GeCIP Detailed Research Plan Form

August 2015

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 in January 2015. On the 18th June 2015 we invited the inaugurated GeCIP domains to develop more detailed research plans working closely with Genomics England. 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 needed 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. Some of you have requested a template for the research plan which we now provide herewith.

We are only expecting one research plan per domain and have designed this form to contain common features with funder application systems to minimise duplication of effort. Please do not hesitate to contact us if you need help or advice.

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 including:

- a cover letter (optional)
- CV(s) from any new domain members which you have not already supplied (required)
- other supporting documents as relevant (optional)

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

Application Summary	
GeCIP domain name	Inherited Cancer Predisposition Domain (InCaP)
Project title *(max 150 characters)*	Integrative studies of Inherited Cancer Susceptibility through the 100KGP: Whole Genome Analyses and translational multidisciplinary characterisation genes and variants

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

Our key objectives come under three themes:
THEME I: Whole Genome analysis of germline genomes for discovery and description
Objective 1: *Comprehensive Whole Genome Analysis of patients with Inherited Cancer Phenotypes recruited to the rare disease program.*
Objective 2: *Comprehensive Whole Genome Analysis of germline genomes from patients recruited to the Cancer program.*
THEME II: Characterisation of genes and variants involved in cancer susceptibility
Objective 1: To develop a definitive resource for interpretation of germline susceptibility variants for clinical decision-making, including explicit methodology for curation, integration of evidence and standardization of classification.
Objective 2: To perform novel multi-disciplinary analyses (clinical, genetic epidemiological, computational and laboratory functional) to advance our understanding of variant function and pathogenicity focused on germline susceptibility variants.
Objective 3: To perform novel genetic epidemiological analyses, harnessing data from 100KGP patients to advance our understanding of the cancer phenotypes and risk associated with variants and CSGs.
THEME III: To evaluate aspects of the ethical, social and economic impact of expansion of germline genetic testing through studies of testing of CPGs undertaken within the 100KGP

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)*

The Inherited Cancer Predisposition Domain (InCaP) will undertake analyses of the whole genome data from families and individuals diagnosed with cancer focusing on looking at changes (mutations) in genes that can be passed down through families and cause people to be at elevated risk of cancer. BRCA1 is a well-recognized cancer susceptibility gene (CSG), recently in the spot light through public discussions lead by Angelina Jolie. Studying these types of genes within the 100,000 Genomes Project will allow us to:
(1) identify the DNA mutations in individual participants which have been passed down through the family which are linked to them having developed cancer. These findings will be shared with the clinicians looking after those patients as they may be important in looking after that patient and/or their family.
(2) look across multiple patients and families to allow us to identify new cancer susceptibility genes and better understand the cancer risks associated with known cancer susceptibility genes.
As more gene and genome sequencing is undertaken in healthcare, including that genetic sequencing performed in cancer patients as part of their routine oncology care, it is essential that accurate information is provided about how genetic changes relate to cancer susceptibility. Knowledge about these genes can influence management of a patient with cancer, but also provide risk regarding other future cancers, as well as risk information for the family. Knowing of inherited risk of cancer may enable unaffected individuals to undertake enhanced screening, risk-reducing surgery or access to drugs which reduce their risk.
The InCaP domain comprises many clinicians and scientists with different but complementary expertise, for example clinical geneticists, laboratory scientists, functional biologists, epidemiologists, computational biologists, statisticians, bioinformaticians. We will work collaboratively to develop knowledge and resources around interpretation of genomic data relating to cancer predisposition. We aim to share these resources with other clinicians and scientists working within the 100,000 Genomes Project and the broader clinical community to enhance patient benefit and clinical research.

Technical summary. *Information from this summary may be displayed on a public facing website. Please include plans for methodology, including experimental design and expected outputs of the research. (max 500 words)*

Clinical testing for cancer susceptibility genes (CSGs) is having an increasing impact on the management of cancer patients, their families and unaffected individuals. However, much of the inherited susceptibility to cancer remains unexplained and the risks and phenotypes associated with established CSGs are yet to be fully delineated. Furthermore, variant interpretation of CSGs is complex and robust resources are lacking; this deficit is problematic in clinical diagnostics and is hampering oncological research. The research activities of the Inherited Cancer Predisposition (InCaP) domain will focus broadly on identification and characterisation of cancer susceptibility genes. This activity will comprise both direct analyses of the whole genomes and clinical data generated from the 100,000 Genomes Project but also broader integrative, multidisciplinary approaches that will advance our understanding of variant interpretation, risk, phenotypic correlates and clinical implementation.

Whole Genome analysis of germline genomes for discovery and description

Patients and their families are being recruited to the 100,000 Genomes Project under >10 phenotypic categories relating to syndromes of cancer susceptibility. For each phenotypic category we shall analyse genomes with the intention of identifying and characterising known and novel susceptibility genes via complex genome-wide analyses of all genes and intergenic regions. In the analysis of each phenotypic group, we shall harness the familial structures and full phenotypic data, analyse iteratively, modelling for variable segregation/penetrance, a breadth of variant types and harnessing functional prediction and pathway-based analyses to prioritise genes and variants for analyses. Any putative novel variants or susceptibility genes identified via the whole genome analyses will be evaluated via relevant in silico and laboratory functional evaluations. Working in collaboration with the leads of the Cancer GeCIPs, we shall undertake analyses of the germline genomes of the patients enrolled in the GeL cancer programs. These analyses will include: (i) evaluation of the contribution of known cancer susceptibility genes and variants to better characterize attributable and absolute risk; (ii) identification of novel susceptibility factors; (iii) application of rich clinical therapeutic and follow-up data captured against the cancer patients to evaluate germline variants (known and novel) for prediction and prognosis; (iv) joint analyses integrating the somatic and germline genomes to correlate the germline characteristics of subgroups defined by somatic markers and vice-versa.

Translational multidisciplinary characterisation genes and variants involved in cancer susceptibility

- We aim to develop a definitive resource for clinical interpretation of germline cancer susceptibility variants, including including explicit methodology for curation, integration of evidence and standardization of classification. Such a resource will require substantial development, including: (i) comprehensive expert clinical curation of extant diverse resources about thousands of variants in CGSs; (ii) establishing an integrated framework for rules for variant classification for CSGs; (iii) development of a comprehensive integrated hierarchical database and web-resource; (iv) engagement with and contribution to established international endeavours in variant classification e.g. InSiGHT BIC, IARC, ENIGMA, Human Variome Society, GA4GH, BRCA-challenge.
- We aim to perform novel integrated multi-disciplinary analyses (clinical, genetic epidemiological, computational and laboratory functional) to advance our understanding of variant function and pathogenicity. These approaches will reflect expertise in the group and include: (i) application and validation of in silico prediction tools for known CSGs; (ii) development of 3D protein modelling approaches; (iii) a systems medicine analysis of variants in CSGs; (iv) functional analyses including development of functional assays to evaluate pathogenicity of key variants in their canonical pathways and genetic and functional interaction studies with other pathways for discovery-based therapeutic intervention; (v) splicing analyses.
- We aim to perform novel genetic epidemiological analyses, harnessing data from 100KGP patients to advance our understanding of the cancer phenotypes and risk associated with variants in CSGs including: (i) creation of virtual registers of mutation carriers and carriers of high-suspicion variants; (ii) prospective follow-up and data linkage of families to outcome data; (iii) coordination with international endeavours for rarely mutated genes.

