GeCIP Detailed Research Plan Form

Background

The Genomics England Clinical Interpretation Partnership (GeCIP) brings together researchers, clinicians and trainees from both academia and the NHS to analyse, refine and make new discoveries from the data from the 100,000 Genomes Project.

The aims of the partnerships are:

1. To optimise:

- clinical data and sample collection
- clinical reporting
- data validation and interpretation.

2. To improve understanding of the implications of genomic findings and improve the accuracy and reliability of information fed back to patients. To add to knowledge of the genetic basis of disease.

3. To provide a sustainable thriving training environment.

The initial wave of GeCIP domains was announced in June 2015 following a first round of applications with expressions of interest in January 2015. In April 2016 we invited the inaugurated Cancer GeCIP domains to develop research plans for 'Gear 2' working closely with Genomics England. Within the Cancer Main Programme, the 'Gear 2' phase of the project refers to recruitment of specific cohorts of patients, inclusion of biopsy tissue (diagnostic/recurrence) and ctDNA in selected cohorts and the initiation of clinical trials in early stage (adjuvant/consolidation) setting. These will be used to ensure that the plans are complimentary and add real value across the GeCIP portfolio and address the aims and objectives of the 100,000 Genomes Project. They will be shared with the MRC, Wellcome Trust, NIHR and Cancer Research UK as existing members of the GeCIP Board to give advance warning and manage funding requests to maximise the funds available to each domain. However, formal applications will then be required to be submitted to individual funders. They will allow Genomics England to plan shared core analyses and the required research and computing infrastructure to support the proposed research. They will also form the basis of assessment by the Project's Access Review Committee, to permit access to data.

Domain leads are asked to complete all relevant sections of the GeCIP Detailed Research Plan Form, ensuring that you provide names of domain members involved in each aspect so we or funders can see who to approach if there are specific questions or feedback and that you provide details if your plan relies on a third party or commercial entity. You may also attach additional supporting documents as relevant (optional).

Additional members can apply to join the GeCIP domain by completing the form on our website found here: http://www.genomicsengland.co.uk/join-a-gecip-domain/.

Genomics England Clinical Interpretation Partnership (GeCIP) Detailed Research Plan Form

Application Summary	
GeCIP domain name	**Upper gastrointestinal cancer**
Project title *(max 150 characters)*	**Upper gastrointestinal cancer research in the 100,000 Genomes Project**

Objectives. *Set out the key objectives of your research. (max 200 words)*

1. To determine the feasibility of analysing biopsy material from a range of Upper gastrointestinal (UGI) cancers. Currently the GeCIP material requirement for DNA would exclude that available from single biopsies. This proposal could be limited to resection specimens, some of which would have received neoadjuvant treatment. This bid would therefore permit a strategy of material harvest beyond biopsy in order to obtain samples yielding robust data.
2. Describe the mutational and other "omic" landscape of UGI cancer
3. To reclassify UGI cancer within and between classical anatomical sites
4. To collaborate with multiple prospective clinical and translational programmes extant and planned in the UGI portfolio, in particular, cross cutting programmes such as prospective lifestyle analyses.

Lay summary. *Information from this summary may be displayed on a public facing website. Provide a brief lay summary of your planned research. (max 200 words)*

Upper gastrointestinal (UGI) cancer is caused by uncontrolled cell growth in the various organs of the upper gastrointestinal tract (commonly the oesophagus, stomach, pancreas, liver, and gallbladder). There are approximately 41,304 UGI cancers diagnosed and 37,736 deaths in the UK per annum. The majority of patients are only identified when they already have advanced disease and carrying out surgery on those whose cancer is suitable for it, can often be challenging. Other treatments, such as the use of various different drugs, have improved outcome but the benefit is negligible.
UGI cancers comprise a number of different conditions that have been grouped by virtue of anatomy and surgical approach. This leads to a mixture of different types of cancer, and therefore to a wide range of research in the field. The overall aim of the Upper Gastrointestinal Cancer GeCIP Domain is to use the data resulting from the 100,000 Genomes Project to describe the landscape of UGI cancer and to allow improved understanding and treatment.

Expected start date	**Q2 2017**
Expected end date	**Q2 2020**

Lead Applicant(s)	
Name	Prof John Bridgewater
Post	Consultant and Senior Lecturer in Medical Oncology
Department	
Institution	University College Hospital
Current commercial links	

Deputy Lead Applicant(s)	
Name	Prof Rebecca Fitzgerald

Post	MRC Programme Leader (tenure) at the MRC Cancer Unit, and an Honorary Consultant in Gastroenterology and General Medicine at Addenbrooke's Hospital, Cambridge.
Department	
Institution	Addenbrooke's Hospital, Cambridge
Current commercial links	

Gear 2 Substudies
GI01: Genomic description of patients receiving cisplatin gemcitabine and pembrolizumab on EORTC -1607-GITCG study
GI02: Genomic description of prognosis and chemosensitivity in resected biliary tract cancer
GI03: Genomic description of prognosis and chemosensitivity in resected biliary tract cancer
GI04: Genomic description of prognosis and chemosensitivity in resected biliary tract cancer
GI05: Evolving genomic landscape as a determinant of the selective therapeutic response of pancreatic cancer.
GI06: Investigating role of radiotherapy in the post-operative setting in gastric cancer
GI07: Investigating role of postoperative radiotherapy in extrahepatic cholangiocarcinoma and gallbladder carcinoma (EHCC and GBCA)
GI08: Mutational evolution in resectable hepatocellular carcinoma
GI09: Mutational evolution in hepatocellular carcinoma
GI10: Mutational evolution in metastatic cholangiocarcinoma
GI11: Impact of therapy on molecular genetics of HCC
GI12: Genetics of high-grade neuroendocrine tumours pre and post chemotherapy.
GI13: Genetics of pancreatic neuroendocrine tumours pre and post systemic therapy.
GI14: Investigation of circulating predictive biomarkers in gastric adenocarcinoma.
GI15: Mutational evolution in synchronous stage 4 metastatic colorectal cancer
GI16: Investigation of the genomic evolution and mechanisms of cisplatin resistance in gastric adenocarcinoma
GI17: Modelling the progression from infection and inflammation to gastric cancer
GI18: Gastric cancer profiling across space and time

GeCIP domain - Expression of interest

Full proposal	
Title *(max 150 characters)*	

Research plans. *Give details of the analyses and experimental approaches, study designs and techniques that will be used and timelines for your analysis. Describe the major challenges of the research and the steps required to mitigate these.*

1 Technical optimisation

It is currently not clear whether tissue biopsies can be reliably utilised for DNA sequencing. The majority of patients with UGI cancer would not have surgery and their exclusion be a significant biological and clinical omission for the proposed programme. The default position would be to link material collection to the surgical studies in the portfolio and to closely monitored cohorts such as the hepatocellular carcinoma transplant population. These are detailed below.

2 Driver gene discovery and exploitation

Somatic driver gene identification will rely on statistical methods of assessing mutation over-representation in genes and pathways. These will be backed up by complementary analyses performed for Aim 3. These data will be linked with drug discovery programmes (academic and commercial) and the cross cutting themes to maximally exploit the dataset, clinical and molecular.

3 Complementary analyses

Further analyses outside genome sequencing will be planned, for instance:

- Serum and plasma for proteomics and metabolomics
- Cell free serum for circulating tumour DNA and to assess tumour recurrence
- Germ-line RNA for transcriptomics
- Lymphocyte DNA for epigenetics
- Tumour for RNA expression profiles, tumour epigenetics and proteomics
- Cancer cell lines for study or xenotransplantation cancer
- Skin biopsies for generation of inducible pluripotent stem cells

5. Clinical utility

An infrastructure that will permit clinical utility of the result to an individual patient will be developed using the framework developed by academic clinical facilities such as UCL Advanced diagnostics. A clinical report will be generated which will include a review by a genomics panel to assess clinical relevance.

6. Cross-cutting analyses

We will aim to leverage cross-cancer analyses with respect to the technology e.g. pan-cancer drivers, method comparisons, new analytical tools, collaboration with inherited cancers GeCIP) and pan-GeL analyses (e.g. population genetics, incidental findings) and also other disciplines (see lifestyle analyses below).

7. Molecular predictors of radiation resistance

Using the prospective clinical data from radiation studies in the UGI portfolio (ABC-07, SCALOP-2, SCOPE-2) we will be able to examine the relationship between tumour genome and the efficacy of chemo-radiation.

Education and Training

To ensure long-term durability and success, we will train a new generation of geneticists, molecular pathologists, clinical bioinformaticians, statisticians and biologists. Trainees from these and other programmes participating in allied research projects will be invited to take part in the UGI domain and gain access to the data as appropriate. It is foreseen that opportunities will arise in the multiple host institutions to support science and clinically based fellowships and PhDs as well as European collaborative projects through, for instance, the European Society for medical Oncology infrastructure and funding and charity funding such as the Cholangiocarcinoma organisation. Key to these programmes will be the integration of multidisciplinary data.

