Genomics england

Genomics England Clinical Interpretation Partnership (GeCIP)

Guidance for expressions of interest to form domains within GeCIP

Authors: Professor Mark Caulfield, Professor Dame Kay Davies, Dr Tom Fowler, Dr Jeanna Mahon-Pearson, Dr Ellen Thomas, Dr Clare Turnbull

Issued by: Office of the Genomics England Chief Scientist

July 2015

Contents

1. Background and Introduction

About this document

This document provides guidance on the Genomics England Clinical Interpretation Partnership (GeCIP). The document defines the goal of GeCIP, describes how GeCIP will be organised and who can express interest in joining GeCIP. In addition we describe the opportunities from such a partnership and our expectations of GeCIP. This document is designed to support expressions of interest from self-organised UK-led domains including UK based researchers, clinicians, healthcare professionals, trainees and their international collaborators.

The 100,000 Genomes Project

In December 2012, the Prime Minister announced a programme of whole genome sequencing as part of the UK Government's Life Sciences Strategy. The principal objective of the project is to sequence 100,000 genomes from patients with cancer, rare inherited disorders, and infectious disease drawn from the NHS in England, and to link the sequence data to a standardised, extensible account of diagnosis, treatment, and outcomes.

This project is designed to produce new capability and capacity for genomic medicine that will transform the NHS. It will also produce new capability for clinical genomics research and training. As part of the proposal, a secure infrastructure will be established for the protection and analysis of clinical and genomic data, and made available for approved academic and industrial research purposes, including those of the contributing clinical organisations from the NHS.

Genomics England – progress with implementation of the 100,000 Genomes Project

The Department of Health has established Genomics England as a wholly owned, limited company to deliver the project. Genomics England is working with NHS England, Public Health England, Health Education England, and NHS Trusts. This is to ensure that the project is fully aligned with NHS transformation and sits within a programme of related initiatives in clinical and laboratory genetics, molecular pathology, service innovation, disease registration, clinical audit, training and technology funding. To identify and enrol participants, NHS England is creating NHS Genomic Medicine Centres (GMCs) in England, which will start recruiting patients and collecting samples and data in early 2015. In August 2014 Genomics England announced a partnership with Wellcome Trust, The Wellcome Trust Sanger Institute and Illumina to create a Genomic Medicine Sequencing Centre in Hinxton which will undertake whole genome sequencing for the main programme.

The Genomics England Clinical Interpretation Partnership

Genomics England now wishes to initiate work with research groups and funding organisations, to ensure that this new research opportunity is fully realised. In addition, Genomics England and its partners will ensure that the tools provided within the secure data infrastructure offer opportunities to both accelerate scientific progress and support focused, interdisciplinary collaboration needed for highest calibre clinical interpretation and maximise patient benefit. To ensure we maximise the value of this programme we have created the Genomics England Clinical Interpretation Partnership (GeCIP). GeCIP is organised into disease-specific and function-specific domains. This is UK-wide (not limited to England) and brings funders, researchers, NHS clinicians/healthcare professionals, trainees and potentially industrial partners together to enhance the value of this dataset for healthcare benefit. The Genomics England Clinical Interpretation Partnership was launched on 27 June 2014. Further information can be found at www.genomicsengland.co.uk.

Maximising the reach of this programme

Our mission is focused on combining clinical and whole genome sequencing data in rare disease, cancer and infection from patients under the care of the NHS in England and it does not extend to other diseases at present. However, we have developed this infrastructure with the opportunity to expand the programme to other diseases and approaches if separate funding is generated. In particular our MRC-funded data infrastructure award is designed to provide the UK Clinical Research Data Infrastructure in Genomic Medicine and could be used to store whole genome sequences from other disorders and a range of associated multi-omic datasets. Furthermore, our sequencing contract contains some elasticity to go beyond 100,000 whole genome sequences, although we do not have funding for this additional activity at present. Genomics England is not able to pay for sequencing of patients who are not from the NHS in England at this time, or from diseases that are outside the scope of rare disease, cancer or infectious disease. Sequencing patients from the devolved nations will require separate funding and discussions are underway.