Evaluation of aspects of the ethical, social and economic impact of expansion of germline genetic testing through studies of testing of CSGs undertaken within the 100KGP

This include activities around development of practice around communication of secondary GPG findings to 100KGP recruits as well as cost-benefit and cost-effectiveness of clinical interpretation of CSG variants.

Expected start date	**1/1/2016**
Expected end date	**31/12/2018**

Lead Applicant(s)	
Name	Dr Clare Turnbull
Post	Senior Lecturer in Genomic Medicine
Department	(i) William Harvey Research Institute (ii) Division of Genetics and Epidemiology
Institution	(i) Queen Mary University London (ii) Institute of Cancer Research
Current commercial links	nil

Administrative Support	
Name	nil
Email	
Telephone	

Subdomain leads		
Name	**Workstream**	**Institution**
(1) Marc Tischkowitz (2) Katie Snape	Clinical phenotyping, eligibility and recruitment	(1) University of Cambridge (2) St Georges University Hospital London
Yvonne Wallis	Clinical Interpretation	Birmingham Regional Genetics Laboratory
(1) Richard Houlston (2) Ian Tomlinson	Large-scale WGS analysis	(1) Institute of Cancer Research (2) University of Oxford
Michael Sternburg	Computational Prediction	Imperial College, London
Anthony Carr	Functional analyses	***University of Sussex:*** Genome Damage and Stability Centre (GDSC)
Antonis Antoniou	Risk and familial analyses	University of Cambridge
(1) Ingrid Slade (2) Anneke Lucassen	ELSE	(1)University of Oxford (2) University of Southampton
Kate Tatton Brown, Ian Frayling	Education and Training	

Detailed research plan

Full proposal (total max 1500 words per subdomain)	
Title *(max 150 characters)*	**Integrative studies of Inherited Cancer Susceptibility: Whole Genome Analyses and translational multidisciplinary characterisation of genes and variants**

Importance. ***Explain the need for research in this area, and the rationale for the research planned. Give sufficient details of other past and current research to show that the aims are scientifically justified. Please refer to the 100,000 Genomes Project acceptable use(s) that apply to the proposal (page 6).***

Genetic architecture of Cancer Susceptibility

Cancer susceptibility genes (CSGs) have been highly informative in clinical practice as they enable us to define the genetic cause of segregation of cancer within families. This enables management of individuals with cancer to be informed on the basis of the molecular aetiology of their cancer, whilst unaffected family member can have predicative testing. This enables unaffected mutation carriers to be offered enhanced screening, risk-reducing surgery or prophylactic drugs, whilst those not carrying the mutation can be reassured. While >100 genes and >400 genomic variants conferring an increased risk of cancer are recognised, for most cancers, the majority of the inherited predisposition remains unexplained. While many explanations for the missing heritability can be advanced none is likely to be all encompassing. Whole genome analysis, which has thus far only been applied to comparatively small sample numbers in familial or syndromic cancer phenotypes, offers the prospect of comprehensively cataloguing all sequence changes affecting cancer risk. Harnessing large multiplex families, constellations of highly specific pleomorphic syndromic presentations, and the power of whole genome sequencing, it is timely to evaluate the power of whole genome analysis for gene discovery and characterization in cancer susceptibility.

Interpretation and classification of genes and variants conferring susceptibility to Cancer

Clinical testing of CSGs in genetics clinics has been undertaken to varying degrees in the UK for over 20 years. Myriad data have accumulated in the literature and databases pertaining to variant pathogenicity in CSGs, much of which is arcane, erroneous and/or unvalidated for clinical application. Moreover, there are no established or consistent standards by which to define pathogenicity for variants in CSGs. Accordingly, many of the rare variants detected through clinical testing are currently reported as 'variants of uncertain significance', a broad category which creates ambiguity and inconsistency of management. This is especially an issue problematic for particular CSGs which are highly variant such that rare missense variants are commonly seen in the healthy population.

Clinical Interpretation of genomic data is acknowledged to be one of the most pressing and challenging aspect of the genomic revolution. This is especially relevant in cancer susceptibility genes for the following reasons:

- These are typically autosomal dominant genes of intermediate penetrance conferring susceptibility to phenotypes which are indistinguishable from the non-genetic forms. Accordingly, the mechanisms available for ascribing pathogenicity in rare distinctive fully-penetrant monogenic childhood disorders are not available. More sophisticated genetic epidemiologic and functional analyses are required to determine pathogenicity.
- On account of the predictive and prognostic information afforded by the genomic analyses of tumours, and the requirement of the constitutional genome for subtraction for these analyses, there will over the next decade be a tsunami of germline genomes from cancer patients requiring interpretation. Clearly, the most pertinent genes for examination within these data will be the cancer susceptibility genes.
- Oncological agents with activity predicated on germline status, such as PARP inhibitors, are increasingly being used clinically. Relating to this scenario alone, there is a rapidly increasing requirement for robust germline interpretation of relevant germline DNA repair genes to determine drug eligibility and predict treatment response.

A multitude of data sources pertaining to variant classification currently exist. These include rapidly increasing volumes of sequencing and genotyping data, numerous in silico prediction tools, curated resources such as the human gene mutation database (HGMD) and ClinVar and decades of literature containing valuable genetic epidemiologic analyses and a variety of functional assays. These data resources are of variable quality with differing levels of curation, disparately located, complicated in presentation and access, with often conflicting assignations of pathogenicity for any given variant. Accordingly, clinical interpretation of variant pathogenicity can be time-consuming and often leads to inconsistent conclusions. Pan-genome bioinformatic tools are emerging which seek to automate collation of relevant data sources in order to inform classification of variants detected on sequencing. Whilst having utility for variant prioritisation for gene discovery experiments, such tools lack both the robustness and the specificity required for clinical variant classification and interpretation. Clinical variant classification is context-

dependent and the relevance of different types of data is necessarily driven by gene-specific contexts. These contexts relate to **biological parameters**, such as mechanism of action and location of functional domains, **epidemiological parameters**, such as the rarity of the phenotype, attribution of the phenotype to the disease and penetrance for disease of pathogenic mutations and **genomic parameters** such as gene size, mutability and mode of inheritance. Furthermore, **clinical parameters** are also important: what are the clinical implications of misaligning a variant as pathogenic versus failing to ascribe pathogenicity in a given clinical scenario. Interpretation of these data are highly complex and requires expert knowledge of clinical phenotype, of the availability, location and relevance of sources of variant-level data and of how these should be assimilated in decision-making around pathogenicity.

To support clinical implementation of broader germline genetic testing for cancer susceptibility and enable and develop research activities in cancer, robust, accessible resources for interpretation of cancer susceptibility genes are urgently required.

