis of cyst wall samples improved the diagnostic accuracy to 89%.

The cyst wall histology correlated well with the surgical specimens; identified high risk cysts (4/24) that needed surgical intervention and 4/24 had benign cysts who did not need surveillance. However, successful diagnostic histology was obtained only in 79% of patients. Molecular analysis of the cyst wall may be an important adjunct in improving the diagnostic accuracy of cysts in the rest of the patients. Hence, histology in conjunction with NGS may help to accurately diagnose high risk pancreatic cysts that need urgent intervention, mucinous cysts that need surveillance and low risk non-mucinous cysts who can be reassured and discharged. Large multi-centre studies with longer follow-up are needed to confirm or refute the above hypothesis.

The low DNA yield obtained from cyst fluid samples is likely due to the highly variable number of cells shed from the cyst. The source of DNA is due to exfoliation of cyst epithelial cell into the cyst cavity and this may vary from cyst to cyst. It may also be dependent on the amount of fluid aspirated from the cyst. Moreover, this was a bio-bank study hence; only surplus volume of fluid was sent for NGS analysis. This may have contributed to low yield. The performance of NGS on both cyst fluid and cyst wall biopsies are almost similar but the rate limiting factor may be aspiration of sufficient fluid for NGS analysis (after sending it for

7

biochemical analysis). Further studies may need to address the right amount of fluid needed for successful NGS analysis. Obtaining cyst wall tissue with 'through the needle' micro-forceps is a feasible method for increasing the DNA yield from pancreatic cysts. It was possible to do NGS analysis in all cyst wall biopsy samples even though they were small. It requires minimal training and can be carried out by endosonographers who are trained to do FNA of cysts.

Up to 15% of pancreatic cancers are secondary to mucinous pancreatic cysts [25]. As pancreatic cancers are aggressive and associated with poor survival; the current standard of care is to assess the cyst and survey them with either 6 month or annual CT/MRI scans until the patients become unfit for pancreatic resection [11]. This is associated with increased cost burden to the health care system. It may cause significant anxiety to patients not knowing the type of cysts they have. Molecular analysis of the cyst wall samples with histology may aid to alleviate the above problem.

In this study, the median follow up period was relatively short (491 days) for pancreatic cysts. Further, during the follow up period there was no overt progression of the cyst lesions. The follow up scans did not suggest any signs of complications such as pancreatitis/ intra-cystic bleeding. Those who died had pancreatic adenocarcinoma and IPMN with low and high-grade dysplasia. Except for one case of pancreatitis there were no other serious adverse events. This is in line with other studies where the quoted pancreatitis risk following cyst wall biopsy is 3-7%[26]. This is slightly higher than the risk of pancreatitis (2.6%) following FNA of pancreatic cyst[27], although most recent multicentre study using flexible needles reported serious adverse event rate of 1.2% (0.2%-3.5%) [28].

The AmpliSeq panel utilised for this study targets multiple hotspots of 50 genes, these have been narrowed down to genes linked to pancreatic cysts and cancer previously using the COSMIC database [24], and the remainder are included in supplementary data. Mutations in KRAS and GNAS have been proposed as a potential diagnostic test for IPMN but have also been shown to lack the specificity required to detect mucinous cystic neoplasms [29]. Similar to previous studies using cyst fluid [12, 14, 30], we identified KRAS and/or GNAS mutations in 100% of IPMN cases. As expected, all detected KRAS mutations were at codons 12 and 61, the most frequently linked to pancreatic cancer according to the COSMIC database [24]. All but one of the detected GNAS mutations were found at codon 844, which has also been linked to pancreatic cancer and IPMN [31]. A single pseudocyst sample carried a mutation at codon 866, which to our knowledge has not been linked to any cancers. Similar to work in cyst fluid, of the four cysts diagnosed as "probable mucinous cyst" by pathologist, 2 were not found to contain any KRAS or GNAS mutations. TP53 mutations were observed in 3 samples (1 at codon 85 and 2 at codon 217), neither of these sites are associated with pancreatic cancer in COSMIC, but somatic mutations in TP53 at any point have been shown to aid in prediction of outcome of pancreatic adenocarcinoma [32].

8

We endeavoured to test matched cyst fluid samples from patients where samples could be obtained and sufficient DNA could be extracted. No additional mutations were detected in cyst fluid that were not present in cyst wall, however confirmatory work was limited to amplicons where mutations were already detected, and a full screen with the AmpliSeq panel would be required to confirm this definitively. Interestingly, in two cases inspection of the raw variant call data output from Mutect2 revealed mutations were detected in cyst fluid, before being removed by the default quality filter. To our knowledge this is the first study directly comparing cyst wall and cyst fluid DNA samples from the same patient, and suggests that cyst wall biopsy can provide an increased level of detection of somatic mutations compared to cyst fluid. Additionally, variant allele frequencies were significantly different between cyst wall and cyst fluid. There are several potential explanations for this. Cysts with a greater solid component will shed more DNA into the cyst fluid, this could affect variant allele frequencies simply through differences in the amount of DNA sampled. This bias is avoided by sampling the tissue directly [33]. However, extensive genetic heterogeneity within IPMNs has been demonstrated using single-cell sequencing, meaning results could still be biased depending on the part of the cyst that was sampled [34].

The limitations of the study are recruitment from a single centre, a small sample size and a relatively short follow up period. The authors are aware that this may lead to selection bias and an increased intervention effect, therefore the results cannot be generalised to a wider population. The authors tried to minimise these effects by blinding the histopathologist, radiologist and bioinformatician to patient information. In addition, Moray micro-forceps can only be passed through 19-gauge needle, which may not be feasible in a proportion of cases depending on the site and size of the pancreatic cyst. A recent study demonstrated that use of flexible needle increased the success of EUS-FNA to 89% compared to 75% with the standard needle that we used in this study [28]. Another limitation is the size of the sample obtained. The samples obtained were too small for conclusive histological diagnosis in about a fifth of the patients. To our knowledge, one other study has assessed the utility of NGS of cyst wall biopsies in diagnosis, using formalin-fixed paraffin embedded (FFPE) tissue. Although DNA concentrations were not reported, this study demonstrated that almost 20% of cyst wall specimens recovered from FFPE were unsuitable for NGS, whereas our use of fresh frozen tissue resulted in no excluded samples and a greater average sequencing depth [35].