Collaborations including with other GeCIPs. *Outline your major planned academic, healthcare, patient and industrial collaborations. This should include collaborations and data sharing with other GeCIPs. Please attach letters of support.*

Role	Experience
Domain leads John Bridgewater j.bridgewater@ucl.ac.uk	JB has lead the clinical research programme in the UK for biliary tract cancers, setting the current standard of care for advanced disease and running a series of studies that are establishing paradigms of care for biliary tract cancer. He chairs the hepatobiliary subgroup of the upper gastrointestinal clinical studies group (UGI CSG) and the internation biliary tract cancer collaborators group. He has been a member of the UGI CSG since 2002 and has multiple translational collaborations through the gastrointestinal portfolio. He is part of the colorectal domain GeCIP bid, specifically an embedded project to harvest material from patients with colorectal liver metastasis as part of a prospective study. He has had a research interest in carcinoma of unknown primary (CUP) which is

	represented in the hepatobiliary sub-group. He has also established a translational programme through a pan-study material collection (funded by CRUK 2014) and a genomic analysis of the ABC-03 series (data in preparation). He has enlisted the supoort of the AMMF (Cholangiocarcinoma user group) and the CUP foundation.
Domain deputy Rebecca Fitzgerald RCF29@MRC-CU.cam.ac.uk	RF works on the underlying clinical, genetic and cell environmental factors that lead to the conversion of a low-risk pre-malignant state into invasive cancer using new diagnostic tools that will identify those patients who are at an increased risk of developing can improve the molecular characterisation of oesophageal adenocarcinoma and identify novel approaches for tumour classification, monitoring and therapy. As such, she has published extensively on the genetics of oesophageal cancer and leads the OCCAMS programme, part of the International Cancer Gene Consortium (ICGC) that has recently puvblished on oesophgeal cancer. The intention is to combine the ICGC and GEL data to provide added value, in particular exploiting the clinical data and further proposed "omics" proposed by GEL.
GMC representative Manuel Rodriguez-Justo	He will also be supported by Manuel Rodriguez-Justo who leads the Genomics England programme at UCL. Manuel is the lead for the Genomics Medicine Centre and is part of the Genomics England pilot. He is also lead pathologist at UCL Advanced Diagnostics (UCLAD) which is the immunohistochemistry, in situ and molecular diagnostics service arm of the UCL-Cancer Institute and a bridge between the Cancer Institute and University College Hospitals NHS Trust (UCLH). It is also the host laboratory for United Kingdom National External Quality Assessment Service for fluorescence immunocytochemistry & in-situ hybridisation.
Patient representative Helen Morement	In 2002 Helen Morement founded AMMF – The Cholangiocarcinoma Charity. Registered with the Charity Commission, no 1091915, AMMF is the UK's only charity dedicated to cholangiocarcinoma. Together with a small team she worked to achieve AMMF's aims of raising the awareness of cholangiocarcinoma, encouraging and supporting research, and providing information and support to patients. Helen Morement acts as patient representative in the NCRI UGI HPB Clinical Trials sub group, the BTC Clinical Trials Management group and the International Biliary Tract Cancer Collaborators group. Helen Morement reviews patient clinical trials information, and information provided by larger, more generalized cancer charities. As well as the AMMF, we have excellent links with: • CUP foundation: John Symons is a member of the clinical studies group • Oesophageal patients association (www.opa.org.uk) • Pancreatic Cancer UK (www.pcuk.org.uk) If successful we will consult the user groups to develop the project.
Education and Training lead Tim Underwood	Timothy Underwood is a Gastro Oesophageal surgeon and MRC Clinician Scientist at the University of Southampton. His laboratory programme focuses on the tumour microenvironment. He is part of the steering committee and the leading surgeon-scientist in the ICGC/OCCAMS collaboration led from Cambridge and a co-author of the recent Nature Genetics papers describing the genetic landscape of oesophageal cancer. He is currently supervising several PhD students. He is a member of the NCRI Upper GI Oesophago-Gastric Sub-Group, a member of the AUGIS research committee and on the Upper GI TMG for the international MRC/CRUK/NIHR AddAspirin trial.
Validation and Feedback representative John Primrose	JP is Professor of Surgery at the University of Southampton with a clinical interest in hepatobiliary cancer. He is Director of the NIHR Hampshire and Isle of Wight Comprehensive Local Research Network and currently Chair of the National Cancer Research Institute Upper GI Clinical Studies Group. He is President of the Association of Surgeons of Great Britain and Ireland. As such he has experience of the greater research community and has worked closely with user groups in delivering clinical research. He is therefore perfectly placed to negotiate the transfer of information in a useable form to the patient.

Other domain members	**Pancreas** S Falk (Chair, clinical trials) S Mukherjee (Clinical oncology) A Biankin (Surgery, molecular pathology for the PrecisionPanc and TCGA programme) P Ghaneh (Surgery) P Corrie (clinical trials) J Valle (clinical trials) **Oesophagogastric** D Cunningham (Chair, clinical trials) T Crosby (Clinical oncology) R Fitzgerald (molecular genetics, on behalf of OCCAMS and TCGA programme) **Hepatobiliary (Biliary, hepatocellular carcinoma, CUP)** J Bridgewater (Chair, clinical trials) M Hawkins (Clinical oncology) T Meyer (medical oncology) J Valle (Clinical trials) J Primrose (Surgery) J Symons (CUP user group) H Morement (Cholangiocarcinoma user group) L Wood (Molecular pathology) B Teh (Molecular pathology) P Ross (Medical oncology) The Bin Tean (National Cancer Centre of Singapore on behalf of TCGA programme) **NET** J Valle (Chair) C Thirlwell (Molecular pathology) D Sarkar(clinical trials)
Potential international collaborators	The Bin Tean (National Cancer Centre of Singapore on behalf of biliary tract TCGA programme). Laura Wood is translational pathologist at John's Hopkins University. It is intended to work closely with the ICGC collaborations to add value to the entire project.

- The OCCAMS and PrecisionPanc programmes have established an extant programme for the assembly of material prospectively in cancer patients. We will collaborate closely with these, ensuring the optimal use of parallel datasets.
- The proposed GIOTTO programme (resected colorectal liver metastasis, also colorectal GeCIP) aims to compare parallel datasets in lifestyle, genomics, circulating DNA and immunological environment prospectively.
- In collaboration with the Faculty of Population and Health Sciences at UCL, we will be examining the potential interaction between genomics and lifestyle (part of prospective data collection in parallel clinical studies.
- In collaboration with the UK HCC transplant programme, we propose to assemble resection material from all transplanted and resected patients within the UK.
- CUP is an area of unmet need for which an ICGC consortium is unlikely ever to be feasible. Molecular data are few and there is an opportunity to describe a significantly under investigated cancer. In collaboration with the CUP Foundation, we would propose a prospective programme to collect sufficient fresh tissue and material for parallel analysis.

If successful we will organise a series of meetings with pharma to formally engage their participation. Confidentiality agreements should permit the early commitment of pharma for compound development. Extant negotiations include the PrecisionPanc programme.

Detailed research plan – Cancer Main Programme Gear 2 studies

Full proposal (total max 1500 words per Gear 2 Substudy)	
Title *(max 150 characters)*	**GI01: Genomic description of patients receiving cisplatin gemcitabine and pembrolizumab on EORTC -1607-GITCG study.**
Cohort details and scientific case	
Cohort eligibility definition (disease type, sub-type, presentation, stage, treatment, clinical characteristics, epidemiological characteristics)	Advanced biliary tract cancer consenting to sequential biopsy at diagnosis, 3 months and at progression
Samples per patient at first ascertainment (primary tumour, LNs, metastatic sites) *It is assumed that in addition there will be one germline sample per patient.*	Metastatic biopsy x3 (3 cores/biopsy)
# cores per tumour (if multi-region biopsying proposed)	
Follow-up samples following first ascertainment	
Purpose of analysis WGS and clinical data from this cohort of patients (brief)	Description of chemotherapy and immunotherapy on genomic landscape
Scientific case and insights that will be gained from this cohort (more details, as indicated)	Description of cohort benefiting from combination treatment. Relationship of genomic landscape to T cell infiltrate, circulating T cell population, change in population with treatment Differential impact on anatomical subgroups (Gall bladder vs hilar CCA vs peripheral CCA)
Is this sub-study dependent on a particular turnaround-time for sample→clinical result? State necessary TAT	No
Is this sub-study dependent on using diagnostic biopsy material?	At diagnosis: Yes Post diagnosis: Yes
Is follow-up with ctDNA important/essential in this cohort?	Yes
Is any other type of molecular analysis of the tumour proposed as part of study of this (e.g. RNA- seq, methylation analysis)?	T-Cell staining for subgroups
Numbers of WGS proposed and recruitment projection	
How many patients do you plan to recruit in total to this cohort? How many WGS would it be proposed were used on this cohort?	70 (10 x germline, 30 x 75x)
How many patients meeting this cohort eligibility present in England per year?	120 (based on ABC studies recruitment)
What number of patients of this cohort eligibility do you anticipate recruiting to 100KGP per year (factoring in non-consent, failure to obtain appropriate samples, technical failure of samples)?	This will be a selected cohort of 10 patients in EORTC -1607-GITCG study
Alignment to clinical trials	
Is it proposed that this sub-study is aligned to	As above

an existing clinical trial/sample collection study? Study details (incl Phase, commercial partner, geographic recruitment, remuneration). Is co-recruitment to this trial/study optional/mandatory for recruitment to this cohort eligibility?	
Is this sub-study a new therapeutic trial?	No