Risk and phenotype association

Even for well characterised genes such as *BRCA1* and *BRCA2*, studies of risk have largely been derived from retrospective analyses of high-risk families collected through biased phenotype-driven ascertainment. Furthermore, risk estimates are predicated on presumption of equivalent risk for all pathogenic variants. Large-scale unbiased prospective studies are required to better estimate the cancer risk and phenotypic spectrum in all contexts and establish gene- and mutation-specific risks.

Public Health, health economics and ethics

Delivery of constitutional genetic testing for cancer predisposition has to date been small-scale, confined to the realm of clinical genetics and focused on the counselling of families with multiple affected individuals. Already, it is becoming clear that finding hereditary cases amongst common tumours, such as colorectal and breast cancer, may be better achieved by systematic testing of incident cases, rather than relying on family histories. Genomic sequencing of cancer patients is therefore set to become routine and the 100KGP marks the advent of delivery of whole genome sequencing as part of routine medical and public health.

Importance of Cancer Susceptibility to 100KGP

More so than any other rare disease phenotypic theme, inherited predisposition to cancer is of central relevance to the 100KGP for the following reasons:

- Firstly, Inherited Cancer is a category within the Rare Diseases programme which will involve sequencing of individuals and families with unexplained familial and/or syndromic presentations of cancer.
- Secondly, within the Cancer programme, a proportion of the patients analysed will have developed their tumour due to inherited germline susceptibility. A broad number of susceptibility genes require analysis for each patient in the Cancer programme to identify these germline susceptibility mutations: return of these 'pertinent' findings is central to the cancer reporting.
- Thirdly, reporting of mutations in 13 high-penetrance CSGs as secondary (i.e. looked-for) findings will be offered to all adult participants within 100KGP.

Accordingly, robust analysis and interpretation of variation in known and novel CSGs is a prerequisite for delivery of the entire 100KGP as well as providing rich opportunities for discovery and translational research studies.

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.*

We have divided the domain into eight work-streams which will interact in a complimentary fashion to address the three overarching themes, as detailed below. Hence, the InCaP domain will not function as subdomains, and in view of this we are submitting a unified proposal:

Workstreams

(1) Clinical phenotyping, eligibility and recruitment (Lead: Snape, Tischkowitz)
(2) Variant Interpretation (Lead: Wallis)
(3) Whole Genome analysis (Leads: Houlston, Tomlinson)
(4) Computational Prediction (Lead: Sternberg)
(5) Functional analyses (Lead: Carr)
(6) Risk and familial analyses (Lead: Antoniou)
(7) ELSE (ethical, legal, social, economic) (Leads: Slade, Lucassen)
(8) Education and Training (Leads: Tatton Brown,Frayling)

Research Themes, Objectives and Plans

THEME I) Whole Genome analysis of germline genomes for discovery and description

The whole genome analyses will span two datasets, described here as separate objectives

OBJECTIVE 1: Comprehensive Whole Genome Analysis of patients with Inherited Cancer Phenotypes recruited to the rare disease program

We have delineated 10 phenotypic eligibility groups in Inherited Cancers, for which there is potential unmet diagnostic need and opportunity for discovery:

- Familial breast cancer
- Familial colorectal cancer
- Multiple bowel polyps
- Familial rhabdomyosarcoma or sarcoma
- Genodermatoses with malignancy: including mutation-negative Gorlin, Cowden and Muir Torre syndromes
- Multiple endocrine tumours: including mutation-negative MEN1- and MEN2 spectrum
- Neuro-endocrine tumours- phaeochromocytoma and paraganglioma
- Parathyroid Cancer
- Multiple primary tumours
- Paediatric congenital malformation-dysmorphism-neoplasia syndromes
- Others may be added

For each category we shall analyse genomes with the intention of identifying and characterising predisposition variants in:

(i) Cancer susceptibility genes already implicated in that phenotype (as per gene panels which we have supplied)
(ii) Other cancer susceptibility genes not conventionally associated with the phenotype
(iii) Novel susceptibility genes and intergenic regions

In each phenotypic analysis, we shall harness the familial structures and analyse iteratively, modelling for incomplete segregation and incomplete penetrance. We shall analyze for different variant types and use functional prediction and pathway-based analyses to prioritize genes and variants for analyses. Any putative novel variants or susceptibility genes identified via the whole genome analyses will be evaluated via functional studies. This may include multi-omic analyses, harnessing the additional specimens collected.

OBJECTIVE 2: Comprehensive Whole Genome Analysis of germline genomes from patients recruited to the

Cancer program

The InCaP domain is currently providing the entirety of the support to the Genomics England Validation and Feedback domain (Drs Baple, Wright) in gene curation and variant curation of the genes for germline reporting for pertinent findings for the Cancer Programme: commensurate with this, we anticipate that the InCap domain will undertake analyses of the germline genomes of the patients enrolled in the cancer program, working in collaboration with the leads and relevant members of the Cancer GeCIPs (multiple intercalating memberships of the Inherited Cancer domain and the respective Cancer Domains will service this cross-talk). These analyses will include:

- Evaluation of the contribution of known cancer susceptibility genes and variants to the cancer
- Analysis for novel susceptibility factors. These analyses will comprise full analysis of a multitude of inheritance mechanisms and variant types, and again, we shall deploy functional prediction and pathway-based analyses to prioritize genes and variants for analyses
- Application of rich clinical treatment and follow-up data captured against the cancer patients to evaluate germline variants (known and novel) not just for association with disease but for prediction and prognosis.
- Collaborating with those analyzing the tumour genomes, we shall perform joint analyses integrating the somatic and germline genomes to correlate the germline characteristics of subgroups defined by somatic markers and vice-versa.

Timelines for analysis: Analyses of accrued familial will be undertaken as soon as there is access to the data embassy. Thereafter, analyses will be repeated at three month intervals or sooner (on addition of additional or noteworthy families). Analysis of germline cancer genomes will be undertaken every three months. **Challenges:** As described above, much of the genetic susceptibility to common cancers may reside in variants of low or modest effect size and familial clusters may be due to shared common variation, shared environment or chance. As phenotypes are often non-specific and typically transmission of risk alleles will have incomplete penetrance via an autososmal dominant inheritance, analyses may require considerable power to identify the causative variants.

THEME II) Translational multidisciplinary characterisation of genes and variants involved in cancer susceptibility:

Broadly we have three objectives for translational multidisciplinary characterisation of genes and variants involved in cancer susceptibility::

1. **To develop a definitive resource for interpretation of germline susceptibility variants for clinical decision-making, including explicit methodology for curation, integration of evidence and standardization of classification.**
2. **To perform novel multi-disciplinary analyses (clinical, genetic epidemiological, computational and laboratory functional) to advance our understanding of variant function and pathogenicity focused on germline susceptibility variants.**
3. **To perform novel genetic epidemiological analyses, harnessing data from 100KGP patients to advance our understanding of the cancer phenotypes and risk associated with variants and CSGs**

OBJECTIVE 1: To develop a definitive resource for interpretation of germline susceptibility variants for clinical decision-making, including explicit methodology for curation, integration of evidence and standardization of classification.