In conclusion, this study demonstrates that it is feasible to obtain cyst wall samples for histological analysis and NGS. The quantity of DNA obtained from tissue was significantly greater than the cyst fluid samples. Larger studies are needed to assess if NGS and cyst wall histology will aid in accurately characterising cysts as high-risk and requiring intervention, those requiring surveillance, and low risk where patients can be reassured and discharged.

Additional information

Authors' contributions

SA obtained funding, designed the study, acquired and interpreted the data, drafted and revised the manuscript and provided supervision. AB acquired and interpreted the data and drafted the manuscript. JIG obtained funding, designed the study, interpreted the data, revised the manuscript and provided supervision. PK acquired and interpreted the data and revised the manuscript. ADA acquired and interpreted the data and revised the manuscript. MWJ interpreted the data and revised the manuscript. CC acquired and interpreted the data and revised the manuscript. GPA obtained funding, provided supervision and revised the manuscript. SVV obtained funding, designed the study, acquired and interpreted the data, drafted and revised the manuscript and provided supervision.

Ethics approval and consent to participate

Ethical approval was obtained from the Nottingham Health Science Biobank access committee (REC reference: 15/NW/0685, approval number: ACP000282). All work was carried out in accordance with the Declaration of Helsinki.

Data availability

Anonymised deep sequence data is available from the NCBI Sequence Read Archive under submission SUB11952288.

10

Competing interests

The authors declare no conflict of interest.

Funding

Supplementary data

Supplementary table 1: Illumina AmpliSeq Cancer Hotspot Panel V2 primer sequences. Primers used for confirmatory sequencing in cyst fluid are in bold. Available as attached Excel file suppl_table_1.xlsx

References

1. Vincent A, Herman J, Schulick R et al. Pancreatic cancer. Lancet 2011; 378: 607-620. doi:10.1016/S0140-6736(10)62307-0

2. Carrato A, Falcone A, Ducreux M et al. A systematic review of the burden of pancreatic cancer in Europe: real-world impact on survival, quality of life and costs. Journal of gastrointestinal cancer 2015; 46: 201-211

3. Walter FM, Mills K, Mendonca SC et al. Symptoms and patient factors associated with diagnostic intervals for pancreatic cancer (SYMPTOM pancreatic study): a prospective cohort study. Lancet Gastroenterol Hepatol 2016; 1: 298-306. doi:10.1016/S2468-1253(16)30079-6

4. Canto MI, Schulick RD, Kamel IR et al. 415g: screening for familial pancreatic neoplasia: a prospective, multicenter blinded study of EUS, CT, and secretin-MRCP (The NCI-Spore Lustgarten Foundation Cancer of the Pancreas CAPS 3 Study). Gastrointestinal Endoscopy 2010; 71: AB119

5. Sheel ARG, Harrison S, Sarantitis I 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: 155-164. doi:10.1038/s41395-018-0395-y

6. Laffan TA, Horton KM, Klein AP et al. Prevalence of unsuspected pancreatic cysts on MDCT. American Journal of Roentgenology 2008; 191: 802-807

7. de Jong K, Nio CY, Mearadji B et al. Disappointing interobserver agreement among radiologists for a classifying diagnosis of pancreatic cysts using magnetic resonance imaging. Pancreas 2012; 41: 278-282

8. Farrell JJ. Prevalence, Diagnosis and Management of Pancreatic Cystic Neoplasms: Current Status and Future Directions. Gut Liver 2015; 9: 571-589. doi:10.5009/gnl15063

9. Grenacher L, Strauss A, Bergmann F et al. Cyst Features and Risk of Malignancy in Intraductal Papillary Mucinous Neoplasms of the Pancreas: Imaging and Pathology. Viszeralmedizin 2015; 31: 31-37. doi:10.1159/000375254

10. Keane MG, Afghani E. A Review of the Diagnosis and Management of Premalignant Pancreatic Cystic Lesions. J Clin Med 2021; 10. doi:10.3390/jcm10061284

11. group TEs. European evidence-based guidelines on pancreatic cystic neoplasms. Gut 2018; 67: 789-804. doi:10.1136/gutjnl-2018-316027

12. Singhi AD, McGrath K, Brand RE et al. Preoperative next-generation sequencing of pancreatic cyst fluid is highly accurate in cyst classification and detection of advanced neoplasia. Gut 2017. doi:10.1136/gutjnl-2016-313586. doi:10.1136/gutjnl-2016-313586

13. Jones M, Zheng Z, Wang J et al. Impact of next-generation sequencing on the clinical diagnosis of pancreatic cysts. Gastrointest Endosc 2016; 83: 140-148. doi:10.1016/j.gie.2015.06.047

14. Rosenbaum MW, Jones M, Dudley JC et al. Next-generation sequencing adds value to the preoperative diagnosis of pancreatic cysts. Cancer Cytopathol 2017; 125: 41-47. doi:10.1002/cncy.21775

15. Paniccia A, Polanco PM, Boone BA et al. Prospective, Multi-Institutional, Real-Time Next-Generation Sequencing of Pancreatic Cyst Fluid Reveals Diverse Genomic Alterations That Improve the Clinical Management of Pancreatic Cysts. Gastroenterology 2023; 164. doi:10.1053/j.gastro.2022.09.028

16. Khalid A, McGrath KM, Zahid M et al. The role of pancreatic cyst fluid molecular analysis in predicting cyst pathology. Clin Gastroenterol Hepatol 2005; 3: 967-973. doi:10.1016/s1542-3565(05)00409-x

17. Zhang ML, Arpin RN, Brugge WR et al. Moray micro forceps biopsy improves the diagnosis of specific pancreatic cysts. Cancer Cytopathol 2018; 126: 414-420. doi:10.1002/cncy.21988

18. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009; 25: 1754-1760. doi:10.1093/bioinformatics/btp324

19. DePristo MA, Banks E, Poplin R et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 2011; 43: 491-498. doi:10.1038/ng.806

20. Benjamin DS, T.; Cibulskis, K.; Getz, G.; Stewart, C.; Lichtenstein, L. Calling Somatic SNVs and Indels with Mutect2. bioRxiv 2019. doi:10.1101/861054 doi:10.1101/861054

