Study Protocol, Version 5

Research Study to Determine the Genetic Causes of Primary Sclerosing Cholangitis and Childhood Autoimmune Liver Disease

Lead investigators:

- Dr Palak Trivedi – Chief Investigator / Cohort Utilisation Lead

- Dr Graeme Alexander
- Dr Carl Anderson
- Dr Roger Chapman
- Dr Mark Hudson
- Professor Deirdre Kelly
- Dr George Mells
- Dr Simon Rushbrook
- Dr Marianne Samyn
- Dr Richard Sandford
- Dr Douglas Thorburn

All enquiries to:
Dr Palak Trivedi
Office 6, Level 2
Institute of Translational Medicine
Heritage Building (Old Queen Elizabeth Hospital)
Mindelsohn Way
Edgbaston,
Birmingham
B15 2THTel: 0121 371 8173

E-mail: P.J.Trivedi@bham.ac.uk

1. Background

1.1 Epidemiology

Primary Sclerosing Cholangitis (PSC), a condition characterised by progressive inflammation, obstruction and fibrosis of the intra- and extra-hepatic bile ducts of the liver, presents typically in the fourth or fifth decades leading often to cirrhosis, end stage liver disease and a need for liver transplantation[1-6]. However presentation across all ages, including young children, is well described. In children the label autoimmune sclerosing cholangitis is also used.

The incidence and prevalence of the disease have been validated in 2 large series. In the first, a study from Norway, 17 of 130 000 inhabitants developed PSC over a 10-year period, giving a mean annual incidence 1.3/100,000 and a point prevalence of 8.5/100,000[7, 8]. A second study in Minnesota identified 22 patients with PSC between 1976 and 2000; the age adjusted incidence was 1.25/100,000 for men and 0.54/100,000 for women. Little is known of the effect of race on the incidence and prevalence of primary sclerosing cholangitis, although it is rare in Alaskans[9]. In the USA in children the incidence and prevalence of autoimmune biliary diseases such as PSC are reported as being 0.3 and 2.1 cases per 100,000 children.[10]

PSC is associated with Ulcerative colitis. It is estimated that 2 - 26% of patients with inflammatory bowel disease (ulcerative colitis in particular) have PSC; population based studies from Scandinavia are consistent and suggest 2 - 7.6% of patients with ulcerative colitis have PSC[11-15]. The pattern of ulcerative colitis associated with PSC appears to differ from that without PSC in that rectal sparing is more common, the inflammatory bowel disease is more quiescent and tends to be right sided with backwash ileitis. In general the prevalence of PSC is much lower in Crohn's disease (1%).

1.2 Clinical Presentation

The symptoms, when present, are characteristic of cholestasis and include pruritus, jaundice and fatigue, but PSC may be asymptomatic (15 - 45%). Deficiency of fat soluble vitamins is common. Osteoporosis is a common complication[16] and a third of patients have fractures following liver transplant. Patients may have cirrhosis at presentation, although these are more likely to be symptomatic.

Liver failure manifest as deep jaundice, ascites, variceal bleeding[17] and encephalopathy are the ultimate consequence of progressive liver injury and evolution to cirrhosis after 12 years. Recurrent bacterial cholangitis due to biliary outflow obstruction is a serious and common complication[4].

Patients with PSC face an increased risk of malignancy. Chronic biliary inflammation is associated with dysplasia and with subsequent evolution to cholangiocarcinoma. The lifetime risk of cholangiocarcinoma in PSC is elevated massively, such that between 4 - 20% of cases develop this complication [18-20]. There is also a heightened risk of hepatocellular carcinoma (2% of patients with PSC and cirrhosis), pancreatic carcinoma[21] (a 10 - 14 fold increased risk) and gallbladder carcinoma. In addition, the risk of colorectal carcinoma and dysplasia are increased in patients with PSC. The first study to demonstrate this, from Sweden, revealed that the absolute cumulative risk for developing colorectal carcinoma in those with PSC and ulcerative colitis was 9%, 31% and 50% at 10, 20 and 25 years of disease duration; in contrast, in those with ulcerative colitis alone, the corresponding risk was 2%, 5% and 10% respectively.[22]

The coexistence of classical features of both AIH and PSC have been described in paediatric and adult settings. Despite studies suggesting an immune mechanism, there is currently no definitive pathogenic basis for their association. Children with sclerosing cholangitis and added features of autoimmunity, account for up to 49% of childhood onset autoimmune liver disease and are often termed autoimmune sclerosing cholangitis (ASC), inferring a distinct clinical entity. Whereas, in adult practice, the concomitant occurrence of AIH and PSC is designated AIH-PSC overlap and more commonly considered a phenotypic variant of either condition. Moreover, the paediatric term ASC may not be synonymous with the adult terminology of AIH-PSC overlap. Particularly, as overlap has only been described in 8-18% of adults diagnosed with AILD.