Full proposal (total max 1500 words per Gear 2 Substudy)	
Title *(max 150 characters)*	**GI02: Genomic description of prognosis and chemosensitivity in resected biliary tract cancer**
Cohort details and scientific case	
Cohort eligibility definition (disease type, sub-type, presentation, stage, treatment, clinical characteristics, epidemiological characteristics)	Prospective CRUK funded adjuvant study ACTICCA-1
Samples per patient at first ascertainment (primary tumour, LNs, metastatic sites) *It is assumed that in addition there will be one germline sample per patient.*	Patients will be suitable for randomisation to surveillance or adjuvant chemotherapy with cisplatin and gemcitabine following surgery. As such primary tumour with nodes will be available.
# cores per tumour (if multi-region biopsying proposed)	Re-biopsy of patients with disease progression proposed.
Follow-up samples following first ascertainment	ctDNA at standard follow up intervals to predict recurrence
Purpose of analysis WGS and clinical data from this cohort of patients (brief)	Define subgroups who benefit or not from adjuvant therapy. Define prognosis within these and following subgroups. Define genomic landscape of anatomical subgroups (gall bladder vs hilar CCA vs peripheral CCA) and their sensitivity/prognosis Identify molecular subgroups (e.g.HER-2 driven, IDH-1 related, BRCA driven). Identify longtitudinal variation in genomic landscape (biopsy on progression)
Scientific case and insights that will be gained from this cohort (more details, as indicated)	ICGC data on biliary tract cancer (Singapore Mayo consortium ~60 genomes) soon to be available but not in the context of randomised treatment study. ICGC data will inform direction of interrogation. The proposed study will expand and further inform a currently neglected field.
Is this sub-study dependent on a particular turnaround-time for sample→clinical result? State necessary TAT	No
Is this sub-study dependent on using diagnostic biopsy material?	At diagnosis: No Post diagnosis: Yes
Is follow-up with ctDNA important/essential in	Yes (at study entry, point 2 TBD by BILCAP

this cohort?	outcome)
Is any other type of molecular analysis of the tumour proposed as part of study of this (e.g. RNA- seq, methylation analysis)?	No
Numbers of WGS proposed and recruitment projection	
How many patients do you plan to recruit in total to this cohort? How many WGS would it be proposed were used on this cohort?	5/patient 100 patients
How many patients meeting this cohort eligibility present in England per year?	70 (based on BILCAP recruitment figures)
What number of patients of this cohort eligibility do you anticipate recruiting to 100KGP per year (factoring in non-consent, failure to obtain appropriate samples, technical failure of samples)?	35
Alignment to clinical trials	
Is it proposed that this sub-study is aligned to an existing clinical trial/sample collection study? Study details (incl Phase, commercial partner, geographic recruitment, remuneration). Is co-recruitment to this trial/study optional/mandatory for recruitment to this cohort eligibility?	As above
Is this sub-study a new therapeutic trial?	

Full proposal (total max 1500 words per Gear 2 Substudy)	
Title *(max 150 characters)*	**GI03: Genomic description of prognosis and chemosensitivity in resected biliary tract cancer**
Cohort details and scientific case	
Cohort eligibility definition (disease type, sub-type, presentation, stage, treatment, clinical characteristics, epidemiological characteristics)	Prospective studies in advanced disease on biliary tract NCRI portfolio
Samples per patient at first ascertainment (primary tumour, LNs, metastatic sites) *It is assumed that in addition there will be one germline sample per patient.*	Patients will be suitable for palliative systemic treatment on a number of prospective studies: (Lilly Ramicirumab/Merestinib study, ABC-09 study, EORTC 1607 study)
# cores per tumour (if multi-region biopsying proposed)	Re-biopsy of patients with disease progression proposed.
Follow-up samples following first ascertainment	ctDNA at standard follow up intervals to predict progression
Purpose of analysis WGS and clinical data from this cohort of patients (brief)	Define subgroups who benefit or not from palliative therapy. Define prognosis within these and following subgroups. Define genomic landscape of anatomical subgroups (gall bladder vs hilar CCA vs peripheral CCA) and their sensitivity/prognosis

	Identify molecular subgroups (e.g.HER-2 driven, IDH-1 related, BRCA driven). Identify longtitudinal variation in genomic landscape (biopsy on progression)
Scientific case and insights that will be gained from this cohort (more details, as indicated)	ICGC data on biliary tract cancer (Singapore Mayo consortium ~60 genomes) soon to be available but not in the context of randomised treatment study. ICGC data will inform direction of interrogation. The proposed study will expand and further inform a currently neglected field.
Is this sub-study dependent on a particular turnaround-time for sample→clinical result? State necessary TAT	No
Is this sub-study dependent on using diagnostic biopsy material?	At diagnosis: No Post diagnosis: Yes
Is follow-up with ctDNA important/essential in this cohort?	Yes
Is any other type of molecular analysis of the tumour proposed as part of study of this (e.g. RNA- seq, methylation analysis)?	No
Numbers of WGS proposed and recruitment projection	
How many patients do you plan to recruit in total to this cohort? How many WGS would it be proposed were used on this cohort?	3/patient 50 patients
How many patients meeting this cohort eligibility present in England per year?	70 (based on BILCAP recruitment figures)
What number of patients of this cohort eligibility do you anticipate recruiting to 100KGP per year (factoring in non-consent, failure to obtain appropriate samples, technical failure of samples)?	35
Alignment to clinical trials	
Is it proposed that this sub-study is aligned to an existing clinical trial/sample collection study? Study details (incl Phase, commercial partner, geographic recruitment, remuneration). Is co-recruitment to this trial/study optional/mandatory for recruitment to this cohort eligibility?	As above
Is this sub-study a new therapeutic trial?	

Full proposal (total max 1500 words per Gear 2 Substudy)	
Title *(max 150 characters)*	**GI04: Genomic description of prognosis and chemosensitivity in resected biliary tract cancer**

Cohort details and scientific case	
Cohort eligibility definition (disease type, sub-type, presentation, stage, treatment, clinical characteristics, epidemiological characteristics)	Prospective studies in advanced disease on biliary tract NCRI portfolio
Samples per patient at first ascertainment (primary tumour, LNs, metastatic sites) *It is assumed that in addition there will be one germline sample per patient.*	Patients will be suitable for palliative systemic treatment on a number of prospective studies: (Lilly Ramicirumab/Merestinib study, ABC-09 study, EORTC 1607 study)
# cores per tumour (if multi-region biopsying proposed)	Re-biopsy of patients with disease progression proposed
Follow-up samples following first ascertainment	ctDNA at standard follow up intervals to predict progression
Purpose of analysis WGS and clinical data from this cohort of patients (brief)	Define subgroups who benefit or not from palliative therapy. Define prognosis within these and following subgroups. Define genomic landscape of anatomical subgroups (gall bladder vs hilar CCA vs peripheral CCA) and their sensitivity/prognosis Identify molecular subgroups (e.g.HER-2 driven, IDH-1 related, BRCA driven). Identify longtitudinal variation in genomic landscape (biopsy on progression)
Scientific case and insights that will be gained from this cohort (more details, as indicated)	ICGC data on biliary tract cancer (Singapore Mayo consortium ~60 genomes) soon to be available but not in the context of randomised treatment study. ICGC data will inform direction of interrogation. The proposed study will expand and further inform a currently neglected field.
Is this sub-study dependent on a particular turnaround-time for sample→clinical result? State necessary TAT	No
Is this sub-study dependent on using diagnostic biopsy material?	At diagnosis: No Post diagnosis: Yes
Is follow-up with ctDNA important/essential in this cohort?	Yes
Is any other type of molecular analysis of the tumour proposed as part of study of this (e.g. RNA- seq, methylation analysis)?	No
Numbers of WGS proposed and recruitment projection	
How many patients do you plan to recruit in total to this cohort? How many WGS would it be proposed were used on this cohort?	3/patient 50 patients
How many patients meeting this cohort eligibility present in England per year?	70 (based on BILCAP recruitment figures)
What number of patients of this cohort eligibility do you anticipate recruiting to	35

100KGP per year (factoring in non-consent, failure to obtain appropriate samples, technical failure of samples)?	
Alignment to clinical trials	
Is it proposed that this sub-study is aligned to an existing clinical trial/sample collection study? Study details (incl Phase, commercial partner, geographic recruitment, remuneration). Is co-recruitment to this trial/study optional/mandatory for recruitment to this cohort eligibility?	As above
Is this sub-study a new therapeutic trial?	

Full proposal (total max 1500 words per Gear 2 Substudy)	
Title *(max 150 characters)*	**GI05: Evolving genomic landscape as a determinant of the selective therapeutic response of pancreatic cancer.**
Cohort details and scientific case	
Cohort eligibility definition (disease type, sub-type, presentation, stage, treatment, clinical characteristics, epidemiological characteristics)	Presentation and stage: (i) Borderline resectable (pre and post neoadjuvant therapy –ESPAC-5F). (ii) Resectable. (iii) Advanced metastatic (ACELERATE trial). Treatment: (i) Neoadjuvant: GemCap. (ii) Adjuvant: GemCap. Advanced: Gem, Gem protide (ACELERATE trial).
Samples per patient at first ascertainment (primary tumour, LNs, metastatic sites) *It is assumed that in addition there will be one germline sample per patient.*	Biopsy samples at first ascertainment. Primary tumour and metastatic lesion.
# cores per tumour (if multi-region biopsying proposed)	3
Follow-up samples following first ascertainment	Follow up samples: Blood DNA for germline, recurrent and metastatic tumour samples
Purpose of analysis WGS and clinical data from this cohort of patients (brief)	Pancreatic tumours are heterogenous with sub-clones responding to treatment whilst others are resistant to particular chemotherapies. This is dependent on the specific genomic profile of the clone. During treatment resistant clones will increase and sensitive clone will decline. Monitoring the clonal populations according to genomic profile will allow the evaluation of the relationship between specific populations and disease progression. Thereby, in future, enabling adaptive chemotherapy regimens. Because there will be patients from clinical