21. Consortium TGP. A global reference for human genetic variation. Nature 2015; 526: 68-74. doi:10.1038/nature15393

22. Karczewski KJ, Francioli LC, Tiao G et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 2020; 581: 434-443. doi:10.1038/s41586-020-2308-7

23. Cingolani P, Platts A, Wang le L et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) 2012; 6: 80-92. doi:10.4161/fly.19695

24. Tate JG, Bamford S, Jubb HC et al. COSMIC: the Catalogue Of Somatic Mutations In Cancer. Nucleic Acids Res 2019; 47: D941-D947. doi:10.1093/nar/gky1015

25. Singhi AD, Koay EJ, Chari ST et al. Early detection of pancreatic cancer: opportunities and challenges. Gastroenterology 2019; 156: 2024-2040

26. Balaban VD, Cazacu IM, Pinte L et al. EUS-through-the-needle microbiopsy forceps in pancreatic cystic lesions: A systematic review. Endoscopic ultrasound 2021; 10: 19

27. Siddiqui AA, Shahid H, Shah A et al. High risk of acute pancreatitis after endoscopic ultrasound-guided fine needle aspiration of side branch intraductal papillary mucinous neoplasms. Endoscopic ultrasound 2015; 4: 109

28. Al-Haddad M, Wallace MB, Brugge W et al. Fine-needle aspiration of pancreatic cystic lesions: a randomized study with long-term follow-up comparing standard and flexible needles. Endoscopy 2021; 53: 1132-1140. doi:10.1055/a-1311-9927

29. Singhi AD, Nikiforova MN, Fasanella KE et al. Preoperative GNAS and KRAS testing in the diagnosis of pancreatic mucinous cysts. Clin Cancer Res 2014; 20: 4381-4389. doi:10.1158/1078-0432.CCR-14-0513

30. Faias S, Duarte M, Albuquerque C et al. Clinical Impact of KRAS and GNAS Analysis Added to CEA and Cytology in Pancreatic Cystic Fluid Obtained by EUS-FNA. Dig Dis Sci 2018; 63: 2351-2361. doi:10.1007/s10620-018-5128-y

31. Tan MC, Basturk O, Brannon AR et al. GNAS and KRAS Mutations Define Separate Progression Pathways in Intraductal Papillary Mucinous Neoplasm-Associated Carcinoma. J Am Coll Surg 2015; 220: 845-854 e841. doi:10.1016/j.jamcollsurg.2014.11.029

32. Shoucair S, Habib JR, Pu N et al. Comprehensive Analysis of Somatic Mutations in Driver Genes of Resected Pancreatic Ductal Adenocarcinoma Reveals KRAS G12D and Mutant TP53 Combination as an Independent Predictor of Clinical Outcome. Ann Surg Oncol 2022; 29: 2720-2731. doi:10.1245/s10434-021-11081-z

33. Khalid A, Zahid M, Finkelstein SD et al. Pancreatic cyst fluid DNA analysis in evaluating pancreatic cysts: a report of the PANDA study. Gastrointest Endosc 2009; 69: 1095-1102. doi:10.1016/j.gie.2008.07.033

34. Kuboki Y, Fischer CG, Beleva Guthrie V et al. Single-cell sequencing defines genetic heterogeneity in pancreatic cancer precursor lesions. J Pathol 2019; 247: 347-356. doi:10.1002/path.5194

35. Rift CV, Melchior LC, Kovacevic B et al. Targeted next-generation sequencing of EUS-guided through-the-needle-biopsy sampling from pancreatic cystic lesions. Gastrointest Endosc 2022. doi:10.1016/j.gie.2022.08.008. doi:10.1016/j.gie.2022.08.008