The lack of a uniform nomenclature remains a key obstacle to understanding the long-term outcomes of children transitioned into adult liver care and furthermore, hinders paediatric involvement in emerging large clinical registries. AIH and PSC may each progress to end-stage liver disease necessitating liver transplantation. Immunosuppression may alter the disease course in AIH, however, their efficacy in ASC remains unknown, while no effective therapy exists for PSC. To date, only two longitudinal studies of childhood onset AILD have been performed, having median observations periods of less than ten years and no adult follow-up. We therefore need to also characterize the nature and progression of childhood onset autoimmune liver disease in patients transitioned from paediatric to adult liver care, while clarifying if ASC is a distinct entity at onset and follow-up.

1.3 Diagnosis

There is no specific blood test. Most cases of PSC come to the attention of the clinician because of abnormal liver blood tests. A raised alkaline phosphatase of biliary origin is typical. Anti-nuclear antibody or smooth muscle antibody is detected in 20 - 70% of cases. pANCA is positive in 65 - 95% of cases[23-26] and can be detected in up to 25% of first degree relatives.

The diagnosis of PSC is based on a number of modalities of which radiology is undoubtedly the most useful. The biliary tree can be imaged by both ERCP [27] and MRCP[28-32] which reveal multifocal strictures within the biliary tree with associated intervening dilatations. MRCP is non-invasive and the preferred imaging of choice with a reported sensitivity of 88% and 99% specificity. ERCP carries a complication rate which can be as high as 14% in symptomatic patients and 2% in asymptomatic patients. Paediatric practice requires a related but slightly extended diagnostic approach

The role of liver histology can be informative in the diagnosis and clinical management of patients with PSC. However, some studies have argued against this[33]. One potential problem with liver biopsy is that regional sampling may occur. In one study cirrhosis was under-estimated in 37% of cases from a cohort of 55 patients biopsied twice on the same occasion. Biopsy may show bile duct injury, with proliferation, periportal inflammation, fibrosis and the almost diagnostic onion skin fibrosis (seen in 13%). Overlapping features with auto-immune hepatitis may be observed.

1.4 Differential Diagnosis

Primary sclerosing cholangitis must be distinguished from disorders that cause similar cholangiographic features on imaging (i.e. HIV, gallstones, etc).

Genetics

*P*rimary Sclerosing Cholangitis is thought to develop as a result of the interaction between the human genome and the environment. If we are to begin to understand PSC we must

understand the concept that individual genetic variants of the human genome are neither sufficient nor necessary for this complex disease to develop but instead act as disease risk (*i.e.* susceptibility) factors. PSC, like the vast majority of human diseases, is complex genetically. Complex diseases are multifactorial, the result of interplay between genes and the environment. Thus, the strong correspondence between genotype and disease phenotype characteristics of Mendelian disorders is not present in this complex disease.

Currently, the best means for quantifying the genetic and environmental influences on complex disease is comparison of disease concordance between monozygotic (i.e. identical) and dizygotic (i.e. fraternal twins). Monozygotic twins share 100% of their DNA, so that disease concordance is suggestive of genetic influence and conversely, discordance illustrates the likely extent of environmental effect. At present there is no peer review study of this kind in PSC.

In addition to twin studies, familial aggregation provides a means to estimate the level of genetic influence in complex diseases. Because family members are more likely to share genetic material among themselves than with the general population, they are also more likely to share genetic characteristics associated with increased disease risk. However, family members (especially close siblings) share environmental exposure that may also contribute to familial disease aggregation.

To date, genome-wide studies have uncovered 23 susceptibility loci for PSC-IBD, the majority of which have been previously reported as risk factors in other immune-mediated disorders. For most candidates, the pathological relationship to PSC-IBD remains largely unknown. Several candidate genes appear to be liver related but the large majority relate to immunity and reaffirm that alterations to immune function, trafficking, and tolerance are likely to influence susceptibility and presentation of PSC-IBD. Similar to most immune-mediated diseases, the strongest association in PSC-IBD resides within the human leukocyte antigen complex and suggests that disease-specific antigens drive pathogenic immune responses. Although genetic predisposition influences disease, genetic determinants account for less than 10% of total disease liability in PSC-IBD, clearly emphasizing the predominant role of environmental factors on disease susceptibility.

Genetic studies define PSC-IBD as a unique disease to IBD, mirroring clinical observations. Most risk loci harbour immune-related genes and disease variants are likely to perturb immune function, tolerance, and/or trafficking. Additional studies in patients and novel experimental systems are needed to identify the origin and impact of environmental factors in relation to genetic predisposition in PSC-IBD[34].