2. Aims, activities and expectations of the Genomics England Clinical Interpretation Partnership

Aims of GeCIP

The overall aim of the Genomics England Clinical Interpretation Partnership is to create a thriving, sustainable environment for researchers and clinical (NHS) disease experts and trainees and a pre-competitive industrial consortium. This community will analyse and constantly refine the clinical interpretation of the 100,000 genomes dataset. There are three overarching aims of GeCIP:

(i) To optimise clinical data and sample collection, clinical reporting and data interpretation for return to clinicians and patients

(ii) To perform research to further improve our understanding of the implications of the findings for genomic medicine in the clinical setting and

(iii) To provide a rich training environment for trainees both within the Genomics Education Programmes of Health Education England.

Scientific activities of GeCIP

It is anticipated that the scientific activities undertaken within GeCIP will include the following areas:

- Research providing enhanced clinical interpretation focused on **rare disease**, including clinically or genomically-driven deeper phenotyping, novel approaches to interpretation and annotation, validation and functional characterisation of variants adding multi-omic data, identification of novel therapeutic targets or repurposing of existing therapies.
- Research providing innovative clinical interpretation in **cancer**, including multi-omic datasets (e.g. transcriptomics, epigenetics, proteomics), optimising molecular pathology pathways, analysis of circulating tumour DNA, sequential biopsy of metastatic disease to elucidate the evolving genetic architecture of cancer, validation and characterisation of variants, identification of novel therapeutic targets or repurposing of existing therapies.
- Research providing improved clinical interpretation in **infectious disease**, focused upon individuals with severe outcomes in sepsis or, in partnership with Public Health England, bacterial and viral genomic sequencing to allow a greater understanding of the spread of antimicrobial resistance and phylogenetic tracking of transmission across the whole of the health economy.
- Research into **stratified healthcare approaches** including innovative clinical trial designs into any of the diseases in scope facilitating externally funded programmes to drive new affordable healthcare in rare disease, cancer and infection.
- Capitalising upon **electronic health records research infrastructure**, such as, the Farr Institute and the Oxford Big Data Centre to build upon and add value to the clinical, laboratory, and health records data.
- Development of **algorithms, models, tools and resources** for clinical genomics research including those designed to improve data quality assurance, annotation, interpretation, and presentation of genomic, clinical, and laboratory data in combination.
- **Training** in any relevant area to the 100,000 Genomes project including individual fellowships or studentships directly related to the Genomics England Programme especially those funded by the Health Education England Programme in Genomics that runs alongside.

The GeCIP Board are also open to scientific proposals outside the range of activities described above.

Expectations from GeCIP

The Genomics England Clinical Interpretation Partnership Board anticipates that this initiative will add value to the 100,000 Genomes Project through activity in the following areas:

- Supporting Genomics England by advising on diseases nominated from the NHS in England ensuring enrolment of appropriate patients, collection of **optimal samples and clinical data** relevant to each disease-specific domain.
- Informing NHS feedback to clinicians and the multidisciplinary teams by providing disease-specific expertise to inform **clinical reporting and data interpretation**. The process of NHS feedback and issue of reports will be managed by Genomics England and the details for this process will be outlined in the 100,000 Genomes Project protocol.
- Engaging clinical teams in the NHS who directly care for patients to assist in **evaluation and development of the evidence** for variants of expected or uncertain but potential pathogenicity. It is conceivable that enrolment of additional family members will allow segregation patterns and parallel focused functional assays may contribute to confirmation of pathogenicity status. Of note, these findings will only be fed back to patients if they transfer from expected pathogenic to known pathogenicity and are validated, therefore becoming clearly clinically actionable.
- Generating high impact research outputs, new grant funding and fellowship awards and new avenues of investigation.
- Priming therapeutic innovation and stratified healthcare opportunities.

3. Constituency of the Genomics England Clinical Interpretation Partnership

The Genomics England Clinical Interpretation Partnership is a partnership of funders, researchers, trainees, clinicians and healthcare professionals.

GeCIP Funders

Genomics England, which is delivering the 100,000 Genomes Project, is working with key funders to establish the Genomics England Clinical Interpretation Partnership. GeCIP is a mechanism for public, charity, philanthropic, and the researchers that they support, to engage with the Genomics England Programme and add value in terms of clinical interpretation and collaborative research and training. There are no funds specifically available through this call from any funder.