Gene	Forward primer (5'-3')	Reverse primer (5'-3')
MPL_1	AGAGCGAACCAAGAATGCCTGTTTA	CTTGGTGACCGCTCTGCATCTAGTG
NRAS_3	AACAAGTGTGATTTGCCAACAAGGA	TTGCACAAATGCTGAAAGCTGTACC
NRAS_2	ACACCCCCAGGATTCTTACAGAAAA	GCCTTCGCCTGTCCTCATGTATTGG
NRAS_1	GGTTCTTGCTGGTGTGAAATGACTG	CCTCACCTCTATGGTGGGATCATAT
ALK_2	CCCTGGAAGAGTGGCCAAGATTGGA	ACAGGGTACCAGGAGATGATGTAAG
ALK_1	TTTGCCCAGACTCAGCTCAGTTAAT	GGGTCTCTCGGAGGAAGGACTTGAG
IDH1_1	GAAATCACCAAATGGCACCATACGA	CCAACATGACTTACTTGATCCCCAT
ERBB4_8	TGGGAACTGATGACCTTTGGAGGAA	ATTTGACCATGACCATGTAAACGTC
ERBB4_7	CAGGGTCCTGACAACTGTACAAAGT	TTGGCCAGCAAGAATGCTTACCCTT
ERBB4_6	CTCAATCCCCTAACTCTGAGTCTTG	CCATGAATACCAGTGACTAGAAAGA
ERBB4_5	TGTTTTGAGCTTGTTTGCTGAATGT	TGGGCAAATGTCAGTGCAAGGTTTA
ERBB4_4	TGCAGGCCTGCATGAATTTCAATGA	GTTACTTACGTGGACATTTCTTGAC
ERBB4_3	GTGACGGCAGATGCTACGGACCTTA	CTGAATCAAATAGGGAAGGAAAGGA
ERBB4_2	GACACCATTCATTGGCAAGATATTG	CATCGCCACATAGGGTAGAACATTT
ERBB4_1	AACTTTGGACTTCAAGAACTTGGAT	CTTAGAGTGTTCCTCAATGTAACAA
VHL_1	**GCCCGTACCTCGGTAGCTGTGGATG**	**GCCCTCCCAGGTCATCTTCTGCAAT**
VHL_2	AACGTCAGTACCTGGCAGTGTGATA	GTGTGGCTCTTTAACAACCTTTGCT
VHL_3	**CTGACGATGTCCAGTCTCCTGTAAT**	**TGTTGGCAAAGCCTCTTGTTCGTTC**
MLH1_1	CCTGCCACTAGAAATATCTGTCTTA	GTCTGACCTCGTCTTCTACTTCTGG
CTNNB1_1	CATCCTCTTCCTCAGGATTGCCTTT	TCGTATTTATAGCTGATTTGATGGA
PIK3CA_1	AAGGGTTGAAAAAGCCGAAGGTCAC	GCATGAACTATTTAAAGAAGCAAGA
PIK3CA_2	TACGAAGGTATTGGTTTAGACAGAA	TTTAAAAGTAATTGAACCAGTAGGC
PIK3CA_3	ATCAGCATTTGACTTTACCTTATCA	TCCAGACGCATTTCCACAGCTACAC
PIK3CA_4	ATCAGCGGTATAATCAGGAGTTTTT	AGGTGGAATGAATGGCTGAATTATG
PIK3CA_5	AACAAGTTTATATTTCCCCATGCCA	ATAGAGATGATTGTTGAATTTTCCT
PIK3CA_6	TAAGAGAGAAGGTTTGACTGCCATA	GGCTTTGAATCTTTGGCCAGTACCT
PIK3CA_7	GCACTTACCTGTGACTCCATAGAAA	GTAACAGACTAGCTAGAGACAATGA
PIK3CA_8	TTTGTGTTTCATCCTTCTTCTCCTG	GATTCTTTTAGATCTGAGATGCACA
PIK3CA_9	GAAAACCATTACTTGTCCATCGTCT	ATGATGCAGCCATTGACCTGTTTAC
PIK3CA_10	GAAGATCCAATCCATTTTTGTTGTC	AATGCCAGAACTACAATCTTTTGAT
PIK3CA_11	ATGCTGTTCATGGATTGTGCAATTC	GATCTTCCACACAATTAAACAGCAT
FGFR3_1	GGATGTGGGGCTGTGCGTCACTGTA	AGGGGGTGGCCCCTGAGCGTCATCT
FGFR3_2	CGGTGGGGGAGCCCAGGCCTTTCTT	CCGAGGAGGAGCTGGTGGAGGCTGA
FGFR3_3	CGCCAGGCGTCCTACTGGCATGACC	TGCTGGTGACCGAGGACAACGTGAT
FGFR3_4	GGGATGCCACTCACAGGTCGTGTGT	TCCTGCTCTGGGAGATCTTCACGCT
FGFR3_5	GGCTGCTAGGGACCCCTCACATTGT	GTCGGCGCCTTTCGAGCAGTACTCC
PDGFRA_1	CATCTCTTGGAAACTCCCATCTTGA	GTGCACTGGGACTTTGGTAATTCAC
PDGFRA_2	ACCACATGTGTCCAGTGAAAATCCT	AGTGAAAAACAAGCTCTCATGTCTG
PDGFRA_3	GTGTGCTTTCATCAGCAGGGTTCAA	CTTTTGTCCCCATAGGCCCCATTTA
PDGFRA_4	TTAAAGTGAAGGAGGATGAGCCTGA	CCATGCAGTGTGTCCACCGTGATCT
KIT_1	TCCATTTGACAAAGCCCGGATCAGT	CGCCAAGGAAGAAGATCATACTCAA

KIT_2	GATATGGTAGACAGAGCCTAAACAT	ATGGCACGGTTGAATGTAAGGCTTA
KIT_3	TGTCTCAGTCATTAGAGCACTCTGG	ATCCCCACACCCTGTTCACTCCTTT
KIT_4	TTTCATACTGACCAAAACTCAGCCT	AGGTGATCTATTTTTCCCTTTCTCC
KIT_5	CTTGGACACGGCTTTACCTCCAATG	TTTTTGCTAAAATGCATGTTTCCAA
KIT_6	TTGAACAAATAAATGAATCACGTTT	CTTTCTAACCTTTTCTTATGTGCTT
KIT_7	GGACAACATAAGAAACTCCAGGTTT	GATGAGGAGGTAGAGCATGACCCAT
KIT_8	TGTCAAGCAGAGAATGGGTACTCAC	AGCCAGAAATATCCTCCTTACTCAT
KIT_9	CCTAAAGAGAACAGCTCCCAAAGAA	GTGCTTCTATTACAGGCTCGACTAC
KDR_9	ACCACCGTGTACTCCAGTGAGGAAG	CCTCTCATGTGATGTCCAGGAGTTG
KDR_8	TGAGGGTAAAAAGCAAAAGAATTGT	ACGGTGGTGTCTGTGTCATCGGAGT
KDR_7	TGAGCATGGAAGAGGATTCTGGACT	GGATGTTAGGCCATATACAGTACCT
KDR_6	CAGCATTCAGGAAGAAAGAGGCATT	AAGAGATTTCCCAAATGTTCCACCA
KDR_5	TGGAGCAATCCCTGTGGATCTGAAA	GAAGCACTAGCCAGTACCTTCCTCT
KDR_4	GGGACCCCAATTATTGAAGGAAATG	GAAGGTTGACCACATTGAGATGGTG
KDR_3	GGCTGCGTTGGAAGTTATTTCTAAG	CACTTCTCCATTCTTCACAAGGGTA
KDR_2	GCTTTGGAAGTTCAGTCAACTCTTT	CCGGGTTACACCATCTATAGTTAAG
KDR_1	TATCTGTTGGAGAAAAGCTTGTCTT	GCCCCTATCTCTCAAGCAAACTTCA
FBXW7_5	TCTTGATACATCAATCCGTGTTTGG	GTAGAATCTGCATTCCCAGAGACAA
FBXW7_4	TCCCTCTGCAGAGTTGTTAGCGGTT	CACTAACAACCCTCCTGCCATCATA
FBXW7_3	TCATCACAAATGAGAGACAACATCA	GTCCCAACCATGACAAGATTTTCCC
FBXW7_2	CATCTTTCTTATAGGTGCTGAAAGG	CCTGTGACTGCTGACCAAACTTTTA
FBXW7_1	TCATTGATAGTTGTGAACCAACACA	ACAATGTTTAAAGGTGGTAGCTGTT
APC_1	GTATGAATGGCTGACACTTCTTCCA	GGAGAGAGAACGCGGAATTGGTCTA
APC_2	CATCTTCTTGACACAAAGACTGGCT	GCACTGATGATAAACACCTCAAGTT
APC_3	GGATTTGGTTCTAGGGTGCTGTGAC	ATTATCATCTTTGTCATCAGCTGAA
APC_4	GTGAACTGACAGAAGTACATCTGCT	TCCAGCAGACTGCAGGGTTCTAGTT
APC_5	GACCCTCTGAACTGCAGCATTTACT	ATAAGCCCCAGTGATCTTCCAGATA
APC_6	GTCATTTTCCTGAACTGGAGGCATT	AGTTCAGAGGGTCCAGGTTCTTCCA
APC_7	TTACGTGATGACTTTGTTGGCATGG	TGAAACAGAATCAGAGCAGCCTAAA
CSF1R_2	AGAGCTCTAGTGAGCACCTGACCTG	CCCCATCCATGGAGGAGTTGAAGTT
CSF1R_1	TCTTCCCACTAATGCCAGATGCTTG	CCTCCACCATGACTTTGAGGTTGAG
NPM1_1	CAGCCAGATATCAACTGTTACAGAA	ATGTCTATGAAGTGTTGTGGTTCCT
EGFR_1	GTCCGGTTTTATTTGCATCATAGTT	ATGTCCTCATTGCCCTCAACACAGT
EGFR_2	CCCAAAGACTCTCCAAGATGGGATA	ACCCACCACGTACCAGATGGATGT
EGFR_3	ACTTCCAGACCAGGGTGTTGTTTTC	GACATGCATGAACATTTTTCTCCAC
EGFR_4	CTGTGCCAGGGACCTTACCTTATAC	CAGCTTGTGGAGCCTCTTACACCCA
EGFR_5	TGGGCCTGAGGTTCAGAGCCATGGA	TAACGTCTTCCTTCTCTCTCTGTCA
EGFR_6	GTTCCCGGACATAGTCCAGGAGGCA	GGCCACCATGCGAAGCCACACTGAC
EGFR_7	ATATCCCCATGGCAAACTCTTGCTA	TGGACTATGTCCGGGAACACAAAGA
EGFR_8	CATGTGTTAAACAATACAGCTAGTG	ACCGCAGCATGTCAAGATCACAGAT
MET_1	AAAGAAGTTGATGAACCGGTCCTTT	TGCTGACATACAGTCGGAGGTTCAC
MET_2	AGACATCTCACATTGTTTTTGTTGA	ACAAATAGGAGCCAGCCTGAATGAT