At the present time the genetic contribution to PSC is therefore recognised but is limited due to a failure to derive a unifying hypothesis and up to now by a candidate gene approach.

2. Methodology

2.1 Aim and Hypothesis

The UK PSC Study began recruitment in 2008, as the **PSC Genetics Study,** with the aim of sourcing DNA from over 3000 patients with Primary Sclerosing Cholangitis. To date over 2000 patients have participated in the study. This is a significant step towards our goal of identifying genes of potential significance in the pathogenesis of this condition.

The primary aim of this study is to identify genetic and clinical factors which contribute to the development and outcome of Primary Sclerosing Cholangitis and childhood autoimmune

liver disease. The study design is a case control genetic association study, aligned to clinical and laboratory data on patients.

The null hypothesis is that there is no association between genetic and clinical factors and PSC. The alternative hypothesis is that there is an association between genetic and clinical factors and PSC.

A consortium of investigators has been established that are involved directly or indirectly in the clinical care of adult patients with Primary Sclerosing Cholangitis and Autoimmune Sclerosing Cholangitis, and children with autoimmune liver disease within the United Kingdom.

The Chief Investigator is Dr Palak Trivedi of the University of Birmingham. The sponsor of the study is the University of Cambridge and Cambridge University Hospitals NHS Foundation Trust. The infrastructure support to maintain recruitment and data analysis is from support to the Investigators by the NIHR Birmingham Biomedical Research Centre, as well as funded research collaborations with Industry.

Over time there has been better recognition and understanding of how PSC presents, and its impact across all ages. Since 2016 we have been working with dedicated, expert paediatric clinicians, to offer young people with an established diagnosis of PSC (also referred to in paediatric practice sometimes as autoimmune sclerosing cholangitis) the chance to take part in this research. We are now opening up the study to children with AIH, with or without possible biliary disease because of the recognition that such children over time may ultimately develop a biliary disease.

2.2 Inclusion Criteria

Adults

- Adults with a confirmed diagnosis of Primary Sclerosing Cholangitis or Autoimmune Sclerosing Cholangitis, who have been informed of their diagnosis.

Children

- Children (aged 0-17 years) with a confirmed diagnosis of any of the following:
  - Primary Sclerosing Cholangitis
  - Autoimmune Sclerosing Cholangitis
  - Autoimmune Hepatitis with possible biliary disease
  - Autoimmune Hepatitis with no biliary disease at time of recruitment

Parents or guardians and/or children must have been informed of the child's diagnosis. The above criteria have been set because of the difficulty in obtaining a confirmed PSC diagnosis in children.

2.3 Exclusion Criteria

All cases must be distinguished from causes of secondary sclerosing cholangitis which include: Congenital Biliary tree abnormalities, previous biliary surgery, bile duct carcinoma, HIV cholangiopathy, Primary Biliary Cholangitis, Sarcoid, graft-versus-host disease, sickle cell disease and immunodeficiency related cholangiopathy and drug reactions. The individual collaborators at each centre will ensure that secondary causes of an abnormal cholangiogram have been excluded.

In children other conditions such as viral hepatitis, Wilson disease, drug induced liver injury need to be excluded by collaborators at each centre.

2.4 Confirmation of Disease Phenotype

Patients will be identified as having PSC or the disease variants outlined in section 2.2 based on a number of criteria; analysis will be stratified to the clinical phenotype collected. However, a number of supplementary investigations can support or refute the diagnosis; further, the use of supplementary investigations may help to identify sub groups within the group with a diagnosis of PSC based on MRCP/ERCP/liver biopsy.

2.4.1 Appropriate biliary imaging

The diagnosis of Sclerosing Cholangitis is based on a number of criteria of which perhaps radiological criteria are the most useful.

The biliary tree can be imaged by either Endoscopic Retrograde Cholangio-pancreatography (ERCP) or by Magnetic Resonance Cholangio-pancreatography (MRCP). Both these imaging modalities show characteristic features which allow correct disease phenotyping.

2.4.2 Liver histology

Providing the patient has had appropriate positive biliary imaging, liver biopsy will not be considered an essential criterion for inclusion. In children with suspicion of autoimmune liver disease a liver biopsy will be performed, ideally prior to commencing treatment but providing there are no specific contra-indications including presence of a coagulopathy.

Patients with an abnormal biopsy yet normal biliary imaging will be collected but will be labelled as Small Duct Primary Sclerosing Cholangitis.

2.4.3 Blood tests

LFTs will be recorded, but will not be part of the diagnostic inclusion criteria because, whilst a raised alkaline phosphatase of biliary origin is a typical abnormality, 10% of patients with PSC have a normal Alkaline Phosphatase. Positive Anti-nuclear-antibody (ANA) or smooth muscle antibodies (SMA) occur in 20 - 70% of cases. Similarly pANCA is positive in 65 - 95% of PSC. The presence or absence of these will be recorded for each patient, but will not be a diagnostic criterion as they are neither 100% specific or sensitive.