	trials we will be able to carry out multi-variate analysis to link outcomes to genotypes.
Scientific case and insights that will be gained from this cohort (more details, as indicated)	We will link the genome analysis to other ongoing research outputs relating to the same patients in order to characterise the changes in genomic profile of pancreatic tumours before and after therapy. This project will provide sequencing of : 1. Biopsy samples from advanced cancer at the time of diagnosis pre-treatment and at disease progression (n=10). 2. Biopsy samples pre and post neoadjuvant therapy (ESPAC-5F) (n=5). 3. Biopsy samples at time of pancreatic resection, at recurrence and disease progression (n=20). This data will be compared to data acquired from matched circulating tumour cells analysis and relevant expression profiles (hENT1, CDA and DPD) carried out in Liverpool.
Is this sub-study dependent on a particular turnaround-time for sample→clinical result? State necessary TAT	Necessary TAT=six months.
Is this sub-study dependent on using diagnostic biopsy material?	At diagnosis: No Post diagnosis: No
Is follow-up with ctDNA important/essential in this cohort?	No
Is any other type of molecular analysis of the tumour proposed as part of study of this (e.g. RNA- seq, methylation analysis)?	All other analysis is taking place in Liverpool with other funding.
Numbers of WGS proposed and recruitment projection	
How many patients do you plan to recruit in total to this cohort? How many WGS would it be proposed were used on this cohort?	1. Pre and post neoadjuvant therapy (ESPAC-5F). Patients=5. WGS=25 2. Pancreatic resection, at recurrence and or disease progression. Patients=20. WGS=100 3. Advanced cancer at the time of diagnosis pre-treatment and at disease progression Patients=10. WGS=50. Total number of patients = 35. Total WGS=175.
How many patients meeting this cohort eligibility present in England per year?	Total number of pancreatic cancer patients in England in 2013 was 7,887. Approximately 15% patients have resectable disease = 1183 patients. Approximately 5-10% patients have borderline resectable disease = 500 patients. Approximately 50-60% patients have metastatic disease = 4,000 patients
What number of patients of this cohort eligibility do you anticipate recruiting to 100KGP per year (factoring in non-consent,	35

failure to obtain appropriate samples, technical failure of samples)?	
Alignment to clinical trials	
Is it proposed that this sub-study is aligned to an existing clinical trial/sample collection study? Study details (incl Phase, commercial partner, geographic recruitment, remuneration). Is co-recruitment to this trial/study optional/mandatory for recruitment to this cohort eligibility?	ESPAC-5F Feasibility phase II, CR-UK funded. ACELERATE. Phase III. Commercial Nucana Biomed Ltd.
Is this sub-study a new therapeutic trial?	All other analysis is taking place in Liverpool with other funding.

Full proposal (total max 1500 words per Gear 2 Substudy)	
Title *(max 150 characters)*	**GI06: Investigating role of radiotherapy in the post-operative setting in gastric cancer**
Cohort details and scientific case	
Cohort eligibility definition (disease type, sub-type, presentation, stage, treatment, clinical characteristics, epidemiological characteristics)	Patients with locally advanced gastric cancer undergoing surgery
Samples per patient at first ascertainment (primary tumour, LNs, metastatic sites) *It is assumed that in addition there will be one germline sample per patient.*	Pre-treatment sample Main surgical sample (tumour and LNs)
# cores per tumour (if multi-region biopsying proposed)	3
Follow-up samples following first ascertainment	follow-up sample at relapse (metastases-peritoneal/liver)
Purpose of analysis WGS and clinical data from this cohort of patients (brief)	determine the effects of neo-adjuvant therapies on cancer genomes, including their evolution, the identification of resistance to systemic and radiation treatment, identification of markers of metastasis EBV status and associated DNA hypermethylation
Scientific case and insights that will be gained from this cohort (more details, as indicated)	The adjuvant therapy of locally advanced gastric cancer has been investigated in successive clinical trials over the last four decades, however the role of adjuvant chemoradiation is still questioned in UK despite a 2013(Ohri) meta-analysis suggesting the addition of radiotherapy to surgery improved OS (HR=0.78(0.70-0.86)p<0.001. We aim to determine features that would indicate a higher risk of locoregional relapse that would aid selecting patients that would benefit from radiotherapy
Is this sub-study dependent on a particular	No

turnaround-time for sample→clinical result? State necessary TAT	
Is this sub-study dependent on using diagnostic biopsy material?	At diagnosis: No Post diagnosis: No
Is follow-up with ctDNA important/essential in this cohort?	Very important. Relating ctDNA burden/mutation to response
Is any other type of molecular analysis of the tumour proposed as part of study of this (e.g. RNA- seq, methylation analysis)?	RNA-seq (inc. tumour stroma) very important.
Numbers of WGS proposed and recruitment projection	
How many patients do you plan to recruit in total to this cohort? How many WGS would it be proposed were used on this cohort?	25 patients, 150WGS
How many patients meeting this cohort eligibility present in England per year?	7000 cases/year ~1500/year having surgery
What number of patients of this cohort eligibility do you anticipate recruiting to 100KGP per year (factoring in non-consent, failure to obtain appropriate samples, technical failure of samples)?	10-15
Alignment to clinical trials	
Is it proposed that this sub-study is aligned to an existing clinical trial/sample collection study? Study details (incl Phase, commercial partner, geographic recruitment, remuneration). Is co-recruitment to this trial/study optional/mandatory for recruitment to this cohort eligibility?	No, however could align to trial
Is this sub-study a new therapeutic trial?	No

Full proposal (total max 1500 words per Gear 2 Substudy)	
Title *(max 150 characters)*	**GI07: Investigating role of postoperative radiotherapy in extrahepatic cholangiocarcinoma and gallbladder carcinoma (EHCC and GBCA)**
Cohort details and scientific case	
Cohort eligibility definition (disease type, sub-type, presentation, stage, treatment, clinical characteristics, epidemiological characteristics)	Patients undergoing radical resection for (EHCC and GBCA)
Samples per patient at first ascertainment (primary tumour, LNs, metastatic sites) *It is assumed that in addition there will be one germline sample per patient.*	Pre-treatment sample Main surgical sample (tumour and LNs)
# cores per tumour (if multi-region biopsying proposed)	3
Follow-up samples following first ascertainment	follow-up sample at relapse (metastases-peritoneal/liver)

Purpose of analysis WGS and clinical data from this cohort of patients (brief)	identification of markers of metastasis identification of genetic markers of poor prognosis
Scientific case and insights that will be gained from this cohort (more details, as indicated)	The benefit of adjuvant chemotherapy and radiation following curative resection has not been established. There is some evidence that adjuvant chemotherapy alone appears to benefit only patients with node positive disease, whilst adjuvant chemotherapy followed by radiation correlates with improved survival for all patients regardless of resection margin status (Hoehn 2015). There is prospective phase II data that systemic chemotherapy following by radiation is tolerated and produces high level of local control (Ben-Joseph SWOG S0809 2015). We aim to determine features that would that would aid selecting patients that would benefit from the additional radiotherapy.
Is this sub-study dependent on a particular turnaround-time for sample→clinical result? State necessary TAT	No
Is this sub-study dependent on using diagnostic biopsy material?	At diagnosis: No Post diagnosis: No
Is follow-up with ctDNA important/essential in this cohort?	Very important. Relating ctDNA burden/mutation to response
Is any other type of molecular analysis of the tumour proposed as part of study of this (e.g. RNA- seq, methylation analysis)?	RNA-seq (inc. tumour stroma) very important.
Numbers of WGS proposed and recruitment projection	
How many patients do you plan to recruit in total to this cohort? How many WGS would it be proposed were used on this cohort?	25 patients, 150WGS of each histological subtype
How many patients meeting this cohort eligibility present in England per year?	~1800 new patients/year in the UK>200 patients undergoing surgery
What number of patients of this cohort eligibility do you anticipate recruiting to 100KGP per year (factoring in non-consent, failure to obtain appropriate samples, technical failure of samples)?	10-15
Alignment to clinical trials	
Is it proposed that this sub-study is aligned to an existing clinical trial/sample collection study? Study details (incl Phase, commercial partner, geographic recruitment, remuneration). Is co-recruitment to this trial/study optional/mandatory for recruitment to this cohort eligibility?	No, however could align to trial
Is this sub-study a new therapeutic trial?	No

Full proposal (total max 1500 words per Gear 2 Substudy)	
Title *(max 150 characters)*	**GI08: Mutational evolution in resectable hepatocellular carcinoma**
Cohort details and scientific case	
Cohort eligibility definition (disease type, sub-type, presentation, stage, treatment, clinical characteristics, epidemiological characteristics)	Hepatocellular carcinoma Defined as surgically resectable or appropriate for transplantation according to specialist HPB MDT
Samples per patient at first ascertainment (primary tumour, LNs, metastatic sites) *It is assumed that in addition there will be one germline sample per patient.*	(i) Primary tumour and background liver sampled at time of resection or transplantation i. Samples from primary lesion (samples from multifocal nodules where appropriate) ii. 5 cores per lesion (ii) Biopsy of metastatic lesion(s) at time of presentation +/- after systemic therapy i. 5 cores per lesion
# cores per tumour (if multi-region biopsying proposed)	As above
Follow-up samples following first ascertainment	As above
Purpose of analysis WGS and clinical data from this cohort of patients (brief)	• Characterise mutational evolution that leads to disease progression • Characterise mutational profile of patients who enjoy long term survival after resection/transplantation for HCC • Identify profiles associated with response/resistance to systemic therapy
Scientific case and insights that will be gained from this cohort (more details, as indicated)	Although the mutational burden within hepatocellular cancer is increasingly well characterised, less is known about the mutational evolution that occurs with progression from primary lesion to metastatic disease (both intra- and extra-hepatic). Sequencing studies in breast cancer have shown that metastases accumulate new mutations, raising the possibility that sequencing advanced disease in hepatocellular carcinoma may uncover additional mutations that are important therapeutic targets or prognostic indicators, and may shed light on the biological processes subverted in the metastatic step. Although surgical resection of hepatocellular carcinoma is associated with improved long-term survival, a significant proportion of patients experience rapid disease recurrence and may therefore derive no benefit from aggressive surgery. Existing methods of identifying patients in whom surgery is likely to