We aim to address the current gross deficit in resources available for interpretation and clinical classification of germline variants in cancer susceptibility genes. Pending appropriate support, we aim to use the structures of the domain and the ongoing interpretation activity already underway for the 100KGP program, to develop a definitive resource to address this deficit. ***Turnbull and Sultana have already developed a prototype of such a resource, the CaVaDa system: Cancer Predisposition Gene Variant Database (http://cavada.dynalias.org/cavada/) to which many InCaP members have contributed. This is in widespread use in clinical genetics units and oncological research groups and already holds extensive data and clinical classifications on >1.2 million variants in CSGs.***

Advancement of this activity will harness the breadth of the domain and require interactions between domain clinicians, genetic epidemiologists, functional biologists, structural computing experts and other related disciplines, as well as expert bioinformaticians to develop the database structures and integrate resources. In

brief, the activity towards development of this resource would be structured as follows:

1. **Comprehensive expert clinical curation of extant diverse resources pertaining to variant classification and clinical variant interpretation of CSGs:**
    a. Comprehensive, systematic literature review to identify the key gene-specific inputs that will
    b. inform variant classification for each CSG
    c. Rigorous review of locus-specific databases for each CSG
2. **Development of integrated framework for rules for variant classification for CSGs**
    a. Development of global generic rules for classification of variant pathogenicity in CSGs
    b. Development of gene-specific classification decision-trees for each CSG.
3. **Development of a comprehensive integrated resource for clinical use for clinical variant classification of CSGs (as per CaVaDa prototype)**
    a. Generation of every possible variant within each CSG in a single data-system
    b. Integration of multiple sources of broad variant level data (incl 100KGP allele frequencies) and gene-specific variant-level data (from Activity 1) in this data-system
    c. Generation and incorporation into data-system of clinical classifications for each variant (from Activity 2b)
    d. Framework by which data-system can be continually annotated and updated by gene-experts.
    e. Data-system to be made publically and maintained available via web-application.
4. **Engagement with and contribution to established international endeavours in variant classification** *e.g.* the InSiGHT Variant Interpretation Committee, BIC, IARC, ENIGMA, Human Variome Society, GA4GH, BRCA-challenge (Frayling, Eccles, Tischkowitz Baralle, Turnbull already members of these groups). This will thus enhance clinical interpretation of 100KGP data, and in turn ensure that 100KGP data contributes to international endeavours in the interpretation of variants.

OBJECTIVE 2: To perform novel multi-disciplinary analyses (clinical, genetic epidemiological, computational and laboratory functional) to advance our understanding of variant function and pathogenicity focused on germline susceptibility variants.

The core resource development, data curation and integration will be enhanced by work by the Computational Prediction workstream (Lead: Sternberg) and the Functional analyses work stream (Lead: Carr), with the aim of using novel computational and functional analyses to improve our understanding of variant pathogenicity and risk:

1. **Application and validation of in silico prediction tools for known CSGs**
    a. Development and extension of state of the art bioinformatic *in silico* prediction tools (such as SuSPECT (Sternberg Laboratory)) with algorithms specifically designed for analysis/prediction of different mutational types, i.e. truncating mutations, missense mutations, mutations influencing splicing
    b. Cross-evaluation using multiple methodologies such as ROC for integrative prediction of different in silico tools.
2. **Development of 3D protein modelling**
    a. Evaluation and prediction of effects of mutations in CSGs on experimental and predicted tertiary and quarternary structures
    b. Evaluation and prediction of effects of mutations in CSGs on structure of key domains

    The structural protein analyses will be coordinated with and undertaken in broader cross-cutting domains and that related to core work of the Sternburg group.
3. **A systems medicine analysis**
    a. Computational analysis of the effects of mutations in CSGs in terms of pathways and the interactome
    b. Reprioritisation of GWAS results based on systems analysis
4. **Functional analyses**
    a. Development of functional assays to evaluate pathogenicity of key variants in their canonical pathways
    b. Genetic and functional interaction studies with other pathways for discovery-based therapeutic

intervention

The functional analyses work will be coordinated with and undertaken in broader cross-cutting domains and that related to core work of the Genome Damage and Stability Centre (GDSC) will be undertaken by Carr, Pearl, O'Driscoll, Downs

5. **Splicing analyses**
    a. Development of RNA and minigene assays for key variants in CSGs
    b. Correlation of in vitro data with in silico-predictions to train and improve in silico prediction of splice disruption

This work in splicing analyses will likely be incorporated and complimentary to work undertaken in a broader cross-cutting domain focused on splicing (led by Barralle)

OBJECTIVE 3): To perform novel genetic epidemiological analyses, harnessing data from 100KGP patients to advance our understanding of the cancer phenotypes and risk associated with variants and CSGs

The proposed work around risk will be led by Antoniou, Pharoah et al comprises the following elements:

1. Creation of virtual registers of mutation carriers and carriers of high-suspicion variants.
    i. Prospective follow-up and data linkage of families to outcome data
    ii. Coordination with international endeavours for rarely mutated genes (e.g. the Consortium of Investigators of Modifiers of BRCA1/2 – CIMBA (led by Antoniou), and the ENIGMA endeavours in risk estimation (Antoniou, Tischkowitz, Turnbull involved))
2. Risk estimation analyses by gene and by genotype in different contexts (conditioned for ascertainment).
3. Integration of information on common genetic variants identified as modifiers of risk through GWAS in risk prediction.
4. Integration of information from 8.a-c into comprehensive cancer risk prediction algorithms for predicting future cancer risks in unaffected individuals in all contexts (e.g. BOADICEA)

THEME III: To evaluate aspects of the ethical, social and economic impact of expansion of germline genetic testing through studies of testing of CSGs undertaken within the 100KGP (Luccassen, Slade et al)

1. **Ethics**

Proposed activities around development of practice around communication of secondary GPG findings to 100kGP recruits, in particular to parents of recruit with rare disease e.g. the effects of learning about inherited predisposition when the reason for taking part was diagnosis of a child; diagnosis of adult onset predispositions in minors and the timing of their communication; management of rare incidental findings; record keeping and NHS flagging to ensure future appropriate surveillance.

2. **Health economics**

The cost-benefit and cost-effectiveness of clinical interpretation of CSG variants will be explored in the setting of genomic [and/or systematic genetic] testing.