2.4.4 Other criteria

Patients that have received an orthotopic liver transplant for PSC can be considered as potential study participants.

The presence of inflammatory bowel disease will be recorded, but is not an essential criterion for diagnosis.

Age, gender and ethnic origin will be recorded.

Patients will be recruited from throughout the UK. This is to allow the largest possible number of samples. We expect from epidemiological studies that most patients would be Caucasian.

2.5 How will patients be approached?

The process for recruitment and participation of adult and paediatric patients is outlined on the flow charts in Appendices A and B ('Appendix A – UK PSC Participation Flow Chart – adults v3' and 'Appendix B – UK PSC Participation Flow Chart – children v2').

2.5.1 Adults

The clinical care team in collaborating centres will identify potential participants by screening their electronic patient record systems and databases for patients with any of the above inclusion criteria.

For each patient who is eligible for the study, the clinical care team will record the patient's name and address in the Invitation App and use the app to generate the standard letter of invitation for the patient. The clinical care team will send the standard letter to the patient, inviting him or her to participate in the study. The letter of invitation will be accompanied by a return slip that includes the patient's NHS number (in England and Wales) or CHI number (in Scotland). The invitation and return slip will be accompanied by a freepost envelope addressed to the UK PSC research office.

Patients who are interested in participating in the study will sign the return slip and send it to the UK PSC research office using the freepost envelope provided,

Using the Clinical Database, the clinical care team will be able to see whether (and when) a patient has sent a signed return slip to the UK PSC research team. If they have not, the clinical care team may contact the patient by telephone to discuss the study. Alternatively, the clinical care team may discuss the study with the patient when he or she attends his or her next appointment in the Gastroenterology or Hepatology outpatient clinic.

2.5.2 Children

Appropriate patients will be identified at dedicated paediatric clinics. The child and their parents/guardian will be introduced to the study at a routine liver disease clinic appointment. They will not have an invitation letter sent to them, and the child's personal details will not be logged on the UK PSC database Invitation App. The site may, however, choose to send a PIS to the child and their parents/guardian prior to their appointment.

Prior to approaching children and their parents/guardian about the study, a member of the paediatric research team will contact the UK PSC team to inform them that they have identified a suitable participant. They will not provide any personal data on the patient, e.g. their name, address, date of birth or NHS/CHI number.

The UK PSC Data Manager will create an anonymous participant in the UK PSC clinical database, and thus generate a unique UK PSC study no. for the potential participant. They will issue this to the paediatric research team, who will insert it on the consent and assent forms, questionnaires and blood tube labels.

Children are only able to join the study if the hospital they attend for their liver clinical care is approved for UK PSC paediatric recruitment. It is not possible for children to join the study as self-referring patients.

2.5.3 Adult patient reconsenting and resampling

UK PSC adult participants who joined the study on consent form version 2 are being invited to reconsent, give further samples and/or update their clinical status (see section 6 and

'Appendix C – UK PSC Participation Flow Chart – Adult reconsenting and resampling – version 2).

2.6 Samples

Following consent, patients donate two blood samples for the following analyses:

- DNA extraction for the purposes of a genome wide association, and allied genetic analyses. The goal is that ultimately patient presentation can be correlated with patient genetic data.

- Serum extraction for tests to measure markers in the blood associated with PSC including fibrosis markers and IgG4 levels.

2.7 How do patients participate in the study?

2.7.1 Adult patients (aged 18 years and over)

If a patient declares interest in the study by sending back a signed return slip or contacting the UK PSC team, the UK PSC Data Manager will enter the patient's name, postcode and NHS or CHI number into the Clinical Database. This information will be used to import the patient's record from the Invitation App into the Clinical Database or to add a new participant (in cases of self-referral).

Using the Clinical Database, the data manager will generate a recruitment pack for the patient and send it to their home address.

The study pack contains contact details for the UK PSC team, to enable the patient to raise any questions they may have. If they wish to participate, they will initial and sign the Informed Consent Form at home.

They will then do the following:

Provide:

- Two blood samples (one for DNA and one to measure markers in the blood associated with PSC).These samples will be 18mls each (4 tubes) and can be taken at a routine PSC clinic appointment or by their GP.

AND:

- Fill in two questionnaires regarding their condition: one about their PSC and their health in general, and one on their experience of itching symptoms (pruritus)
- Return the questionnaires and samples, plus the ICF, to the UK PSC study team in Cambridge, using the freepost padded envelope provided.