	offer the greatest survival benefit rely on relatively coarse histopathological features. Genomic profiling of patients will identify signatures associated with good long-term outcomes, improving patient stratification. Patients with progressive disease will be treated with loco-regional and systemic therapies. Assessment of recurrent lesions (including lesions that demonstrate response and those that do not) will allow the identification of profiles associated with treatment response. Direct comparison with background liver tissue will also identify potentially actionable pathways in the development of HCC in diseased liver.
Is this sub-study dependent on a particular turnaround-time for sample→clinical result? State necessary TAT	No
Is this sub-study dependent on using diagnostic biopsy material?	At diagnosis: No Post diagnosis: Yes
Is follow-up with ctDNA important/essential in this cohort?	No
Is any other type of molecular analysis of the tumour proposed as part of study of this (e.g. RNA- seq, methylation analysis)?	Global proteomic analysis using iTRAQ
Numbers of WGS proposed and recruitment projection	
How many patients do you plan to recruit in total to this cohort? How many WGS would it be proposed were used on this cohort?	50 patients, 300 WGS
How many patients meeting this cohort eligibility present in England per year?	300
What number of patients of this cohort eligibility do you anticipate recruiting to 100KGP per year (factoring in non-consent, failure to obtain appropriate samples, technical failure of samples)?	25
Alignment to clinical trials	
Is it proposed that this sub-study is aligned to an existing clinical trial/sample collection study? Study details (incl Phase, commercial partner, geographic recruitment, remuneration). Is co-recruitment to this trial/study optional/mandatory for recruitment to this cohort eligibility?	No
Is this sub-study a new therapeutic trial?	No

Full proposal (total max 1500 words per Gear 2 Substudy)	
Title *(max 150 characters)*	**GI09: Mutational evolution in hepatocellular carcinoma**
Cohort details and scientific case	
Cohort eligibility definition (disease type, sub-type, presentation, stage, treatment, clinical characteristics, epidemiological characteristics)	Hepatocellular carcinoma Defined as surgically resectable or appropriate for transplantation according to specialist HPB MDT
Samples per patient at first ascertainment (primary tumour, LNs, metastatic sites) *It is assumed that in addition there will be one germline sample per patient.*	(i) Primary tumour and background liver sampled at time of resection or transplantation i. Samples from primary lesion (samples from multifocal nodules where appropriate) ii. 5 cores per lesion (ii) Biopsy of metastatic lesion(s) at time of presentation +/- after systemic therapy i. 5 cores per lesion
# cores per tumour (if multi-region biopsying proposed)	As above
Follow-up samples following first ascertainment	As above
Purpose of analysis WGS and clinical data from this cohort of patients (brief)	• Characterise mutational evolution that leads to disease progression • Characterise mutational profile of patients who enjoy long term survival after resection/transplantation for HCC • Identify profiles associated with response/resistance to systemic therapy
Scientific case and insights that will be gained from this cohort (more details, as indicated)	Although the mutational burden within hepatocellular cancer is increasingly well characterised, less is known about the mutational evolution that occurs with progression from primary lesion to metastatic disease (both intra- and extra-hepatic). Sequencing studies in breast cancer have shown that metastases accumulate new mutations, raising the possibility that sequencing advanced disease in hepatocellular carcinoma may uncover additional mutations that are important therapeutic targets or prognostic indicators, and may shed light on the biological processes subverted in the metastatic step. Although surgical resection of hepatocellular carcinoma is associated with improved long-term survival, a significant proportion of patients experience rapid disease recurrence and may therefore derive no benefit from aggressive surgery. Existing methods of identifying patients in whom surgery is likely to offer the greatest survival benefit rely on relatively coarse histopathological features.

	Genomic profiling of patients will identify signatures associated with good long-term outcomes, improving patient stratification. Patients who develop recurrence after surgery will be treated with systemic therapy. Assessment of recurrent lesions (including lesions that demonstrate response and those that do not) will allow the identification of profiles associated with treatment response. Direct comparison with background liver tissue will also identify potentially actionable pathways in the development of HCC in diseased liver.
Is this sub-study dependent on a particular turnaround-time for sample→clinical result? State necessary TAT	No
Is this sub-study dependent on using diagnostic biopsy material?	At diagnosis: No Post diagnosis: Yes
Is follow-up with ctDNA important/essential in this cohort?	No
Is any other type of molecular analysis of the tumour proposed as part of study of this (e.g. RNA- seq, methylation analysis)?	Global proteomic analysis using iTRAQ
Numbers of WGS proposed and recruitment projection	
How many patients do you plan to recruit in total to this cohort? How many WGS would it be proposed were used on this cohort?	50 patients, 300 WGS
How many patients meeting this cohort eligibility present in England per year?	300
What number of patients of this cohort eligibility do you anticipate recruiting to 100KGP per year (factoring in non-consent, failure to obtain appropriate samples, technical failure of samples)?	25
Alignment to clinical trials	
Is it proposed that this sub-study is aligned to an existing clinical trial/sample collection study? Study details (incl Phase, commercial partner, geographic recruitment, remuneration). Is co-recruitment to this trial/study optional/mandatory for recruitment to this cohort eligibility?	No
Is this sub-study a new therapeutic trial?	No

Full proposal (total max 1500 words per Gear 2 Substudy)	
Title *(max 150 characters)*	**GI10: Mutational evolution in metastatic cholangiocarcinoma**

Cohort details and scientific case	
Cohort eligibility definition (disease type, sub-type, presentation, stage, treatment, clinical characteristics, epidemiological characteristics)	Intrahepatic/hilar cholangiocacrcinomas arising without known risk factors (PSC, liver fluke infestation etc.) Defined as surgically irresectable by specialist HPB MDT, either because of: 1)Locally advanced disease 2)Metastatic disease
Samples per patient at first ascertainment (primary tumour, LNs, metastatic sites) *It is assumed that in addition there will be one germline sample per patient.*	(i) Primary tumour sampled at time of presentation i. 5 cores per lesion (ii) Biopsy of metastatic lesion(s) at time of presentation i. 5 cores per lesion (iii) Biopsy of metastatic lesion(s) after palliative therapy
# cores per tumour (if multi-region biopsying proposed)	As above
Follow-up samples following first ascertainment	As above
Purpose of analysis WGS and clinical data from this cohort of patients (brief)	To elucidate molecular mechanisms underlying carcinogenesis in the biliary tract. We know very little of the mutational evolution that contributes to disease progression. We therefore propose to generate complete catalogues of somatic mutations and copy number alterations in matched primary and metastatic cholangiocarcinoma. To guide patient selection for treatment based on genomic profile.
Scientific case and insights that will be gained from this cohort (more details, as indicated)	Cholangiocarcinomas are aggressive malignancies with dismal outcomes. Although limited data exists on the prevalence of genetic alterations in biliary tract cancer, these data rely on heterogenous groups often combining intrahepatic, extrahepatic and gallbladder cancer – now recognised as distinct biological entities, and rely on exome sequencing. There is therefore a paucity of data assessing whole genome sequencing, in particular looking at driver mutations and copy-number changes, malignant progression and clonal evolution as well as predicting response to therapy. Although we are beginning to define the mutational burden across a wide range of primary tumours, we still know very little about the mutational evolution that occurs with disease progression. Sequencing studies in breast cancer have shown that metastases accumulate new mutations, raising the possibility that sequencing advanced disease in other cancers will uncover additional mutations

	that may be important therapeutic targets or prognostic indicators. A detailed knowledge of the mutations enriched for in the evolution from primary cholangiocarcinoma to metastases may shed light on the biological processes subverted in this cancer.
Is this sub-study dependent on a particular turnaround-time for sample→clinical result? State necessary TAT	No
Is this sub-study dependent on using diagnostic biopsy material?	At diagnosis: No Post diagnosis: Yes
Is follow-up with ctDNA important/essential in this cohort?	No
Is any other type of molecular analysis of the tumour proposed as part of study of this (e.g. RNA- seq, methylation analysis)?	Global proteomic analysis using iTRAQ
Numbers of WGS proposed and recruitment projection	
How many patients do you plan to recruit in total to this cohort? How many WGS would it be proposed were used on this cohort?	20 patients, 250 WGS
How many patients meeting this cohort eligibility present in England per year?	>500
What number of patients of this cohort eligibility do you anticipate recruiting to 100KGP per year (factoring in non-consent, failure to obtain appropriate samples, technical failure of samples)?	10
Alignment to clinical trials	
Is it proposed that this sub-study is aligned to an existing clinical trial/sample collection study? Study details (incl Phase, commercial partner, geographic recruitment, remuneration). Is co-recruitment to this trial/study optional/mandatory for recruitment to this cohort eligibility?	No
Is this sub-study a new therapeutic trial?	No

Full proposal (total max 1500 words per Gear 2 Substudy)	
Title *(max 150 characters)*	**GI11: Impact of therapy on molecular genetics of HCC**
Cohort details and scientific case	
Cohort eligibility definition (disease type, sub-type, presentation, stage, treatment, clinical characteristics, epidemiological characteristics)	All patients with advanced HCC undergoing first line systemic therapy. 2 cohorts 1. standard of care sorafenib 2. experimental first line therapy
Samples per patient at first ascertainment (primary tumour, LNs, metastatic sites)	1 primary tumour