Activities will likely be coordinated with ELSE evaluations undertaken within the ELSE domain (Lucassen, Wordsworth and Slade also involved in this GeCIP)

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.*

We shall collaborate with the Cancer GeCIPs to undertake efficient, productive analyses of the cancer germline genomes, as delineated above. In each of the cancer GeCIPs as listed, the key member of the GeCIP with germline expertise is part of InCAP:

Breast: Turnbull, Tischkowitz , Antoniou
Colorectal: Tomlinson, Houlston
Ovarian: Tischkowitz, Pharoah
Lung: Houlston, Turnbull
Renal: Tomlinson, Turnbull, Houlston
Testicular cancer: Turnbull, Tomlinson
Brain Tumours: Houlston

Lucassen, Slade and Wordsworth are also central to the Ethics and Health Economics GeCIPs and will facilitate collaborations as indicated. Carr, Pearl, O'Driscoll, Downs, Baralle and Sterburg are part of cross-cutting/functional domains within which the analyses will be undertaken.
Additional key collaborations will evolve but include international consortia relating to activities that we have detailed in this proposal and include CIMBA (Consortium of Investigators of MOdifiers of BRCA1/2- coordinated by AA), IBCCS (International BRCA1/ Carrier Cohort Study - Coordinated at Netherlands Cancer Institute coordinated by MT), ENIGMA, InSIGHT, ClinVar/ClinGen (through collaboration with Sharon Plon, Heidi Rehm and others)

Training. *Describe the planned involvement of trainees in the research and any specific training that will form part of your plan.*

Tatton Brown and Frayling will be joint training directors and Ingrid Slade will be the trainee representative. We recognise rich opportunities for training within our proposed InCaP domain. It is anticipated that the many trainees will engage in the research we are proposing, including the following:

- Ingrid Slade: Clinical Trainee in Public Health 2012-2017: Clinical Genetics, Ethics, Health Economics. Analyses of patient-centred responses to engagement and return of genomic findings
- Kevin Litchfield: PhD student 2013-2017 (Turnbull/Houlston): Mathematics, bioinformatics, modelling. Analysis of WGS for description/discovery
- Eman Alhuzimi: PhD student 2014-2017 (Sternberg): Development of resource for a systems approach to variant classification.
- Sirawit Ittisoponpisan: PhD student 2014-2017 (Sternberg): Prediction of 3D protein complexes and genetic variants
- Matthew Scales: PhD student 2014-2018 (Houlston/Sternberg): Gene prioritisation from in silico prediction of protein structure

In addition, we shall seek PhD fellowships and Clinical Research fellowships to further engage trainees in learning from and delivering on the proposed body of research and clinical translational activities. Of additional note:

- Snape and Tatton Brown (St. Georges Medical School) are have developed a MOOC in genomic medicine for HEE, lead a Genomic Medicine MSc, lead a BSc module in cancer Genetics and are delivering a national PGCert for Clinical Genetics Trainees. The InCaP domain will likely make invaluable contribution to teaching on and enriching these programs.
- The Genome Damage and Stability Centre runs an MSc in Cancer Cell Biology (University of Sussex: MO'D convenor, class size 2014-15; 19 students) incorporating all aspects of cancer cell biology and will offer projects in functional modelling of cancer-specific variants using multiple platforms (Yeast, Chicken, Mammalian) using the latest gene manipulation technologies (flp-in, Cas9/Crispr, Lentivirus). Students will therefore have the opportunity to pursue research projects related to and inspired by the InCaP domain.
- Sternberg is the Deputy Director of the MSc in Bioinformatics and Theoretical Systems Biology at Imperial College (class size about 15). Research projects related to this related to GeCIP will be offered for example development or enhancement of databases; development and application of novel algorithm for gene prioritisation.

People and track record. *Explain why the group is well qualified to do this research, how the investigators would work together.*

The constituent members listed within this proposal (see Appendix 1 for full details of members) represent a breadth of clinical and academic expertise related to the proposed activities, as described below. The members of the InCaP domain already work closely together clinically within and across their respective GMCs, via research collaborations, via consortia and through regional and national development activities within our speciality, such as those lead by CGG (Cancer Genetics Group), Pan Thames Cancer Genetics Group, and BSGM (British Society of Genetic Medicine). The InCaP domain members are strongly representative across:

A) <u>Clinical and Research areas</u>

- **Discovery analysis for CPGs from WES/WGS data:** IT, RSH, EM, CAT, MT, JS, KS, SE, KTB, PP, AA

- **Integration of somatic and constitutional analyses of WGS/WES of cancer:** IT, RSH, EM, CAT, PP
- **Penetrance analysis for CPGs:** AA, PP, MT, RSH, CAT
- **Survival analysis in genetic cancer predisposition:** PP, AA, RSH, DE
- **Development of cancer risk prediction/survival models and related online tools:** AA, PP
- **Development of relational databases:** MS, RS
- **Development and maintenance of online resources:** CAT, MS, RS
- **Clinical interpretation of mutational data from patients with familial cancer syndromes:** LI, HH, MT, RSH, IT, EM, DE, DBa, PP, AL, AA, IMF, EW, CAT, SE, KS, RR, IW, YW, IB
- **Variant classification systems:** DE, DBa, IMF, MT, CAT, HH, SE, RR, IW, YW, IB
- **International collaborative projects to advance variant interpretation:** DBa, DE, IMF, MT, CAT, WF, HH, IB
- **In vitro analysis of splicing:** DBa
- **In silico variant effect prediction:** MS, RSH, CAT
- **3D protein modelling to predict variant effect:** MS, LP
- **Systems medicine analysis:** MS
- **Functional analyses of genome stability and relevant cell cycle, chromatin and signalling:** AC,LP,MO'D, JD
- **Research studies collecting and analysing genetic data from families with cancer:** IT, RSH, EM, DE, IMF, MT, CAT, PP, AA, WF
- **Economic analyses of models of evaluation of cancer genetic predisposition:** SW, IMF, IS
- **Ethical and psychosocial analysis of impact of testing for cancer predisposition:** AL, IS, KK
- **Clinical management of patients with familial cancer syndromes:** LI, HH, MT, RSH, IT, EM, DE, DBa, AL, KK, WF, EW, CAT, MA
- **Genetic predisposition to Breast-Ovarian Cancer:** HH, MT, DE, FL, DBa, CAT, LI, KS, PP, AL, AA, WF, MA
- **Genetic predisposition to Colorectal Cancer:** RSH, IT, HT, JS, IMF, LI, AL, HH, CAT, MA, FL, IB
- **Genetic predisposition to Endocrine cancers:** FL, EM, LI, EW, KS, DE
- **Genetic predisposition to Renal cancer:** EM, EW
- **Genetic predisposition to Gastric cancer:** MT, PP
- **Genetic predisposition to Endometrial Cancer:** IT, IMF
- **Genetic predisposition to Haematological malignancies:** JF, RSH
- **Genetic predisposition to Childhood cancers and syndromes:** HH, DE, KTB, MA, WF
- **Genetic predisposition to Skin Cancers:** NR, KRO
- **Genetic predisposition to pleomorphic cancer syndromes:** JS, KTB, WF, HH, CAT
- **Clinical/oncological management of patients with CPGs**: DE, HT, CAT

B) GENOMIC MEDICINE CENTRES

The following are clinicians active in the Clinical Genetics Units/Regional Diagnostic Laboratories of the respective GMCs:

- **East of England NHS GMC:** EM, MT, JBa, RH
- **South London NHS GMC:** LI, CAT, KS, KTB, HH
- **North West Coast NHS GMC:** LG
- **Greater Manchester NHS GMC**: FL, DGE
- **North Thames NHS GMC:** LS, AB
- **North East and North Cumbria NHS GMC**: NR, JB
- **Oxford NHS GMC:** IT
- **South West Peninsula NHS GMC:** SE, CB
- **Wessex NHS GMC:** AL, DE, DBa, MA
- **Imperial College Health Partners NHS GMC:** RSH, KK
- **West Midlands NHS GMC:** EW, KRO, YW
- **Yorkshire and Humbs:** IB, RR

C) RELEVANT NHS CLINICAL SPECIALITIES

- **Clinical Laboratory Sciences:** SE, IMF, RR, IB, YW

- **Clinical Genetics:** DE, DBa, MA, EW, KRO, IT, EM, MT, JS, CAT, RSH, LI, KTB, KS, FL, HH
- **Clinical Molecular Pathology:** IMF
- **Genetic Counselling:** KK
- **Related clinical disciplines:** HT (gastroenterology), NR (dermatology)
- **Public Health:** PP, IS

D) KEY STAKEHOLDER GROUPS

LS is Chair of **Cancer Genetics Group** steering committee. DE, DBa, MA, EW, KRO, IT, EM, MT, JS, CAT, RSH, LI, SE, KK, IMF, JB, DGE, KS, AB, IB are members of CGG. CGG is the specialist subgroup of British Society of Genetic Medicine focused on cancer susceptibility.

Clinical interpretation. *(Where relevant to your GeCIP) Describe your plans to ensure patient benefit through clinical interpretation relevant to your domain. This should specifically address variant interpretation and feedback and your interaction with the cross-cutting Validation and Feedback domain.*

As detailed above, several members of the InCaP domain are already working closely with Dr Emma Baple and Dr Caroline Wright from Genomics England and actively contributing to the gene and variant curation required for the gene panels and in particular the genes for which secondary and pertinent findings will be reported for the full programme and cancer programme respectively. Dr Yvonne Wallis, from Birmingham Regional Genetics Laboratory, is leading the Clinical Interpretation Workstream and has engaged clinical scientists and clinical geneticists in this activity.

Beneficiaries. *How will the research benefit patients and healthcare institutions including the NHS, other researchers in the field? Are there other likely beneficiaries?*

Identification of novel cancer susceptibility genes will inevitably make a significant contribution to both our understanding of cancer aetiology and furthermore our ability to investigate and manage patients. Arguably, the activity around development of robust resources for variant interpretation, characterisation and risk proposed by the domain will be of greater and more widespread immediate value as these resources will support clinical diagnostics, management of germline findings in the oncology setting and support broader endeavours in somatic cancer research for which germline input is currently lacking.

Commercial exploitation. *(Where relevant to your GeCIP) Genomics England has a very explicit intellectual property policy. We and other funders need to know if the proposed research likely to generate commercially exploitable results. Do you have commercial partners in place?*

There are no commercial partners currently in place.

References. *Provide key references related to the research you set out.*

Data requirements

Data scope. *Describe the groups of participants on whom you require data and the form in which you plan to analyse the data (e.g. phenotype data, filtered variant lists, VCF, BAM). Where participants fall outside the disorders within your GeCIP domain, please confirm whether you have agreement from the relevant GeCIP domain. (max 200 words)*

Analyses will utilise

- complete phenotype data, full filtered variant lists and VCF files for all individuals recruited under the inherited cancer eligibility criteria. Additionally unfiltered bam files for visualisation of variants and specialised analyses (e.g. for germline CNVs).
- germline (only) data from all patients recruited under the cancer programme. The InCaP domain are already supporting clinical interpretation of the germline data from the cancer program through curation of the 'pertinent findings' gene/variant lists. The InCaP domain will require access to the variant data from these patients in order to offer additional expert clinical interpretation. Members of the InCaP group are also participants in the relevant specific cancer GeCIPs in order to coordinate and optimise analyses.

Should funding become available for additional analyses on the same or similar patients, we shall wish to integrate those data with the genome sequence and it will be necessary to update the analysis, software, data import and computing plans accordingly.

Data analysis plans. *Describe the approaches you will use for analysis. (max 300 words)*

We shall undertake analyses:

- Across family samples using cohorts of similar and overlapping patients
- Under different models of inheritance
- Of linkage and homozygosity mapping within families
- By different tranches of predicted variant severity (T1 burden test, T2 burden test) in order to prioritise disease-causing sequence changes
- Via successive filtering steps and functional prioritisation of variants
- Integrating across different mutational types to maximise power
- Via single variant analysis and haplotype association analysis
- Via pathway-based analyses using curated gene sets catalogued by the Broad gene set enrichment analysis database which is compiled from multiple pathway resources including KEGG, BIOCARTA and Gene Otology biological processes

Key phenotype data. *Describe the key classes of phenotype data required for your proposed analyses to allow prioritisation and optimisation of collection of these. (max 200 words)*

Key phenotype data includes:

Key phenotype data will vary according to the phenotype for which participants are recruited, but should include:

- Family structure
- Detailed family history of cancer and benign tumours, including tumour type(s), histology, age(s) of diagnosis, living/deceased. This should include information on all 1/2/3 degree members, not just those recruited to the programme
- Additional non-cancer phenotype data e.g. congenital abnormalities, skin lesions, bony lesions, mental retardation and any other unusual and/or early-onset phenotype of note.

Alignment and calling requirements. *Please refer to the attached file (Bioinformatics for 100,000 genomes.pptx) for the existing Genomics England analysis pipeline and indicate whether your requirements differ providing explanation. (max 300 words)*

We plan to use the Genomics England analysis pipeline for these samples. Of note we value the contribution to germline variation made by CNVs and should like to ensure that CNV calls are made available and are of high quality.

Tool requirements and import. *Describe any specific tools you require within the data centre with particular emphasis on those which are additional to those we will provide (see attached excel file List_of_Embassy_apps.xlsx of the planned standard tools). If these are new tools you must discuss these with us. (max 200 words)*

We anticipate primarily using bio-informatic tools listed within *List_of_Embassy_apps.xlsx* . However, we wish to retain the option to import other tools as required to meet the demands of data analysis: in particular we may wish to review CNV and breakpoint calling.

Data import. *Describe the data sets you would require within the analysis environment and may therefore need to be imported or accessible within the secure data environment. (max 200 words)*

We may wish to import BAM files and VCF files generated from WES or WGS of relevant patient cohorts to meta-analyse data. We may also wish to import genotype files and microarray files.

Computing resource requirements. *Describe any analyses that would place high demand on computing resources and specific storage or processing implications. (max 200 words)*

It is difficult to predict at this stage, particularly as recruitment numbers and the structure of the data centre are uncertain. However, we anticipate that 30-50 cores will be required in the first instance for analysis using .vcf file-based variant and phenotype data.
We anticipate analysis of the full set of germline exomes from the inherited cancer families (1000 families, 3000 individuals) will require ~30,000 CPU hours for the preliminary analysis and upto 100,000 CPU hours for the more advanced discovery analyses such as T1, T2 burden testing (as this requires intensive permutation).
We anticipate analysis of the full set of germline exomes from the cancer programme (~20,000 individuals) will require ~200,000 CPU hours for the preliminary analysis and up to 600,000 CPU hours for the more advanced analyses.
Depending on the quality of these data, additional resources may be required if additional analysis of raw data (.bam) files is indicated, for example for CNV analyses.