2.7.2 Paediatric patients (aged 0 to 17 years)

At their liver disease clinic appointment the child and parent/guardian will be given the study pack which includes the Parent/Guardian Information Sheet (version 2) and a Participant Information Sheet (PIS) appropriate to the child's age group (PIS age 6-10 version 2, PIS age 11-15, version 2, or PIS age 16-17 years, version 1).

The parents/guardian and child will have the opportunity to read the study pack and to ask questions. They will then be asked if they are willing to participate in the study. If they wish to join, they will sign the following:

Children aged 0-5 yrs

- The Parent/Guardian signs the Parent/Guardian Informed Consent Form (version 2)

Children aged 6-15yrs

- The Parent/Guardian signs the Parent/Guardian Informed Consent Form (version 2)
- The child signs the Child Assent Form (version 2), if they are capable of doing so

Children aged 16-17yrs

- The child signs the PSC Informed Consent Form themselves (version 4).

They will be asked to do the following:

Provide:

- Two blood samples (one for DNA and one to measure markers in the blood associated with childhood AILD). These samples will be 2.5mls each for children aged 0-10 years and 5mls each for children aged 11-17 years. Samples will be taken at a routine liver clinic appointment.
- For their parent/guardian to fill in a questionnaire regarding their child's AILD and health in general (*UK PSC Paediatric Participant Questionnaire,* version 1). If the child is aged 11 or more and capable, they can complete the questionnaire themselves.

The research team at the child's hospital will return to the UK PSC research office:

- The completed participant questionnaire (using the freepost envelope provided)
- The signed Informed Consent Form, any Assent Form and blood samples (using the sample transport box provided).

On receipt of consent the paediatric research team will contact the UK PSC team, providing the patient's name, address, date of birth and NHS/CHI number, for logging in the UK PSC clinical database. The UK PSC Data Manager will then enter these details onto the clinical database, thus enrolling the previously anonymous patient in the study.

Adult and Paediatric Patients

The UK PSC Data Manager will log consents, assent, participant and pruritus questionnaire responses into the clinical database, as events against the patient record. The data manager will scan the ICF and upload the scanned document into the database. The participant's clinical care team will have access to the scanned copy of the initialled and signed ICF via the Clinical Database.

Creating a consent event triggers the generation of a Clinician Questionnaire. The recruiting site is sent a message via the database, asking them to input the data required for the Clinician Questionnaire. Completion of the clinician questionnaire generates an update message to the UK PSC Data Manager, who will then perform quality control checks.

The UK PSC Clinical update will also be sent to participating centres annually, to collect updated information on research participants, including major health events.

2.8 Additional recruitment methods

In addition, PSC Support (www.pscsupport.org.uk) and the UK PSC study website (www.uk-psc.com) will promote details of the study. Patients who wish to take part will then contact the UK PSC team, requesting a study pack.

Finally patients that have had a liver transplant for PSC will be identified. We will request that the UKT and Republic of Ireland Audit and Research Commission identify the name of patients with PSC that have had a liver transplant. The lead clinician for that patient will then write to that individual asking them if they wish to take part.

In addition, a detailed summary of the patient's clinical condition will be sent by the referring clinician to ensure the disease phenotype is accurate. On arrival in Cambridge the sample will be given a unique bar-code; the following data will be recorded:

1) Age
2) Sex
3) How the diagnosis was made
4) Associated Inflammatory Bowel Disease
5) Whether Liver transplantation occurred
6) Associated autoimmune conditions
7) All information from the questionnaire

2.9 Blood donation and DNA and serum storage

The following samples will be collected from patients*.*

Adults (aged 18+)

- Two blood samples (2 x 18mls; 4 tubes)

Children

- Two blood samples (2 x 2.5mls for age up to 10 years; 2 x 5mls for age 11-17 years)

These samples will be sent to Cambridge. Here the sample will be processed by an approved laboratory, as follows:

- Blood sample 1 - DNA extracted and stored.
- Blood sample 2 - serum extracted. Tests will be subsequently undertaken for markers in the blood associated with PSC and the related conditions outlined in section 2.2.

Following analysis, the DNA and serum samples will be transferred to an approved Biorepository for long-term storage, for use in future, ethically-approved studies. The collected samples will be retained in the Biorepository indefinitely.

3. Genetic studies

Using DNA from this collection, compared to control datasets available to investigators, genome-wide association studies, as well as deep sequencing/exome sequencing studies, will be performed. Data may be shared as part of international collaborations. All DNA sent

out for scientific analysis will have a unique bar code attached to it and no patient details will be included.

4. Correlating serum markers with genetic findings and clinical correlates

As outlined above, we will source a second blood sample from participants, for the extraction of serum, to measure markers in the blood associated with PSC. Testing will be carried out for fibrosis markers and IgG4 levels.