MET_3	TTTTCAAAAGGCTTAAACACAGGAT	TGTCAACATCGCTCTAATTCAGAGA
MET_4	AGTCTACAGATTCATTTGAAACCAT	GCCCATGATAGCCGTCTTTAACAAG
MET_5	AAACCACAAAAGTATACTCCATGGT	AATGTTACGCAGTGCTAACCAAGTT
MET_6	TGACTTGGTGGTAAACTTTTGAGTT	GTTGCTGATTTTGGTCTTGCCAGAG
SMO_1	GCCGCGATGTAGCTGTGCATGTCCT	GCCAGAATGAGGTGCAGAACATCAA
SMO_2	GCCAGGCATAGGTGAGGACCACAAA	ACCAGGTAGAGGGAGTACAGAGTGA
SMO_3	CCTGGTCTTCACTCACCTCGGATGA	GGACTCTGTGAGTGGGATTTGTTTT
SMO_4	TGCACCAGGTACGCCTCCAGATGAG	CCATCCCTGACTGTGAGATCAAGAA
SMO_5	TCCTCCAGAAGCTTGAACTCTCATA	AGCGGGCTTGGCCTTTGACCTCAAT
BRAF_2	TGCTCTGATAGGAAAATGAGATCTA	CCACAAAATGGATCCAGACAACTGT
BRAF_1	TCTTTTTCTGTTTGGCTTGACTTGA	CATACTTACCATGCCACTTTCCCTT
EZH2_1	CTAGCATCTATTGCTGGCACCATCT	CAATGCCACCTGAATACAGGTTATC
FGFR1_2	CCACCCCTCTTTAGCCATGGCAAGG	GGTCACTGTACACCTTACACATGAA
FGFR1_1	CTAGGGGTCCCTAGGAGGAACCTCA	GACCCAAAGGGCAGTAAGATAGGAA
JAK2_1	ACACTGACACCTAGCTGTGATCCTG	TTGAAGCAGCAAGTATGATGAGCAA
CDKN2A_2	CTTCCTGGACACGCTGGTGGTGCTG	CGGCATCTATGCGGGCATGGTTACT
CDKN2A_1	CCCACCCTGGCTCTGACCATTCTGT	GCAGCACCACCAGCGTGTCCAGGAA
GNAQ_1	TTGTTAACCTTGCAGAATGGTCGAT	ATCCATTGCCTGTCTAAAGAACACT
ABL1_1	ACCGTCAGGCTGTATTTCTTCCACA	ACTGTCTATGGTGTGTCCCCCAACT
ABL1_2	CAGCTTCTTTCAAGAACTCTTCCAC	TATGCGCTGAAGCTCCATTTTGCAT
ABL1_3	GCCTGGCCAGGCCCCTACCTGTGGA	ACCCGGGAGCCCCCGTTCTATATCA
ABL1_4	CGGACTTGATGGAGAACTTGTTGTA	GGGGAGAACCACTTGGTGAAGGTAG
NOTCH1_3	CCACCTGGGCCGGAGCTTCCTGAGT	GGCGTCAGGAACTGGGCTGCGGTCA
NOTCH1_2	GCTGCCTCACCATGTCCTGACTGTG	GCCTGCACACACTGCCGGTTGTCAA
NOTCH1_1	CGGGCTGGACTGTGCGGAGCATGTA	TGCGTCACGCTTGAAGACCACGTTG
RET_1	CCTCCCTTGTTGGGACCTCAGATGT	CGGGGGATTAAAGCTGGCTATGGCA
RET_2	ACTTGTGGTAGCAGTGGATGCAGAA	AGGCAGAGCATACGCAGCCTGTACC
RET_3	GCAGGCCCCATACAATTTGATGACA	AAGGGGCTTCCAGGAGCGATCGTTT
RET_4	GCCTGCCCCATGGTGCACCTGGGAT	CTCACAGCTCGTTCATCGGGACTTG
RET_5	CTTTGCGTGGTGTAGATATGATCAA	CTCTGGTTACTGAAAGCTCAGGGAT
PTEN_1	GCAGCCGCAGAAATGGATACAGGTC	TGCCATCTCTCTCCTCCTTTTTCTT
PTEN_2	ACCTCACTCTAACAAGCAGATAACT	TAATGGTGGCTTTTTGTTTGTTTGT
PTEN_3	GCCCCGATGTAATAAATATGCACAT	AGACCATAACCCACCACAGCTAGAA
PTEN_4	GCCACTGGTCTATAATCCAGATGAT	AATGGCTACGACCCAGTTACCATAG
PTEN_5	TCTTCACCTTTAGCTGGCAGACCAC	GATCAAGATTGCAGATACAGAATCC
PTEN_6	ATTTCTCCCAATGAAAGTAAAGTAC	AAGATATATTCCTCCAATTCAGGAC
PTEN_7	CGCTCTATACTGCAAATGCTATCGA	GGTAAATACATTCTTCATACCAGGA
PTEN_8	TCACATACATACAAGTCAACAACCC	GCAGTATAGAGCGTGCAGATAATGA
FGFR2_4	GATGAAGATGATTGGGAAACACAAG	ATCTGGAAGCCCAGCCATTTCTAAA
FGFR2_3	TTCCTCCTCCTGTGATCTGCAATCT	TGGCTTCTTGGTCGTGTTCTTCATT
FGFR2_2	TCCAGTGGATCAAGCACGTGGAAAA	ATCATCCTCTCTCAACTCCAACAGG
FGFR2_1	TATTGGTCTCTCATTCTCCCATCCC	ATCACTGTAAACCTTGCAGACAAAC