Omics samples

Analysis of omics samples. *Summarise any analyses that you are planning using omics samples taken as part of the Project. (max 300 words)*

This is currently difficult to predict and will be predicated on signals and potential discovery from the WGS data. Utilising omics samples will be driven by gene discoveries and/or suspicious variants. Potential exemplars include utilisation of PAXgene samples to undertake reverse transcription and sequencing of cDNA to evaluate the functional impact of putative splicing variants.

Data access and security	
GeCIP domain name	[from previous entry]
Project title *(max 150 characters)*	[from previous entry]

Applicable Acceptable Uses. *Tick all those relevant to the request and ensure that the justification for selecting each acceptable use is supported in the 'Importance' section (page 3).*

✓ *Clinical care*

☐ *Clinical trials feasibility*

☐ *Deeper phenotyping*

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

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

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

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

☐ *Non hypothesis driven R&D - health*

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

☐ *Other health use - clinical audit*

☐ *Public health purposes*

☐ *Subject access request*

☐ *Tool evaluation and improvement*

Information Governance

✓ The lead and sub-leads of this domain will read and signed the Information Governance Declaration form provided by Genomics England and will submit by e-mail signed copies to Genomics England alongside this research plan.

Any individual who wishes to access data under your embassy will be required to read and sign this for also. Access will only be granted to said individuals when a signed form has been processed and any other vetting processes detailed by Genomics England are completed.

Other attachments

Attach other documents in support of your application here including:

- a cover letter (optional)
- CV(s) from any new domain members which you have not already supplied (required)
- other supporting documents as relevant (optional)

Appendix 1: Domain members, affiliations and research/clinical expertise

	Academic institution	NHS affiliation/ contract	Research expertise	Clinical discipline/ expertise	Email
Domain Lead Clare Turnbull (CAT)	QMUL, London ICR, London	Hon. NHS contract GSTT and Royal Marsden	***Molecular and Statistical Genomics:*** Genetic susceptibility to cancer (breast cancer, ovarian cancer, testicular cancer, sarcoma, Wilms Tumour). Computational approaches to variant classification	***Clinical Genetics:*** Breast/Ovarian/colorectal Cancer predisposition	c.turnbull@qmul.ac.uk
Deputy Lead Ian Tomlinson (IT)	University of Oxford	Hon. NHS contract Radcliffe Hospitals NHS trust	***Molecular Genomics:*** Genetic susceptibility to colorectal cancer, bladder cancer, renal cancer, endometrial cancer, benign tumours including GI polyposis and Barrett's oesophagus. Pharmacogenomics. Copy number variation and cancer susceptibility	***Clinical Genetics:*** Colorectal Cancer predisposition	iant@well.ox.ac.uk
Deputy Lead Richard Houlston (RSH)	ICR, London	Hon. NHS contract Royal Marsden	***Molecular and Statistical Genomics:*** Genetic susceptibility to colorectal cancer, leukaemia, lung cancer, brain tumours Computational methodology in genetic epidemiology, statistical genetics	***Clinical Genetics:*** Colorectal Cancer predisposition	richard.houlston@icr.ac.uk
Eamonn Maher (EM)	University of Cambridge	Addenbrooks NHS Trust	***Molecular Genomics:*** Genetic susceptibility to Renal Cancer, Paraganglioma/phaeochromocytoma, multiple tumours	***Clinical Genetics***	erm1000@medschl.cam.ac.uk
Julian Sampson (JS)	Cardiff	Cardiff and Vale University Health Board	***Molecular Genomics:*** Genetic susceptibility to cancer syndromes, colorectal cancer. Trials of therapeutics in cancer predisposition syndromes	***Clinical Genetics***	Sampson@cardiff.ac.uk
Diana Eccles (DE)	University of Southampton	University Hospital Southampton Foundation Trust	***Molecular and Clinical Genomics:*** Genetic susceptibility to breast cancer, molecular pathology of breast cancer in relation to germline mutation, childhood tumours, IARC, variant reporting guidelines, ENIGMA consortium, chemoprevention for cancer, Human Variome Project BRCA Challenge.	***Clinical Genetics***	d.m.eccles@soton.ac.uk
Marc Tischkowitz (MT)	University of Cambridge	Cambridge	***Molecular Genomics:*** Genetic susceptibility to gastric cancer, breast cancer, ENIGMA consortium, Cancer Genetics Group Steering Committee	***Clinical Genetics***	mdt33@medschl.cam.ac.uk

Huw Thomas (HT)	Imperial college, London	St Marks, London	***Molecular and Clinical Genomics:*** Genetic susceptibility to colorectal cancer, polyposis, Screening and management of syndromes of high risk genetic predisposition	***Gastroenterology***	huw.thomas@imperial.ac.uk
Sian Ellard (SE)	University of Exeter Medical School	Royal Devon & Exeter NHS Foundation Trust	***Molecular Genomics:*** Genetic susceptibility to diabetes and hyperinsulinism	***Clinical Scientist*** Molecular testing for genetic forms of diabetes and endocrine disorders	sian.ellard@nhs.net
Fiona Lalloo (FL)		Central Manchester University Hospitals, NHS Foundation Trust	***Clinical Genomics***: Genetic susceptibility to endocrine tumours, breast cancer, bowel cancer	***Clinical Genetics***	fiona.lalloo@cmft.nhs.uk
Katie Snape (KS)	St Georges Medical School London	St George's University Hospitals NHS Foundation Trust, London	***Molecular and Clinical Genomics:*** Genetic susceptibility to breast cancers, Cancer Genetics Group Steering Committee ***Medical Education***	***Clinical Genetics***	ksnape@sgul.ac.uk
Kate Tatton Brown (KTB)	St Georges Medical School London	St George's University Hospitals NHS Foundation Trust, London	***Molecular and Clinical Genomics:*** Genetic susceptibility to childhood cancers, overgrowth syndromes ***Medical Education***	***Clinical Genetics***	katrina.tattonbrown@stgeorges.nhs.uk
Diana Baralle (DBa)	University of Southampton	University Hospital Southampton Foundation Trust	**Molecular Genomics:** Splicing mechanisms, analyses, minigene assays. ENIGMA consortium	***Clinical Genetics***	d.baralle@soton.ac.uk
Ian Frayling (IMF)	Cardiff University	Cardiff and Vale University Health Board	***Molecular Pathology:*** Colorectal cancer, polyposis and Lynch Syndrome molecular genetic pathology. Member of InSiGHT Council and Variant Interpretation Committee (VIC). Adjunctive testing of tumours in the clinical context. ***Health economics*** of genetic testing and disease screening. Cancer Genetics Group Steering Committee	***Genetic Pathology***	fraylingIM@cardiff.ac.uk