It is assumed that in addition there will be one germline sample per patient.	
# cores per tumour (if multi-region biopsying proposed)	1
Follow-up samples following first ascertainment	1 follow up core
Purpose of analysis WGS and clinical data from this cohort of patients (brief)	To determine the effect of first line therapy on molecular evolution of HCC and mechanisms of resistance
Scientific case and insights that will be gained from this cohort (more details, as indicated)	To define predictive markers of response and resistance to first line therapy.
Is this sub-study dependent on a particular turnaround-time for sample→clinical result? State necessary TAT	No
Is this sub-study dependent on using diagnostic biopsy material?	At diagnosis: Yes Post diagnosis: No
Is follow-up with ctDNA important/essential in this cohort?	Yes at start of therapy and during follow up – 3 samples
Is any other type of molecular analysis of the tumour proposed as part of study of this (e.g. RNA- seq, methylation analysis)?	No
Numbers of WGS proposed and recruitment projection	
How many patients do you plan to recruit in total to this cohort? How many WGS would it be proposed were used on this cohort?	50 before sorafenib and 50 before experimental therapies
How many patients meeting this cohort eligibility present in England per year?	500
What number of patients of this cohort eligibility do you anticipate recruiting to 100KGP per year (factoring in non-consent, failure to obtain appropriate samples, technical failure of samples)?	100
Alignment to clinical trials	
Is it proposed that this sub-study is aligned to an existing clinical trial/sample collection study? Study details (incl Phase, commercial partner, geographic recruitment, remuneration). Is co-recruitment to this trial/study optional/mandatory for recruitment to this cohort eligibility?	For cohort 2, patients will be receiving experimental therapy in the context of clinical trials.
Is this sub-study a new therapeutic trial?	No

Full proposal (total max 1500 words per Gear 2 Substudy)	
Title *(max 150 characters)*	**GI12: Genetics of high-grade neuroendocrine tumours pre and post chemotherapy.**
Cohort details and scientific case	
Cohort eligibility definition (disease type, sub-type, presentation, stage, treatment, clinical characteristics, epidemiological characteristics)	Metastatic high-grade neuroendocrine tumours undergoing systemic chemotherapy.

Samples per patient at first ascertainment (primary tumour, LNs, metastatic sites) *It is assumed that in addition there will be one germline sample per patient.*	1
# cores per tumour (if multi-region biopsying proposed)	3
Follow-up samples following first ascertainment	3 – if possible post systemic therapy
Purpose of analysis WGS and clinical data from this cohort of patients (brief)	To determine the genetic evolution and development of resistance to therapy in high-grade NETs and identification of novel therapeutic targets and pathways.
Scientific case and insights that will be gained from this cohort (more details, as indicated)	NETs are relatively rare tumours with an incidence of 5/100000 population. However, they have a higher prevalence than any other upper GI cancer at 30/100000. High-grade NETs (ki-67 >20%) have a very poor clinical outcome with the majority progressing after first line chmoetherapy within ~3 months. The median overall survival is around 6 months, making this NET subgroup a significant clinical area of unmet need. Exome and WGS has never been performed in this tumour type, therefore this study would be very informative regarding the pathogenesis of this NET subgroup. Also, factors leading to the development of resistance to chemotherapy have never been undertaken in high-grade NETs, which is the universally agreed first line treatment in this tumour type.
Logistical aspects	
Is this sub-study dependent on a particular turnaround-time for sample→clinical result? State necessary TAT	No
Is this sub-study dependent on using diagnostic biopsy material?	At diagnosis: Yes Post diagnosis: No
Is follow-up with ctDNA important/essential in this cohort?	Follow up with ctDNA is desirable, however it is not essential. ctDNA will be taken pre-treatment, at 3 time points during treatment and post treatment. It would then be taken at 6 weekly intervals during follow up until disease progression. Total – up to 10 ctDNA samples per patient.
Is any other type of molecular analysis of the tumour proposed as part of study of this (e.g. RNA- seq, methylation analysis)?	RNA-seq and methylation analysis will performed on tissue samples and paid for through other grant income. CTC analysis is possible in this patient cohort.
Numbers of WGS proposed and recruitment projection	
How many patients do you plan to recruit in total to this cohort? How many WGS would it be proposed were used on this cohort?	50 – all receiving systemic chemotherapy 150 WGS

How many patients meeting this cohort eligibility present in England per year?	250
What number of patients of this cohort eligibility do you anticipate recruiting to 100KGP per year (factoring in non-consent, failure to obtain appropriate samples, technical failure of samples)?	25
Alignment to clinical trials	
Is it proposed that this sub-study is aligned to an existing clinical trial/sample collection study? Study details (incl Phase, commercial partner, geographic recruitment, remuneration). Is co-recruitment to this trial/study optional/mandatory for recruitment to this cohort eligibility?	No
Is this sub-study a new therapeutic trial?	This is a possibility, however confirmation of funding for this trial is awaited.

Full proposal (total max 1500 words per Gear 2 Substudy)	
Title *(max 150 characters)*	**GI13: Genetics of pancreatic neuroendocrine tumours pre and post systemic therapy.**
Cohort details and scientific case	
Cohort eligibility definition (disease type, sub-type, presentation, stage, treatment, clinical characteristics, epidemiological characteristics)	Locally advanced and metastatic pancreatic neuroendocrine tumours (NETs) pre and post systemic therapy.
Samples per patient at first ascertainment (primary tumour, LNs, metastatic sites) *It is assumed that in addition there will be one germline sample per patient.*	1
# cores per tumour (if multi-region biopsying proposed)	3
Follow-up samples following first ascertainment	3 – if possible post systemic therapy
Purpose of analysis WGS and clinical data from this cohort of patients (brief)	To determine the genetic evolution and development of resistance to therapy in pancreatic NETs.
Scientific case and insights that will be gained from this cohort (more details, as indicated)	NETs are relatively rare tumours with an incidence of 5/100000 population. However, they have a higher prevalence than any other upper GI cancer at 30/100000. Pancreatic NETs are known to harbour mutations in ATRX/DAXX (43%), Menin (44%) and mTOR pathway genes (14%). Tumours harbouring mutations in ATRX/DAXX at surgical resection are known to have a poorer clinical outcome in terms of PFS and OS. To date, it has not been determined whether tumours harbouring mutations in ATRX/DAXX have a higher incidence of developing resistance to systemic therapy (chemotherapy and molecularly targeted

	therapy) which is commonly used in this tumour type. Factors leading to the development of resistance to systemic therapy have never been undertaken in pancreatic NETs. also WGS has never been performed in this tumour type. Our group has previously undertaken integrated genomic analysis of pancreatic NETs and successfully sequenced mutations in ATRX/DAXX and other genes in ctDNA extracted from plasma of patients undergoing treatment. Therefore this study of paired WGS pre and post systemic therapy in pancreatic NETs will inform us for the first time on mechanisms of resistance to therapy in this tumour group. It will also determine whether tumours harbouring ATRX/DAXX mutations more commonly develop resistance to therapy.
Is this sub-study dependent on a particular turnaround-time for sample→clinical result? State necessary TAT	No
Is this sub-study dependent on using diagnostic biopsy material?	At diagnosis: Yes Post diagnosis: No
Is follow-up with ctDNA important/essential in this cohort?	Follow up with ctDNA is desirable, however it is not essential. ctDNA will be taken pre-treatment, at 3 time points during treatment and post treatment. It would then be taken at 6 weekly intervals during follow up until disease progression. Total – up to 10 ctDNA samples per patient.
Is any other type of molecular analysis of the tumour proposed as part of study of this (e.g. RNA- seq, methylation analysis)?	RNA-seq and methylation analysis will performed on tissue samples and paid for through other grant income. CTC analysis is possible in this patient cohort.
Numbers of WGS proposed and recruitment projection	
How many patients do you plan to recruit in total to this cohort? How many WGS would it be proposed were used on this cohort?	60 (20 chemotherapy, 20 sunitinib therapy and 20 everolimus therapy) Total: 400 – 480 (dependent on availability of multi-regional biopsies)
How many patients meeting this cohort eligibility present in England per year?	Approx 300
What number of patients of this cohort eligibility do you anticipate recruiting to 100KGP per year (factoring in non-consent, failure to obtain appropriate samples, technical failure of samples)?	30
Alignment to clinical trials	
Is it proposed that this sub-study is aligned to an existing clinical trial/sample collection study? Study details (incl Phase, commercial partner,	No – clinical trials have previously been completed in chemotherapy, sunitinib and everolimus therapy).

geographic recruitment, remuneration). Is co-recruitment to this trial/study optional/mandatory for recruitment to this cohort eligibility?	
Is this sub-study a new therapeutic trial?	No

Full proposal (total max 1500 words per Gear 2 Substudy)	
Title *(max 150 characters)*	**GI14: Investigation of circulating predictive biomarkers in gastric adenocarcinoma.**
Cohort details and scientific case	
Cohort eligibility definition (disease type, sub-type, presentation, stage, treatment, clinical characteristics, epidemiological characteristics)	Eligible patients: Patients with locally advanced or metastatic gastric adenocarcinoma undergoing palliative chemotherapy. Tumour location: Fundus/body/distal stomach, excluding Siewert type 3/Gastro-oesophageal junction tumours Treatment: Cisplatin-based palliative chemotherapy Clinical characteristics: Diffuse and intestinal subtypes patients with locally advanced/metastatic disease
Samples per patient at first ascertainment (primary tumour, LNs, metastatic sites) *It is assumed that in addition there will be one germline sample per patient.*	Pre-chemotherapy samples: 1 X endoscopic biopsies at diagnostic OGD from different tumour regions.
# cores per tumour (if multi-region biopsying proposed)	1
Follow-up samples following first ascertainment	Biopsy of recurrent/progressive disease: 1 X biopsy of any accessible areas of recurrent/progressive disease (e.g. liver metastases)
Purpose of analysis WGS and clinical data from this cohort of patients (brief)	To identify mutational signatures predictive of benefit to palliative chemotherapy in tumour tissue and blood in gastric adenocarcinoma.
Scientific case and insights that will be gained from this cohort (more details, as indicated)	The Cancer Genome Atlas (TCGA) has identified four molecular subtypes of gastric adenocarcinoma: Epstein-Barr virus associated, Microsatellite Unstable (MSI), Genomically stable and those exhibiting Chromosomal Instability (CIN). However, the role of these subgroups in prognosis and response to therapy has yet to be defined. We propose to sequence a cohort of gastric adenocarcinomas in order to define their molecular subgroup and collate this information with clinical data regarding response to chemotherapy. At present, approximately 45% of patients respond to standard cisplatin-based triplet therapy but it is unclear which molecular