5. Informing patients of any clinically-actionable genetic findings

If a clinically-actionable genetic abnormality is identified during analysis of a DNA sample, the research team will inform the participant's clinical care team. In turn, the clinical care team will inform the participant and their parent/guardian (in paediatric cases)and refer them to a clinical geneticist for confirmation of the finding, if the patient so desires. However, we appreciate that some participants might prefer not to know if they have a clinically actionable genetic abnormality. The study ICF (v3 onwards) and Parent/Guardian ICF (version 1 onwards) therefore allows the participant to opt out of being informed. Our policy is based on the Genomics England (GEL) initiative.

6. Follow-up with existing participants

6.1 Patient Resampling

The Informed Consent Form (version 3 onwards) and the Parent/Guardian Consent Form (version 1 onwards) include permission for patients to be recontacted in the future for follow up blood samples. They may be approached at a frequency no greater than annually, to provide one blood sample (maximum 18mls in adults; 2.5mls in children aged 0-10 years; 5mls in children aged 11-17 years) for serum and/or DNA extraction. These samples will be stored in an approved Biorepository, and analysed for markers of disease activity and severity, with correlation to clinical course.

6.2 Patient Recall

To maximise the value of this study, we seek consent to invite participants to take part in future research seeking effective therapies for PSC and/or childhood AILD. The Informed Consent Form (version 3 or higher) or Parent/Guardian Consent Form for child participants (version 1 or higher), includes consent to be directly contacted and invited to participate in other research studies. Such invitations will be based on information about the patient stored in the research database. Invitations may include
ethically-approved patient information sheets for clinical trials of investigational drugs.

7. Data capture

7.1 Data capture from electronic medical records

In some centres, it may be possible to download most of the information collected using the UK PSC Clinician Questionnaire direct from electronic medical records (EMR). In these centres, the information technology (IT) department will be asked to provide these data in bulk every six to twelve months. To obtain clinical information in bulk from EMR, the UK PSC research team will provide the IT department with the NHS (or CHI) numbers of consented participants recruited from that hospital. The IT department will send the bulk data to the UK PSC research team via the N3 network. A list of the data that will be collected in bulk from EMR is provided in Appendix E.

7.2 Data capture from centralized data repositories

In some regions, results of medical investigations are stored in centralized data repositories. For example, results of medical investigations from hospitals across Northern Ireland are stored in the Northern Ireland Electronic Care Record (NIECR) (http://www.nidirect.gov.uk/northern-ireland-electronic-care-record-niecr). Similarly, the results of medical investigations from hospitals in Wales are channeled into the NHS Wales Clinical Portal (http://www.wales.nhs.uk/nwis/page/52547). In England, results of medical investigations from some patients are stored in a data warehouse known as Patient View. Wherever possible, results of medical investigations for participants in the UK-PSC Study will be obtained from centralized data repositories. To obtain linked data, the data repository will be given a list of the NHS (or CHI) numbers of consented participants. The data repository will send the linked data to the UK PSC research team via the N3 network. A list of the data that will be collected in bulk from centralized data repositories is provided in Appendix E.

7.3 Data capture from NHS information centres

Linked information about consented participants will be also obtained from NHS Digital in England (https://digital.nhs.uk/); the Information Services Division (ISD) for NHS National Services Scotland (http://www.isdscotland.org/), NHS Wales Informatics Service (http://www.wales.nhs.uk/nwis/) and the Health and Social Care Board (HSCB) Northern Ireland (http://www.hscboard.hscni.net/). Linked data obtained from these data centres will include information about hospital admissions, outpatient attendances, investigations, prescribing, and date and cause of death. These data will provide some of the information required to fully characterise the participant's disease.

To obtain these linked data, the UK PSC research team will provide NHS Digital, ISD, NHS Wales Informatics Service or HSCB with a list of the NHS numbers of participants in the study (or CHI numbers for Scottish patients). NHS Digital, ISD, NHS Wales Informatics Service or HSCB will provide the requested data linked to these numbers. These data will be transferred to the UK PSC research team using N3, Secure Electronic File Transfer (SEFT) or NSS National Safe Haven. A list of the linked datasets that will be collected from NHS Digital, ISD, NHS Wales Informatics Service and HSCB is provided in Appendix E.

7.4 Data storage

Information collected about participants will be held in the UK-PSC Database. This is a bespoke Microsoft SQL database designed by Dr. Tony Bennett of Illuminaries® Ltd. (http://www.illuminaries.co.uk/). The UK-PSC Database enables electronic data capture in collaborating centres using a secure web-based interface, as well as import of bulk data from obtained from EMR systems. The database is hosted on a dedicated server on the network of the Cambridge University Hospitals NHS Foundation Trust, linked to the N3 network. The N3 network is used for electronic data transfer. The database will be supported by Illuminaries® for the duration of the study. Identifiable information about participants will be destroyed ten years after the end of the study. Non-identifiable information will be retained indefinitely for use in future, ethically-approved studies.