Louise Izatt (LI)		Guy's and St Thomas' London	***Clinical Genetics***: Genetic susceptibility to endocrine tumours, breast cancer, colorectal cancer. On ACBCS. Part of Paediatric Endocrine Tumour Guidelines working group. Cancer Genetics Group Steering Committee	***Clinical Genetics***	Louise.Izatt@gstt.nhs.uk

Kai Ren Ong (KRO)		Birmingham Women's Hospital	***Clinical Genetics***: Genetic susceptibility to cancer, Gorlin syndrome	***Clinical Genetics***	*Kai-ren.ong@bwnft.nhs.uk*
Munaza Ahmed (MA)		University Hospital Southampton Foundation Trust	***Clinical Genetics***: genetic susceptibility to cancer, information needs for informing cancer patients' about genetic testing in mainstream oncology.	***Clinical Genetics:*** Childhood cancer syndromes	*Munaza.Ahmed@uhs.nhs.uk*
Neil Rajan (NR)	Institute of Genetic Medicine, Newcastle	Institute of Genetic Medicine, Newcastle	***Molecular Genetics:*** genetic susceptibility to dermatogenetic disorders	***Dermatology***	*neil.rajan@newcastle.ac.uk*
Emma Woodward (EW)		Birmingham Women's NHS Trust	***Molecular and clinical Genomics:*** Genetic susceptibility to renal cancer, endocrine tumours. Cancer Genetics Group Steering Committee		*E.R.Woodward@bham.ac.uk*
Kelly Kohut (KK)	ICR, London (Hon. Contract)	Royal Marsden NHS Trust, London	**Genetic Counselling and psychosocial aspects of genetic/genomic medicine:** clinical management of families with genetic susceptibility to cancer, informed consent, delivery of test results, family communication, duty to warn	***Genetic Counselling:*** breast-ovarian cancer predisposition	*kohut.kelly@gmail.com*
Michael Sternburg (MS)	Imperial college, London		***Computational biology:*** 3-D protein-protein structural modelling and risk effect prediction. SuSPECAT tool		*m.sternberg@imperial.ac.uk*
Antony Carr (AC)	***University of Sussex:*** Genome Damage and Stability Centre (GDSC)		***Molecular pathology:*** Functional assays, molecular and structural modelling and analysis relating to DNA repair, genome stability, cell cycle chromatin dynamics and intracellular signal transduction pathways.		*a.m.carr@sussex.ac.uk*
Laurence Pearl (LP)	***University of Sussex:*** GDSC		***Molecular Pathology:*** structural modelling and structure-function studies of DNA damage response proteins		laurence.pearl@sussex.ac.uk
Mark O'Driscoll (Mo'D)	***University of Sussex:*** GDSC		***Molecular Pathology:*** molecular pathology, assay development, signalling pathways		m.o-driscoll@sussex.ac.uk

Jessica Downs (JD)	***University of Sussex:*** GDSC		***Molecular Pathology:*** chromatin effects, assay development for DNA damage responses and genome instability		J.A.Downs@sussex.ac.uk
Paul Pharoah (PP)	University of Cambridge	Public Health England	***Public health genetics, genetic epidemiology and clinical epidemiology:*** Survival analysis, risk analysis, data linkage. Genetic susceptibility to breast cancer, ovarian cancer	***Public Health Medicine***	*pp10001@medschl.cam.ac.uk*
Antonis Antoniou (AA)	University of Cambridge		***Genetic epidemiology/statistical genetics:*** Complex segregation analysis, risk prediction, cancer risk model development, genetic susceptibility to breast, ovarian, prostate cancer, risk modifiers for high-risk mutation carriers, CIMBA consortium, ENIGMA consortium		*aca20@medschl.cam.ac.uk*
Razvan Sultana (RS)	Genomics England		Software Engineer, Bioinformatics, Genomics: Bioinformatic analyses of large-scale variant datasets, hierarchical databases		*razvan.sultana@genomicsengland.co.uk*
Anneke Lucassen (AL)	University of Southampton	University Hospital Southampton Foundation Trust	Ethics: Ethical frameworks around delivery of genetic medicine	Clinical Genetics	*annekel@soton.ac.uk*
Ingrid Slade (IS)	University of Oxford	Oxford University Hospitals NHS Foundation Trust	Ethics: Ethical frameworks around delivery of genetic medicine	Public Health Medicine, Clinical Genetics	*ingrid@well.ox.ac.uk*
Jude Fitzgibbon (JF) (JF)	Queen Mary University of London		Molecular Genomics: Genetic susceptibility to familial leukaemia		*j.fitzgibbon@qmul.ac.uk*
Julian Barwell (JBa)		University Hospital Leicester	Cancer Genetics Group Steering Committee	Clinical Genetics	Julian.Barwell@uhl-tr.nhs.uk
Alex Henderson (AH)		Centre of Life, Newcastle		Clinical Genetics	Alex.Henderson@nuth.nhs.uk

Lynn Greenhalgh (LG)		Liverpool Women's Hospital		Clinical Genetics	Lynn.greenhalgh@lwh.nhs.uk
Carole Brewer (CB)		Exeter University Hospital		Clinical Genetics	carole.brewer1@nhs.net
Lucy Side (LS)	University College London	Great Ormond Street Hospital NHS Trust	Chair, Cancer Genetics Group Steering Committee		Lucy.Side@gosh.nhs.uk
Virginia Clowes (VC)		Northwick Park Hospital NHS Trust			v.clowes@nhs.net
Rachel Harrison (RH)		Nottingham University Hospitals			rachel.harrison@nuh.nhs.uk
Angela Brady (AB)		Northwick Park Hospital NHS Trust	Cancer Genetics Group Steering Committee	Clinical Genetics	angela.brady@nhs.net
Ian Berry (IB)	Leeds Genetics Laboratory	St James's Hospital, LeedsTeaching Hospitals NHS Trust	***Molecular Genetics:*** Genetic susceptibility to colorectal cancer, variant classification including bioinformatic methods. Member of InSiGHT and Variant Interpretation Committee (VIC).	***Clinical Scientist***	ianberry@nhs.net
Yvonne Wallis (YW)		Birmingham Women's Hospital		***Clinical Scientist***	YVONNE.WALLIS@bwnft.nhs.uk
Rachel Robinson (RR)	Leeds Genetics Laboratory	St James's Hospital, LeedsTeaching Hospitals NHS Trust		***Clinical Scientist***	rachell.robinson@nhs.net

Will Foulkes (WF)	McGill University, Montreal	Collège des Médecins du Quebec	Molecular Genomics: Genetic susceptibility to cancer (breast cancer, ovarian cancer, pleomorphic cancer syndromes)	Medical Genetics	william.foulkes@mcgill.ca
John Burn (JB)	University of Newcastle	Advisory Members			john.burn@newcastle.ac.uk
Gareth Evans (DGE)	University of Manchester				Gareth.Evans@cmft.nhs.uk