GeCIP Researchers and Trainees

It is the ambition of Genomics England and all its funders to see a powerful research and training programme alongside the Genomics England Programme. This will create opportunities for GeCIP researchers and trainees that will stimulate and support research programmes funded by GeCIP funding partners through response mode and specific funding calls. The researchers leading these programmes will be expected to form and lead appropriate national and international consortia within cognate domains that provide eligible NHS samples, clinical data, and analytical skillsets to maximise the value of the 100,000 Genomes Project dataset. These GeCIP domains must ensure that the clinical interpretation delivered by the programme remains at the forefront of genomic medicine. Genomics England is committed to creating a vibrant training environment within GeCIP and each domain will need a Training Director. All trainees are welcome to contact Genomics England for advice and support with joining domains. It is anticipated that the majority of research and training will occur within the GeCIP domains.

GeCIP Clinicians/Healthcare Professionals

The GeCIP domains include clinicians and healthcare professionals from the NHS and other organisations who enrol patients, contribute samples and/or data to the programme through NHS Genomic Medicine Centres. This approach offers the opportunity for those in contributing NHS organisations to access the 100,000 Genomes Project dataset and Genomics England Knowledge Base and to participate in the analysis and interpretation of sequencing on patients within their disease area (all data in the research data infrastructure will be pseudonymised)*. These NHS clinicians and healthcare professionals will be able to engage in and add value to the GeCIP domain by nominating diseases for the 100,000 Genomes Project, providing samples, data, annotation, tools and expertise. Those funded on specific NHS training programmes related to Genomics England (NHS Trusts, NHE England, Public Health England and Health Education England) will also be GeCIP clinicians/ healthcare professionals.

* This is not the route for NHS organisations to review identifiable patient data. There will be a separate mechanism for that.

Commercial Partners

This call is not the route for industry to form partnerships with Genomics England or for commercial entities to access the 100,000 Genomes Project. Interested commercial organisations or academic-industry collaboratives should contact Genomics England via the website **www.genomicsengland.co.uk.** A key part of the Genomics England mission is to generate health and wealth for the UK. As part of that mission, we have formed a consortium of industrial partners who will work in the pre-competitive environment alongside GeCIP members (described above) with the goal of accelerating diagnostic, analytical and therapeutic advances for healthcare benefit. In the event that we introduce industry partners to a GeCIP domain all work done will be in open innovation space and Genomics England will own the clinical and sequence data and any results derived from these data. Genomics England will license intellectual property back to academics, clinicians and commercial partners if appropriate.

4. The model for the Genomics England Clinical Interpretation Partnership

The Genomics England Office of the Chief Scientist is responsible for establishing and delivering GeCIP. During the 100,000 Genomes Project, Genomics England will support and provide GeCIP members with an informatics platform and datasets as released from the Project. The data generated by this programme is owned by Genomics England, a wholly owned Department of Health company.

Organisation of GeCIP into domains

We expect UK-led multidisciplinary groups of researchers, NHS healthcare professionals, and trainees to come together to propose **disease-specific or function-specific domains** (depending on their area of expertise) with clear evidence of international leading research capability. Each proposed domain must clearly add value to the Genomics England Programme by contributing patients, phenotypes, knowledge, analytic applications, expertise, and undertaking cutting edge research that enhances clinical interpretation. Each domain will self-nominate a UK based domain leader with an NHS contract. This stipulation is applied because this is an NHS transformation programme and the leader (who could be non-clinical or clinical) must be connected to the healthcare system in which the enhanced clinical interpretation will be applied. The GeCIP Board reserve the right to accept or not the domain's nominee.

We are seeking domains focused upon rare diseases, cancer and infectious disease but anticipate there may also be domains that are function-specific rather than disease-specific. These might include domains focused on ethics and social science or advanced analytical methodologies such as machine learning. We anticipate that domains will be broad for example, a rare inherited cardiovascular disease domain or a breast cancer domain. We anticipate that within domains logical subdomains may emerge, which are focused on more specific disease subtypes, syndromes or functional activities. At this time we do not wish to receive applications to form GeCIP domains for diseases that are not within the scope of the 100,000 Genomes Project.

The scale of an individual GeCIP domain

We expect the UK research and NHS community to organise and lead these GeCIP domains and to represent a substantial majority of the domain membership. We envisage to sequence approximately 75,000 people with rare disease and cancer. The precise numbers of individual whole genomes have yet to be determined. However, we do not envisage a rare disease domain to have more than 1500 whole genome sequences and not more than 4000 whole genome sequences in cancer to evaluate. This is subject to change as the final figures have not been approved by the Science Committee.