8. Collaboration with third party investigators

8.1 Sharing of clinical and genetic information with third party investigators

The UK-PSC Database contains a large amount of clinical and genetic information about study participants which could be extremely useful to third-party investigators. The UK PSC Study is part of an National Institute for Health and Research (NIHR) research programme

supporting translational research into rare diseases. The NIHR requires us to share data with third-party investigators. From an ethical perspective, data sharing is immensely important because it maximises research outputs without subjecting patients to additional research procedures.

Anonymized information from the UK-PSC Database is available to third-party investigators, subject to approval by a Data Access Committee consisting of independent experts appointed by the Study Steering Group.

8.2 Sharing of DNA samples with third party investigators

The UK PSC Study DNA collection and serum bank are valuable resources for third-party investigators. As the study is part of an NIHR research programme, we are required to permit sample sharing. From an ethical perspective, sample sharing is immensely important because it maximises research outputs without subjecting patients to re-sampling. For these reasons, anonymized DNA samples will be made available to third-party investigators, subject to approval by a Data Access Committee, consisting of members of the study team, plus independent experts appointed by the Study Steering Group and a member of the Information Governance or Data Protection Team at Cambridge University Hospitals NHS Foundation Trust (the sponsor).

Any data exported from the database for third-party investigators will be anonymized, meaning that all identifiable details will be removed, including the NHS or CHI number. However, shared information will be labelled with the unique study identifier. During the study and for ten years after the study has ended, samples and shared information will therefore be 'linked-anonymized' because the study number will be linked to identifiable information retained in the UK-PSC Database. Thereafter, samples and information will be fully anonymized because identifiable information will have been destroyed.
Anonymized genetic information will also be deposited and stored indefinitely in the European Genome-phenome Archive (EGA), maintained by the European Bioinformatics Institute (EBI) in Hinxton, Cambridge, UK. Anonymized genetic information stored in the EGA may be shared with legitimate, third-party researchers, in the UK or abroad by managed access. Third-party researchers will be made to sign a legally-binding Data Access Agreement in which they commit to protect the confidentiality of participants and use the genetic data for research purposes only.

9. References

1. Chapman, R.W.G., et al., *Primary Sclerosing Cholangitis - a Review of Its Clinical-Features, Cholangiography, and Hepatic Histology.* Gut, 1980. **21**(10): p. 870-877.
2. Wiesner, R.H., et al., *Primary Sclerosing Cholangitis - Natural-History, Prognostic Factors and Survival Analysis.* Hepatology, 1989. **10**(4): p. 430-436.
3. Angulo, P. and K.D. Lindor, *Primary sclerosing cholangitis.* Hepatology, 1999. **30**(1): p. 325-332.
4. MacFaul, G.R. and R.W. Chapman, *Sclerosing cholangitis.* Current Opinion in Gastroenterology, 2004. **20**(3): p. 275-280.
5. Talwalkar, J.A. and K.D. Lindor, *Primary sclerosing cholangitis.* Inflammatory Bowel Diseases, 2005. **11**(1): p. 62-72.
6. Bambha, K., et al., *Incidence, clinical spectrum, and outcomes of primary sclerosing cholangitis in a United States community.* Gastroenterology, 2003. **125**(5): p. 1364-1369.
7. Schrumpf, E. and K.M. Boberg, *Epidemiology of primary sclerosing cholangitis.* Best Practice & Research in Clinical Gastroenterology, 2001. **15**(4): p. 553-562.