HRAS_2	TCCTCCTGCAGGATTCCTACCGGAA	AAGACTTGGTGTTGTTGATGGCAAA
HRAS_1	CCTGAGGAGCGATGACGGAATATAA	GCGGCGCCAGGCTCACCTCTATAGT
ATM_1	GAAATTTCTAAATGTGACATGACCT	CTGCTAGTGAATGAGATAAGTCATA
ATM_2	AGATGAAAGGATTCCACTGAAAGTT	CCTTCAGAAGTCACAGAATGATTTT
ATM_3	ATACCATTCTGGCACGCTTTGGAAA	TTCTTTACATGGCTTTTGGTCTTCT
ATM_4	TCCAGGTTCGTTTGCATCACTAACA	ATCAAAAAGCCATTTGACCGTGGAG
ATM_5	CATACCTGTTTTCCCAATAAGTTTT	AGACTGCTTTCCAAAGATTCTTGTA
ATM_6	TCAGCATTATGAAGGTCCACTGAAG	TATATGTAGAGGCTGTTGGAAGCTG
ATM_7	CTGTAGATAGGCCAGCATTGGATCT	TGGTGTACTTGATAGGCATTTGAAT
ATM_8	ATTTCACATTTTGTGCCTCCACTGT	AGAAACCCTTTTGAAGGCCTGGAT
ATM_9	GCTGTAAAGTGAGCAGCACAAGACT	TTTCTTTAGACCTTCTTCAGGAACA
ATM_10	AAGCGTTTACGATCCTCTTTCAGTG	TCTGTTAAAGTTCATGGCTTTTGTG
ATM_11	TAATTGCAAATTACCTTAATTTCCA	CAGAAGGCATAAATATTCCAGCAGA
ATM_12	CTCACCTTAACAAGCTGTCTCCTCT	AGGAGAATATGGAAATCTGGTGACT
ATM_13	TTTCCGTGTTTCTCTGCAGTAATGT	TATTCTCAGATGACTCTGTGTTTTT
ATM_14	ATCTTGGTAGGCAAACAACATTCCA	TTCTTTTTTCTCCAGTTGGTTACAT
ATM_15	ATACACAGTCTACCTGGTAAGAAAA	TGTTTGACTCTAGATGCTGTGAGAA
ATM_16	AAATTCACTTGTCCACCAACACTGA	CACCTCACTGAAACCTTTGTGTTTT
ATM_17	TGTTGGCAGGTTAAAAATAAAGGCT	ATTTGCTCATACAGCAGGCCATAGA
KRAS_3	CTCTGAAGATGTACCTATGGTCCTA	GTATTTATTTCAGTGTTACTTACCT
KRAS_2	**AAGGTGCACTGTAATAATCCAGACT**	**AGTCCTCATGTACTGGTCCCTCATT**
KRAS_1	**GCCTGCTGAAAATGACTGAATATAA**	**AAGAATGGTCCTGCACCAGTAATAT**
PTPN11_1	TCTTTTAATTGCCCGTGATGTTCCA	TGCCTCCCTTTCCAATGGACTATTT
PTPN11_2	CTGGTACCTGCTCTTCTTCAATCCT	TGATGTTTCCTTCGTAGGTGTTGAC
HNF1A_1	CTCGCTCCTCCTTGCTAGGGTTCTT	GGCTGATTGAAGAGCCCACAGGTGA
HNF1A_2	GGGGGGGCCCGCTGTACGTGTCCAT	GGGGTGTCCCCATCACAGGCACAGG
FLT3_4	CACCCACGGGAAAGTGGTGAAGATA	CACAACACAAAATAGCCGTATAAAA
FLT3_3	GAGAGGCACTCATGTCAGAACTCAA	AACATCCTCTTTGTCATCAAGCTAC
FLT3_2	ACTCATCATTTCATCTCTGAAGCAA	CACATTCCATTCTTACCAAACTCTA
FLT3_1	GAGCTTATTTCACACGTTCTTTTCT	TACTTGGGAGACTTGTCTGAACACT
RB1_1	ACCTTTCCAATTTGCTGAAGAGTGC	TATTCTTTCCTTTGTAGTGTCCATA
RB1_2	CCAAGCAGAGAATGAGGGAGGAGTA	CATTGGTGCTAAAAGTTTCTTGGAT
RB1_3	TGTGCAATACCTGTCTATAGAATCA	GAGAGATGTAATGACATGTAAAGGA
RB1_4	CTGGAGTGTGTGGAGGAATTACATT	GACAACAGAAGCATTATACTGCTTT
RB1_5	GATCTTGATGCCTTGACCTCCTGAT	GCGATACAAACTTGGAGTTCGCTTG
RB1_6	TTGTTAGCCATATGCACATGAATGA	AGGCAACTTGACAAGAGAAATGATA
RB1_7	ACAAGCAGATTCAAGGTGATCAGTT	AAAATTATGCTTACTAATGTGGTTT
RB1_8	TGCCTGTCTCTCATGAGTTCATACT	TAATTTTTCTTATTCCCACAGTGTA
RB1_9	GAGGAAGATCCTTGTATGCTGTTAC	CATGTAATAAAATTCTGACTACTTT
RB1_10	TTTGACCTACCCTGGTGGAAGCATA	CTTCCTCAGACATTCAAACGTGTTT
AKT1_2	ACCATGAACGAGTTTGAGTACCTGA	CCACCTTGGCCACGATGACTTCCTT
AKT1_1	GGCGAGGGTCTGACGGGTAGAGTGT	TTACGCGCCACAGAGAAGTTGTTGA