	subgroups are more chemo-sensitive. We propose to identify a mutational signature capable of predicting sensitivity to chemotherapy which may be used to select patients for treatment. A further challenge is that many patients in the advanced setting are diagnosed using small volume, endoscopic biopsies, unsuitable for next generation sequencing techniques. Alongside the collection of tumour tissue we will store whole blood and plasm for the isolation and characterisation of circulating tumour DNA (ctDNA) and exosomes. We will seek to detect the mutational signatures developed in sequencing the tumour tissues in circulating exosomal DNA or ctDNA. In this way we will develop a non-invasive blood-based biomarker assay to be used to select the most effective therapy for advanced gastric adenocarcinoma patients.
Is this sub-study dependent on a particular turnaround-time for sample→clinical result? State necessary TAT	No
Is this sub-study dependent on using diagnostic biopsy material?	At diagnosis: Yes Post diagnosis: Yes
Is follow-up with ctDNA important/essential in this cohort?	Yes, collection of plasma/ whole blood for ctDNA and exosomal DNA is an essential part of this project. Whole blood/plasma to be collected during chemotherapy and follow-up
Is any other type of molecular analysis of the tumour proposed as part of study of this (e.g. RNA- seq, methylation analysis)?	Targeted sequencing/analysis of ctDNA/exoDNA. Exploratory RNAseq may be carried out.
Numbers of WGS proposed and recruitment projection	
How many patients do you plan to recruit in total to this cohort? How many WGS would it be proposed were used on this cohort?	In order to represent all of the molecular subtypes adequately we would require 300 patients. For profiling of the tumour tissue only would require 300 WGS. A targeted sequencing panel would be used for ctDNA/exoDNA analysis.
How many patients meeting this cohort eligibility present in England per year?	5,600 cases/year ~2500/year are suitable for palliative chemotherapy
What number of patients of this cohort eligibility do you anticipate recruiting to 100KGP per year (factoring in non-consent, failure to obtain appropriate samples, technical failure of samples)?	30-40
Alignment to clinical trials	
Is it proposed that this sub-study is aligned to an existing clinical trial/sample collection study? Study details (incl Phase, commercial partner,	Not currently aligned to a clinical trial

geographic recruitment, remuneration). Is co-recruitment to this trial/study optional/mandatory for recruitment to this cohort eligibility?	
Is this sub-study a new therapeutic trial?	No

Full proposal (total max 1500 words per Gear 2 Substudy)	
Title *(max 150 characters)*	**GI15: Mutational evolution in synchronous stage 4 metastatic colorectal cancer**
Cohort details and scientific case	
Cohort eligibility definition (disease type, sub-type, presentation, stage, treatment, clinical characteristics, epidemiological characteristics)	• Synchronous presentation of liver-limited stage 4 colorectal cancer • Deemed suitable for surgical resection by specialist MDT
Samples per patient at first ascertainment (primary tumour, LNs, metastatic sites) *It is assumed that in addition there will be one germline sample per patient.*	Primary colorectal tumour sampled endoscopically at time of diagnosis and prior to neoadjuvant therapy i. Multi-site sampling of primary specimen after resection (5 cores) ii. Multi-site sampling of liver metastasis(es) after resection (5 cores)
# cores per tumour (if multi-region biopsying proposed)	As above
Follow-up samples following first ascertainment	Biopsy of recurrent lesions after surgery (approx. 2/3rds ptnts @ 24 months)
Purpose of analysis WGS and clinical data from this cohort of patients (brief)	• Characterise mutational evolution that leads to disease progression • Characterise mutational profile of patients who enjoy long term survival after resection of stage 4 CRC • Identify profiles associated with response/resistance to systemic therapy
Scientific case and insights that will be gained from this cohort (more details, as indicated)	Although surgical resection of liver-limited metastatic colorectal cancer is associated with improved long-term survival, a significant proportion of patients experience rapid disease recurrence and may therefore derive no benefit from aggressive surgery. Existing methods of identifying patients in whom surgery is likely to offer the greatest survival benefit rely on coarse histopathological features. Genomic profiling of patients will identify signatures associated with good long-term outcomes, improving patient stratification. Direct comparison with genomic information from primary colorectal lesions (which are easily accessible at time of diagnosis by endoscopy) will improve clinical utility of these findings. Although the mutational burden within colorectal cancer is increasingly well

	characterised, less is known about the mutational evolution that occurs with progression from primary tumour to metastatic disease (both oligometastatic and polymetastatic states). Sequencing studies in breast cancer have shown that metastases accumulate new mutations, raising the possibility that sequencing advanced disease in colorectal cancer may uncover additional mutations that are important therapeutic targets or prognostic indicators, and may shed light on the biological processes subverted in the metastatic step. Patients who develop recurrence after surgery will be treated with systemic therapy. Assessment of recurrent lesions (including lesions that demonstrate response and those that do not) will allow the identification of profiles associated with treatment response.
Is this sub-study dependent on a particular turnaround-time for sample→clinical result? State necessary TAT	No
Is this sub-study dependent on using diagnostic biopsy material?	At diagnosis: No Post diagnosis: Yes
Is follow-up with ctDNA important/essential in this cohort?	No
Is any other type of molecular analysis of the tumour proposed as part of study of this (e.g. RNA- seq, methylation analysis)?	Global proteomic analysis using iTRAQ
Numbers of WGS proposed and recruitment projection	
How many patients do you plan to recruit in total to this cohort? How many WGS would it be proposed were used on this cohort?	50 patients, 300 WGS
How many patients meeting this cohort eligibility present in England per year?	800
What number of patients of this cohort eligibility do you anticipate recruiting to 100KGP per year (factoring in non-consent, failure to obtain appropriate samples, technical failure of samples)?	20
Alignment to clinical trials	
Is it proposed that this sub-study is aligned to an existing clinical trial/sample collection study? Study details (incl Phase, commercial partner, geographic recruitment, remuneration). Is co-recruitment to this trial/study optional/mandatory for recruitment to this cohort eligibility?	No
Is this sub-study a new therapeutic trial?	No

Full proposal (total max 1500 words per Gear 2 Substudy)	
Title *(max 150 characters)*	**GI16: Investigation of the genomic evolution and mechanisms of cisplatin resistance in gastric adenocarcinoma**
Cohort details and scientific case	
Cohort eligibility definition (disease type, sub-type, presentation, stage, treatment, clinical characteristics, epidemiological characteristics)	Patients with localised gastric adenocarcinoma undergoing neo-adjuvant chemotherapy followed by surgical resection. Tumour location: Fundus/body/distal stomach, excluding Siewert type 3/Gastro-oesophageal junction tumours Treatment: Cisplatin-based neo-adjuvant chemotherapy Clinical characteristics: Diffuse and intestinal subtypes patients with operable disease
Samples per patient at first ascertainment (primary tumour, LNs, metastatic sites) *It is assumed that in addition there will be one germline sample per patient.*	Pre-chemotherapy samples: 3 X endoscopic biopsies at diagnostic OGD or staging laparoscopy from different tumour regions. 1X endoscopic biopsy from normal gastric mucosa and 1 X endoscopic biopsy from any areas of intestinal metaplasia. Post-chemotherapy samples: 3 X biopsies from different areas of the resected tumour tissue. 1X biopsy from normal gastric mucosa and 1 X biopsy from any areas of intestinal metaplasia. Biopsy of recurrent disease: 1 X biopsy of any accessible areas of recurrent disease (e.g. liver metastases)
# cores per tumour (if multi-region biopsying proposed)	As above
Follow-up samples following first ascertainment	As above
Purpose of analysis WGS and clinical data from this cohort of patients (brief)	To track the genomic evolution of gastric adenocarcinoma during neo-adjuvant chemotherapy in order to identify drivers of cisplatin resistance.
Scientific case and insights that will be gained from this cohort (more details, as indicated)	The molecular events driving resistance to chemotherapy in gastric adenocarcinoma are poorly understood but primary and acquired drug resistance are major determinants of clinical outcome. By analysing tissue collected before and after neo-adjuvant chemotherapy we will track the genomic evolution of the tumours in response to the selective evolutionary pressure of neo-adjuvant treatment. The identification of driver mutations and amplifications maintained throughout neo-adjuvant chemotherapy has the potential to reveal novel therapeutic vulnerabilities in gastric adenocarcinoma. The molecular sub-classification of gastric adenocarcinoma performed by The Cancer