International collaborators and GeCIP

We also recognise that genomic research is usually advanced most rapidly by international collaboration. Genomics England do expect individuals forming consortia to include selected international collaborators who would clearly add demonstrable value to the GeCIP domain. Non-UK researchers and scientists who are not invited to become members of a GeCIP domain can apply to access the data but will have to wait six months and pay a subscription for access.

Oversight of GeCIP domains

The Genomics England Office of the Chief Scientist will support and oversee each domain to ensure compatibility and optimise alignment of the activity within each disease across the full Clinical Interpretation Partnership. These domains will only exist as long as they as they are productive and clearly deliver the Genomics England goal of enhanced clinical interpretation. We reserve the right to change the leadership, dissolve and reform domains that are ineffective or dysfunctional.

5. The role and value of Genomics England Clinical Interpretation Partnership domains

The role of GeCIP domains

The overall aim of this model is to create a thriving, sustainable, integrated, dynamic community of research and clinical (NHS) disease experts, which exchanges information and incorporates pertinent information from the global knowledge base for the disease. This work will encompass activity addressing both the more immediate clinical remit and developing the longer-term scientific activities, as laid out for the GeCIP domain. Within these remits, we anticipate that each GeCIP domain will:

- Form sub-domains facing specific disease sub-groups and specific tasks e.g. analytics, function, deeper phenotyping and training.
- Provide expert insights from the existing global genomic and clinical knowledge base of the disease in order to inform clinical reporting and data interpretation to ensure return of optimal data to NHS clinicians and patients.
- Iteratively clinically validate and improve the knowledge base within the 100,000 Genomes Project dataset through return of enhanced phenotypic patient data.
- Generate new research findings through large-scale analyses of the 100,000 Genomes Project dataset.
- Generate new research findings through specific research projects related to the 100,000 Genomes Project dataset and/or multi-omic samples collected as part of the programme.
- Harness opportunities to extend the value of the 100,000 Genomes Project dataset and generate insights into clinical genomics through international collaborative research.
- Provide research training in genomics to a new generation of research trainees.

Figure 1: A potential structure for a GeCIP domain

Genomics England Clinical Interpretation Partnership Disease-specific Domain

Activity/Group	Key functions and outputs
The Genomics England Office of the Chief Scientist	• Oversight, informatics and logistics
GeCIP Steering Committee	• Coordination and management of domains
Clinical Interpretation including NHS clinicians	• Highest fidelity dynamic reporting system • Contribution to Interpretation, Validation and Feedback domain
Genomics researchers	• Novel genomic discoveries • Engagement of international collaborators
Multiple phenotypic sub-groups	• Deeper phenotyping and sub-analyses
Analysts and Bio-informaticians	• Novel analytic approaches • Development of analytical and reference resources
Functional characterisation Multi-omics	• Single cell or model functional studies • RNA, epigenetics, proteomics
Trainees and training director	• Research projects and higher degrees through the Genomic Medicine Academy
Precompetitive industry partners	• Academic-industry collaboration to accelerate application of new findings

The value of working within a GeCIP domain

The Genomics England Clinical Interpretation Partnership is a rich collaborative environment which is organised into disease-specific and function-specific domains such that multidisciplinary expertise comes together to drive up the fidelity and relevance of whole genome sequencing. The members of a GeCIP domain will help shape and inform the direction and scope of the 100,000 Genomes Project including disease focus, sample collection, clinical phenotypic data acquisition, functional characterisation of findings, innovative analytical approaches, supporting stratified healthcare and therapeutic innovation with the purpose of enhanced clinical data interpretation and feedback in their disease area. It is a condition of membership that each domain interacts with and assists Genomics England in delivering optimal data and clinical interpretation back to NHS clinicians and patients. To support this work, Genomics England will provide a domain-specific embassy and role based access control framework so that members can access the tools, data and knowledge available.