8. Boberg, K.M., et al., *Incidence and prevalence of primary biliary cirrhosis, primary sclerosing cholangitis, and autoimmune hepatitis in a Norwegian population.* Scandinavian Journal of Gastroenterology, 1998. **33**(1): p. 99-103.
9. Hurlburt, K.J., et al., *Prevalence of autoimmune liver disease in Alaska natives.* American Journal of Gastroenterology, 2002. **97**(9): p. 2402-2407.
10. Deneau, M., et al., *Primary Sclerosing Cholangitis, Autoimmune Hepatitis, and Overlap in Utah Children: Epidemiology and Natural History.* Hepatology, 2013. **58**(4): p. 1392-1400.
11. Olsson, R., et al., *Prevalence of Primary Sclerosing Cholangitis in Patients with Ulcerative-Colitis.* Gastroenterology, 1991. **100**(5): p. 1319-1323.
12. Rasmussen, H.H., et al., *Primary Sclerosing Cholangitis in Patients with Ulcerative-Colitis.* Scandinavian Journal of Gastroenterology, 1992. **27**(9): p. 732-736.
13. Broome, U., et al., *Liver-Disease in Ulcerative-Colitis - an Epidemiologic and Follow-up-Study in the County of Stockholm.* Gut, 1994. **35**(1): p. 84-89.
14. Aitola, P., A.L. Karvonen, and M. Matikainen, *Prevalence of Hepatobiliary Dysfunction in Patients with Ulcerative-Colitis.* Annales Chirurgiae Et Gynaecologiae, 1994. **83**(4): p. 275-278.
15. Heikius, B., et al., *Hepatobiliary and coexisting pancreatic duct abnormalities in patients with inflammatory bowel disease.* Scandinavian Journal of Gastroenterology, 1997. **32**(2): p. 153-161.
16. Angulo, P., et al., *Bone disease in patients with primary sclerosing cholangitis: prevalence, severity and prediction of progression.* Journal of Hepatology, 1998. **29**(5): p. 729-735.
17. Zein, C.O., K.D. Lindor, and P. Angulo, *Prevalence and predictors of esophageal varices in patients with primary sclerosing cholangitis.* Hepatology, 2004. **39**(1): p. 204-210.
18. Farrant, J.M., et al., *Natural-History and Prognostic Variables in Primary Sclerosing Cholangitis.* Gastroenterology, 1991. **100**(6): p. 1710-1717.
19. Broome, U., et al., *Natural history and prognostic factors in 305 Swedish patients with primary sclerosing cholangitis.* Gut, 1996. **38**(4): p. 610-615.
20. Broome, U., et al., *Primary Sclerosing Cholangitis and Ulcerative-Colitis - Evidence for Increased Neoplastic Potential.* Hepatology, 1995. **22**(5): p. 1404-1408.
21. Bergquist, A., et al., *Hepatic and extrahepatic malignancies in primary sclerosing cholangitis.* Journal of Hepatology, 2002. **36**(3): p. 321-327.
22. Lindberg, B.U., U. Broome, and B. Persson, *Proximal colorectal dysplasia or cancer in ulcerative colitis. The impact of primary sclerosing cholangitis and sulfasalazine - Results from a 20-year surveillance study.* Diseases of the Colon & Rectum, 2001. **44**(1): p. 77-83.
23. Mulder, A.H.L., et al., *Prevalence and Characterization of Neutrophil Cytoplasmic Antibodies in Autoimmune Liver-Diseases.* Hepatology, 1993. **17**(3): p. 411-417.
24. Mulder, A.H.L., et al., *Antineutrophil Cytoplasmic Antibodies (Anca) in Autoimmune Liver-Disease.* Anca-Associated Vasculitides, 1993. **336**: p. 545-549.
25. Kallenberg, C.G.M., A.H.L. Mulder, and J.W.C. Tervaert, *Antineutrophil Cytoplasmic Antibodies - a Still-Growing Class of Autoantibodies in Inflammatory Disorders.* American Journal of Medicine, 1992. **93**(6): p. 675-682.
26. Tervaert, J.W.C., et al., *Antineutrophil Cytoplasmic Antibodies in Primary Sclerosing Cholangitis, Ulcerative-Colitis, and Autoimmune-Diseases.* Gastroenterology, 1992. **102**(3): p. 1090-1090.
27. Maccarty, R.L., et al., *Primary Sclerosing Cholangitis - Findings on Cholangiography and Pancreatography.* Radiology, 1983. **149**(1): p. 39-44.
28. Angulo, P., et al., *Magnetic resonance cholangiography in patients with biliary disease: its role in primary sclerosing cholangitis.* Journal of Hepatology, 2000. **33**(4): p. 520-527.

29. Vitellas, K.M., et al., *MR cholangiopancreatography in patients with primary sclerosing cholangitis: Interobserver variability and comparison with endoscopic retrograde cholangiopancreatography.* American Journal of Roentgenology, 2002. **179**(2): p. 399-407.
30. Fulcher, A.S., et al., *Primary sclerosing cholangitis: Evaluation with MR cholangiography - A case-control study.* Radiology, 2000. **215**(1): p. 71-80.
31. Revelon, G., et al., *Primary sclerosing cholangitis: MR imaging findings with pathologic correlation.* American Journal of Roentgenology, 1999. **173**(4): p. 1037-1042.
32. Talwalkar, J.A., et al., *Cost-minimization analysis of MRC versus ERCP for the diagnosis of primary sclerosing cholangitis.* Hepatology, 2004. **40**(1): p. 39-45.
33. Burak, K.W., P. Angulo, and K.D. Lindor, *Is there a role for liver biopsy in primary sclerosing cholangitis?* American Journal of Gastroenterology, 2003. **98**(5): p. 1155-1158.
34. Chung, B.K. and G.M. Hirschfield, *Immunogenetics in primary sclerosing cholangitis.* Curr Opin Gastroenterol, 2017. **33**(2): p. 93-98.