IDH2_1	AGTTCAAGCTGAAGAAGATGTGGAA	CCTCTCCACCCTGGCCTACCTGGTC
CDH1_1	AGTCCCAGGCGTAGACCAAGAAATG	TTTGAAGATTGCACCGGTCGACAAA
CDH1_2	ATAGAATTACCGTGGTGGGATTGAA	TTCCTGGTCCTGACTTGGTTGTGTC
CDH1_3	TCAATATGGTGTATACAGCCTCCCA	GTCTTGGTACTTTGTAAATGACACA
TP53_8	ACTTCTCCCCCTCCTCTGTTGCTGC	GGGCCAGGAAGGGGCTGAGGTCACT
TP53_7	CTCTTTTCCTATCCTGAGTAGTGGT	TCCACCGCTTCTTGTCCTGCTTGCT
TP53_6	CCTCATCTTGGGCCTGTGTTATCTC	CAAGTGGCTCCTGACCTGGAGTCTT
TP53_5	**CCCAGGCCTCTGATTCCTCACTGAT**	**GACCCCAGTTGCAAACCAGACCTCA**
TP53_4	GCCCTGTGCAGCTGTGGGTTGATTC	CTGGGCAACCAGCCCTGTCGTCTCT
TP53_3	GCCGTCTTCCAGTTGCTTTATCTGT	CCACAGCTGCACAGGGCAGGTCTTG
TP53_2	TCCAGATGAAGCTCCCAGAATGCCA	ACTTGGCTGTCCCAGAATGCAAGAA
TP53_1	GGGTTGGAAGTGTCTCATGCTGGAT	ATCCACTCACAGTTTCCATAGGTCT
ERBB2_1	GAGTCATATCTCCCCAAACCCCAAT	GAATGTGAAAATTCCAGTGGCCATC
ERBB2_2	GCAGCCATAGGGCATAAGCTGTGTC	TTTGGGGGTGTGTGGTCTCCCATAC
ERBB2_3	CTCCTTGGTCCTTCACCTAACCTTG	AGGGGATGAGCTACCTGGAGGATGT
SMAD4_1	CTACTTACCAATTCCAGGTGATACA	TTCCTCATGTGATCTATGCCCGTCT
SMAD4_2	TTACCTACCATTACTCTGCAGTGTT	GCTACTTCTGAATTGAAATGGTTCA
SMAD4_3	TGCTGGTAGCATTAGACTCAGATGG	TGGTGAAGGATGAATATGTGCATGA
SMAD4_4	ATTAAGGCCCACATGGGTTAATTTG	AGTGAAGGACTGTTGCAGATAGCAT
SMAD4_5	TGAGAAGTGACCCCATAATTCCATT	TTTCTTTAGGGCCTGTTCACAATGA
SMAD4_6	TGTCCTGTGGACATTGGAGAGTTGA	TAGCTCCTGAGTATTGGTGTTCCAT
SMAD4_7	CTTGGGTAGATCTTATGAACAGCAT	TTTTTCTTCCTAAGGTTGCACATAG
SMAD4_8	ATAGCTGGAGCTATTCCACCTACTG	AAAGGTCTTTGATTTGCGTCAGTGT
SMAD4_9	GAAGTACTTCGTCTAGGAGCTGGAG	GCTGCTGCTGGAATTGGTGTTGATG
STK11_1	ACAGCGTCTCCGAGTCCAGCACCTC	AGGGCGAGCTGATGTCGGTGGGTAT
STK11_2	GGGTGCCACCGGTGGTGAGCAGCAG	TTGGGAGGCTCCCAGGCAGCTGCAA
STK11_3	AGCCCTGGCTGGTCCGGCAGGTGTC	GCTCACCACCGGTGGCACCCTCAAA
STK11_4	GAGGGATGAGGCTCCCACCTTTCAG	CTACAACATCACCACGGGTCTGTAC
STK11_5	GGCACCGTGAAGTCCTGAGTGTAGA	CCGGAAGAAACATCCTCCGGCTGAA
GNA11_1	ACATGATGGATGTCACGTTCTCAAA	TCCTGGGATTGCAGATTGGGCCTTG
JAK3_3	ACACTTAGCTTGGAAGCTGACAAGT	ACTGGATGTCAGTCTGCCCTTCTGT
JAK3_2	GGTCTGTGAGCACAAAATTTGGGAT	GGAGCAGCACGAGATGCCGGTACGA
JAK3_1	TACGCAAGGATTTGGCCAGTGCTAT	CCCACCTGATTGCATGCCAGTCCTC
SRC_1	GCCCGCCTGTGCCTAGAGGTTCTCC	CCCGAGTCCCTGCACGACCTCATGT
GNAS_1	**ATGTCAAGAAACCATGATCTCTGTT**	**TGGTGAGATCCATTGACCTCAATTT**
GNAS_2	**CTGATCCCTAACAACACAGAAGCAA**	**GTCCCTCTGGAATAACCAGCTGTCC**
SMARCB1_1	AGACGCACCCTTAGTGTTAGGTTTT	TGCTTTACTCATAGGTGGGAAACTA
SMARCB1_2	GAACTGAAACGTGCTGGAGAACTAA	CAACAGCTCCCACCACTTAGATGCC
SMARCB1_3	CGACTGCCTTGTACCATTCATGTTC	GCTGACTGTTGCTTCCATTTCACTT
SMARCB1_4	GGCCCAATCTTCTGAGATGCTCCGT	CCCTACACTTGGCTGCCCTGTAGAG

Table 1: Cohort characteristics and histological diagnosis based on cyst wall biopsy.