	Genome Atlas has led to insights into the various molecular sub-groups of gastric cancer. However, these groups have not been related to response to therapy or clinical outcome. We seek to identify these molecular subtypes and examine their inherent chemo-sensitivity or resistance to neo-adjuvant treatment. The collection of pre- and post-chemotherapy tissue will allow us to monitor the molecular changes in each group and to identify resistance mechanisms specific to each biology. This work will lead to the identification of drug targets and predictive biomarkers aligned to the molecular classification and will advance revolutionary treatment strategies.
Is this sub-study dependent on a particular turnaround-time for sample→clinical result? State necessary TAT	No
Is this sub-study dependent on using diagnostic biopsy material?	At diagnosis: Yes Post diagnosis: Yes
Is follow-up with ctDNA important/essential in this cohort?	No
Is any other type of molecular analysis of the tumour proposed as part of study of this (e.g. RNA- seq, methylation analysis)?	An integrated genomic approach incorporating RNA-seq (inc. tumour stroma) and methylation status would be desirable.
Numbers of WGS proposed and recruitment projection	
How many patients do you plan to recruit in total to this cohort? How many WGS would it be proposed were used on this cohort?	In order to represent all of the molecular subtypes adequately we would require 50 patients. For profiling of the tumour tissue only this would require 300 WGS
How many patients meeting this cohort eligibility present in England per year?	5,600 cases/year ~1500/year having surgery
What number of patients of this cohort eligibility do you anticipate recruiting to 100KGP per year (factoring in non-consent, failure to obtain appropriate samples, technical failure of samples)?	10-15
Alignment to clinical trials	
Is it proposed that this sub-study is aligned to an existing clinical trial/sample collection study? Study details (incl Phase, commercial partner, geographic recruitment, remuneration). Is co-recruitment to this trial/study optional/mandatory for recruitment to this cohort eligibility?	Not currently aligned to a clinical trial
Is this sub-study a new therapeutic trial?	No

Full proposal (total max 1500 words per Gear 2 Substudy)	
Title *(max 150 characters)*	**GI17: Modelling the progression from infection and inflammation to gastric cancer**
Cohort details and scientific case	
Cohort eligibility definition (disease type, sub-type, presentation, stage, treatment, clinical characteristics, epidemiological characteristics)	Eligible patients: a) Patients with local gastric adenocarcinoma. Tumour location: Fundus/body/distal stomach, excluding Siewert type 3/Gastro-oesophageal junction tumours Treatment: curative intent (surgery +- neoadjuvant chemotherapy) Clinical characteristics: a representative sample of the population will be derived ie those with Epstein-Barr Virus (EBV), (~10%), and/or Helicobacter Pylori (H. Pylori) (~33%), and/or atrophic gastritis (both autoimmune- and H. Pylori-related). Clinical follow-up post-treatment. b) Patients with i) Epstein-Barr Virus (EBV), ii) Helicobacter Pylori (H. Pylori) iii) atrophic gastritis (both autoimmune- and H. Pylori-related) Clinical characteristics: a representative sample of the population will be derived. Clinical follow-up with respect to progression of gastritis/gastric cancer.
Samples per patient at first ascertainment (primary tumour, LNs, metastatic sites) *It is assumed that in addition there will be one germline sample per patient.*	a) Pre-treatment samples: 1 biopsy each from primary site and normal mucosa (plus 1 from gastritis site if present) b) EBV/H Pylori/Gastritis samples: 1 biopsy each from normal mucosa (plus 1 from gastritis site if present)
# cores per tumour (if multi-region biopsying proposed)	As above
Follow-up samples following first ascertainment	As above
Purpose of analysis WGS and clinical data from this cohort of patients (brief)	To identify gene pathways and processes over-represented in progression from infection to inflammation to gastric cancer, with the objective of patient stratification for early intervention and drug targeting.
Scientific case and insights that will be gained from this cohort (more details, as indicated)	We propose to sequence a cohort of gastric adenocarcinomas to facilitate molecular subtyping and compare with a complimentary subtyping of EBV, H Pylori and gastritis patients. Given that EBV and H Pylori infections are present in almost 45% of gastric cancer patients, we wish to identify pathways of progression from infection and inflammatory disease in order to identify at-risk patients for intervention treatment, while considering drug repurposing.

Is this sub-study dependent on a particular turnaround-time for sample→clinical result? State necessary TAT	No
Is this sub-study dependent on using diagnostic biopsy material?	At diagnosis: Yes Post diagnosis: Yes
Is follow-up with ctDNA important/essential in this cohort?	No
Is any other type of molecular analysis of the tumour proposed as part of study of this (e.g. RNA- seq, methylation analysis)?	RNA-seq and methylation.
Numbers of WGS proposed and recruitment projection	
How many patients do you plan to recruit in total to this cohort? How many WGS would it be proposed were used on this cohort?	300 patients
How many patients meeting this cohort eligibility present in England per year?	~7,000 per year (gastric cancer) H Pylori is present in ~ 40% of poulation EBV is present in ~ 90% of population
What number of patients of this cohort eligibility do you anticipate recruiting to 100KGP per year (factoring in non-consent, failure to obtain appropriate samples, technical failure of samples)?	30-40
Alignment to clinical trials	
Is it proposed that this sub-study is aligned to an existing clinical trial/sample collection study? Study details (incl Phase, commercial partner, geographic recruitment, remuneration). Is co-recruitment to this trial/study optional/mandatory for recruitment to this cohort eligibility?	No
Is this sub-study a new therapeutic trial?	No

Full proposal (total max 1500 words per Gear 2 Substudy)	
Title *(max 150 characters)*	**GI18: Gastric cancer profiling across space and time**
Cohort details and scientific case	
Cohort eligibility definition (disease type, sub-type, presentation, stage, treatment, clinical characteristics, epidemiological characteristics)	Diffuse or intestinal pattern gastric cancer of fundus/body/distal stomach excluding GOJ. Any stage with patient suitable for endoscopic biopsy of primary tumour +- biopsy of metastasis and/or recurrent disease at a single or multiple time points. Cohort will consist of 2 sub-classifications. Cohort 1: patients treated with curative intent (surgery +- neoadjuvant chemotherapy) and followed over time with biopsy and sequencing of recurrent disease. Cohort 2: patients with advanced/metastatic disease who will have

	multiple biopsies taken for sequencing from all involved and accessible disease sites at presentation and further biopsy/sequencing of new disease as it occurs.
Samples per patient at first ascertainment (primary tumour, LNs, metastatic sites) *It is assumed that in addition there will be one germline sample per patient.*	Cohort 1: 3 x endoscopic biopsies at diagnostic OGD from different tumour regions (can be at time of staging laparoscopy) followed by 3 x biopsies (8mm punch) from resected tumour specimen at the time of surgery (additional samples for 'omics analysis would be very helpful). Cohort 2: 3 x endoscopic biopsies at diagnostic OGD from different tumour regions (can be at time of staging laparoscopy) + biopsy of metastatic disease from multiple sites (up to 5) preferably taken as excision biopsies if peritoneal disease or visible lesions at staging lap. Biopsy of new disease during follow-up (additional samples for 'omics analysis would be helpful).
# cores per tumour (if multi-region biopsying proposed)	As above
Follow-up samples following first ascertainment	As above
Purpose of analysis WGS and clinical data from this cohort of patients (brief)	To understand the clonal dynamics and evolution of disease recurrence/metastasis with reference to primary tumour heterogeneity and anti-cancer therapy.
Scientific case and insights that will be gained from this cohort (more details, as indicated)	A molecular genetic classification of primary gastric cancer has recently been proposed but requires validation in a clinical context. In addition, although some data exists on the clonal structure and molecular progression of these cancers, very little whole genome sequencing data are available and basic knowledge of the evolution of recurrent and metastatic disease is lacking. To allow the rational design and delivery of new and urgently required novel therapies information regarding driver mutations, copy number changes, clonal dynamics, response to therapy, metastatic evolution and the causes of malignant progression are required. By investigating advanced, metastatic and recurrent gastric cancer over space and time we will gain insight into these processes.
Is this sub-study dependent on a particular turnaround-time for sample→clinical result? State necessary TAT	No
Is this sub-study dependent on using diagnostic biopsy material?	At diagnosis: Yes Post diagnosis: Yes

Is follow-up with ctDNA important/essential in this cohort?	No
Is any other type of molecular analysis of the tumour proposed as part of study of this (e.g. RNA- seq, methylation analysis)?	RNA-seq (inc. tumour stroma) and epigenetic profiling would be highly desirable.
Numbers of WGS proposed and recruitment projection	
How many patients do you plan to recruit in total to this cohort? How many WGS would it be proposed were used on this cohort?	60 patients total (30 in each arm)
How many patients meeting this cohort eligibility present in England per year?	Several hundred
What number of patients of this cohort eligibility do you anticipate recruiting to 100KGP per year (factoring in non-consent, failure to obtain appropriate samples, technical failure of samples)?	25
Alignment to clinical trials	
Is it proposed that this sub-study is aligned to an existing clinical trial/sample collection study? Study details (incl Phase, commercial partner, geographic recruitment, remuneration). Is co-recruitment to this trial/study optional/mandatory for recruitment to this cohort eligibility?	No
Is this sub-study a new therapeutic trial?	No

Data and informatics requirements
All GeCIP Domains have contributed to the construction of the data model for their tumour type and have contributed to ongoing efforts within Genomics England to develop the clinical and research informatics infrastructure.

Data access and security	
GeCIP domain name	**Upper gastrointestinal cancer**
Project title *(max 150 characters)*	**Upper gastrointestinal cancer research in the 100,000 Genomes Project**

Applicable Acceptable Uses. *Tick all those relevant to the request and ensure that the justification for selecting each acceptable use is supported above.*

X *Clinical care*

X *Clinical trials feasibility*

X *Deeper phenotyping*

X *Education and training of health and public health professionals*

X *Hypothesis driven research and development in health and social care - observational*

X *Hypothesis driven research and development in health and social care - interventional*

X *Interpretation and validation of the Genomics England Knowledge Base*

X *Non hypothesis driven R&D - health*

X *Non hypothesis driven R&D - non health*

X *Other health use - clinical audit*

X *Public health purposes*

X *Tool evaluation and improvement*

Information Governance

X The lead for each domain will be responsible for validating and assuring the identity of the researchers. The lead may be required to support assurance and audit activities by Genomics England.

Any research requiring access to the embassy will be required to complete IG Training and read and sign a declaration form. Access will only be granted once these requirements have been met.