Patient demographic data	Median (IQR) or n (%)
Sex	13 female (54%)
Age	71 (59-78)
Ethnicity	
European	16 (67%)
Non-European (Pakistan)	1 (4%)
Unknown	7 (32%)
Cyst aetiology	
Malignant	
Adenocarcinoma	1 (4.1%)
NET	3 (12.5%)
Mucinous cyst	
IPMN	7 (29.2%)
Probable cyst of mucinous origin	4 (16.7%)
Non-mucinous Cyst	
Benign epithelial cyst	1 (4.1%)
Lymphoepithelial cyst	1 (4.1%)
SCA	1 (4.1%)
Pseudocyst	1 (4.1%)
Inadequate tissue for diagnosis	5 (20.8%)

Table 2: Somatic protein coding mutations with over 1% prevalence in pancreatic cyst samples. PSC=pseudocyst, EC=epithelial cyst, IPMN=intraductal papillary mucinous neoplasm, SCA=serous cystadenoma, NET=neuroendocrine tumour, LEC=lymphoepithelial cyst, ADC=adenocarcinoma. * indicates a stop codon. All mutant allele frequencies are expressed as percentage of sequence reads containing the mutation compared to the hg38 reference genome wild type sequence. The NGS was compared against Histology.

Patient	Diagnosis	KRAS	GNAS	TP53	VHL	NRAS	CTNNB1	PTEN
1	ADC	p.G12R (10%)	-	-	-	-	-	-
2	NET	-	-	-	-	-	-	-
3	NET	-	-	-	-	-	-	-
4	NET	-	-	-	-	-	-	-
5	IPMN (high grade)	-	p.A844H (31%)	-	-	-	-	-
6	IPMN (low grade)	p.Q61H (40%)	p.R844C (55%)	-	-	-	-	-
7	IPMN (low grade)	p.G12D (8%)	p.R844H (8%)	-	-	-	-	-
8	IPMN (low grade)	-	p.R844C (3%)	-	-	-	-	-
9	IPMN (low grade)	p.G12D (20%)	-	-	-	-	-	-
10	IPMN (low grade)	p.G12V (6%)	-	-	-	-	-	-
11	IPMN (low grade)	p.G12A (5%)	p.R844H (10%)	-	-	-	p.T41I (6.6%)	-
12	Probable mucinous cyst	-	p.R844H (35%)	-	-	-	-	-
13	Probable mucinous cyst	p.Q61H (2%)	-	-	-	-	-	-
14	Probable mucinous cyst	-	-	p.P85A (2.9%)	-	-	-	-
15	Probable mucinous cyst	-	-	p.V217A (2.5%)	-	-	-	-
16	PSC	-	p.A866G (1.5%)	-	p.L85P (1.5%)	-	-	-
17	EC	-	-	-	-	-	-	-
18	LEC	-	-	-	-	-	-	-
19	SCA	-	-	-	-	-	-	p.Y350H (7%)
20	Unknown	-	-	-	p.R161* (6%)	-	-	-
21	Unknown	-	p.R844H (2.2%)	-	-	-	-	-

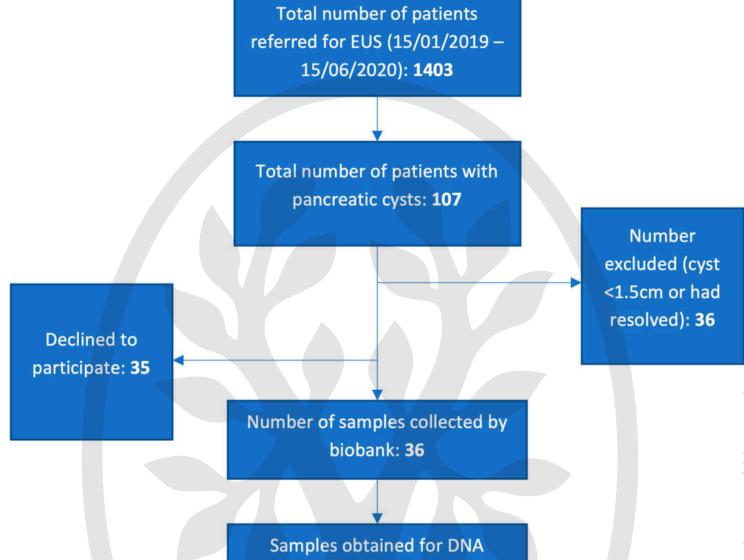| 22 | Unknown | - | - | p.V217A (2.6%) | p.K171* (14.1%) | - | - | - | |
| 23 | Unknown | p.G12V (4%) | - | - | - | - | - | - | |
| 24 | Unknown | p.Q61H (11%) | - | - | - | p.Q31G (4.1%) | - | - | |

Table 3: Confirmation of GNAS, KRAS, TP53 and VHL variants detected in cyst wall DNA using DNA extracted from cyst fluid. [†]indicates variant that is present in NGS data but falls below Mutect2 quality filters.

Patient	Cyst wall variant (MAF)	Present in cyst fluid (MAF)
7	KRAS G12D (8%)	KRAS G12D (45%)
7	GNAS R844H (8%)	GNAS R844H (39%)
8	GNAS R844C (3%)	GNAS R844C (7%)
9	KRAS G12D (20%)	KRAS G12D (40%)
12	GNAS R844H (35%)	NA[†]
13	KRAS Q61H (2%)	NA
16	GNAS A866G (1.5%)	NA
16	VHL L85P (1.5%)	NA
20	VHL R161* (6%)	NA
22	TP53 V217A (2.6%)	NA
22	VHL K171* (14%)	NA[†]

Table 4: Sensitivity, specificity and diagnostic accuracy of different investigations.

Investigation	Sensitivity (%)	Specificity (%)	Diagnostic accuracy (%)
Fluid CEA+ Amylase	12.5	100	30
Cytology	23.08	100	33.3
EUS+ Cytology	40	100	52
NGS	93.3	50	84.2
Standard of care (SOC)	66.6	50	63.1
NGS+ SOC	100	50	89.4