Domains will create and augment a rich dataset for each disease greatly adding value to the 100,000 Genomes Project. In return for the contribution of expertise from a domain towards optimising the clinical reporting and interpretation, the domain will have priority access for a period of six months to the 100,000 Genomes Project dataset pertaining to their domain, with no access fee to conduct and publish research. After six months the data will be released to other users creating an incentive for the GeCIP domains to expedite analysis and publication. In addition, members will have an opportunity to apply for priority access to additional multi–omics samples for research. In addition, domain status will confer advantage when applying for strategic initiatives and response-mode funding from GeCIP funders.

6. Requirements of Genomics England from the GeCIP domains

Research focused on clinical genomic interpretation

It is expected and a condition of membership that GeCIP domains advance research into genomic interpretation through analyses of the 100,000 Genomes Project dataset and we envisage this will include:

- Enhanced clinical interpretation focused on **rare disease**, including clinically or genomically-driven deeper phenotyping, novel approaches to interpretation and annotation, validation and functional characterisation of variants, identification of novel therapeutic targets, or repurposing of existing therapies.
- Innovative clinical interpretation in **cancer**, including sequential biopsy to address the genetic architecture of cancer, validation and characterisation of variants, identification of novel therapeutic targets, or repurposing of existing therapies.
- Improved clinical interpretation in **infectious disease**, focused upon individuals with severe outcomes in sepsis, or – in partnership with Public Health England – greater understanding of the spread of antimicrobial resistance and phylogenetic tracking of transmission across the whole of the health economy.
- Improved validation and patient feedback, ethical and social science research and health economics for whole genome sequence in an NHS context.

Working with the interpretation, validation and feedback domain

A key remit of each disease-facing domain is the provision of expert, disease-specific input to update clinical reporting and data interpretation. This will include provision and maintenance of lists of genes to prioritise for clinical reporting (virtual gene panels). Each GeCIP domain will be expected to regularly review at least the data held in the Genomics England Knowledge Base relating to the specific disease area in question. Over time, the clinical interpretation of variants found in the 100,000 Genomes Project dataset is likely to lead to greater certainty and in some cases, validation, of the clinical impact of the variant. Such results will be committed to the Genomics England Knowledge Base, to be drawn upon by other GeCIP users.

There will be a specific domain, constituted by the Genomics England Chief Scientist which will take over-arching focus on standards and policy for interpretation, validation and feedback of variant data. This GeCIP domain will include in its membership:

- Clinical geneticists, cancer genomics clinicians, research and NHS laboratory scientists
- GMC representation
- Ethics representation
- Validation and feedback representation
- Patient representation

All clinically-facing domains will provide a named member who will contribute to this activity in clinical reporting and data interpretation on behalf of that disease domain. A particular focus of the Interpretation Validation and Feedback domain will be development of the policy on feedback for secondary (looked-for) findings to patients.
This domain will be responsible for defining which variants are fed back as secondary findings and will maintain the list of genes in which these findings are sought. This work will be developed in close consultation with patient representatives, with specialist input from ethics and in line with international developments in the field.

Working with the regulatory authorities

The GeCIP domains are expected to collaborate as needed with regulators, including E-Quality Management System, National Institute of Health and Clinical Excellence, Medicines and Healthcare Regulatory Products Agency, and NHS England to develop standardised approved diagnostic approaches to validate findings from this programme. In addition, this collaboration is expected to support the development of systematic approaches to verification of scientific findings, evaluation of proposed applications of such findings on diagnosis and treatment options, and management of impacts upon commissioning strategies and costs. This joined up approach will ensure swift translation of clinically validated research findings into patient benefit.

Commitment of data to the Genomics England Knowledge Base

Data will be released into GeCIP in a series of builds. Genomics England will encourage and promote an environment of open collaborative approach and GeCIP members will be expected to comply with this approach to enhance and maximise the value of the Genomics England Knowledge Base. At all times the GeCIP domains must adhere to the Genomics England Data Access, Sharing and Acceptable Uses Policy, which can be found on the Libary and Resources page of www.genomicsengland.co.uk. The Genomics England individual level data cannot be removed from the Genomics England Data Centre. Analyses can be run inside the Data Centre and the results exported for publication. In all cases the research results (along with raw data and analytical steps) undertaken within the GeCIP domains must be provided to the Genomics England team, for commitment to the Genomics England Knowledge Base.

- Domains will have specific 'embassies', which comprise of the sub-section of the data related to that domain and computer resources housed within the Genomics England infrastructure.

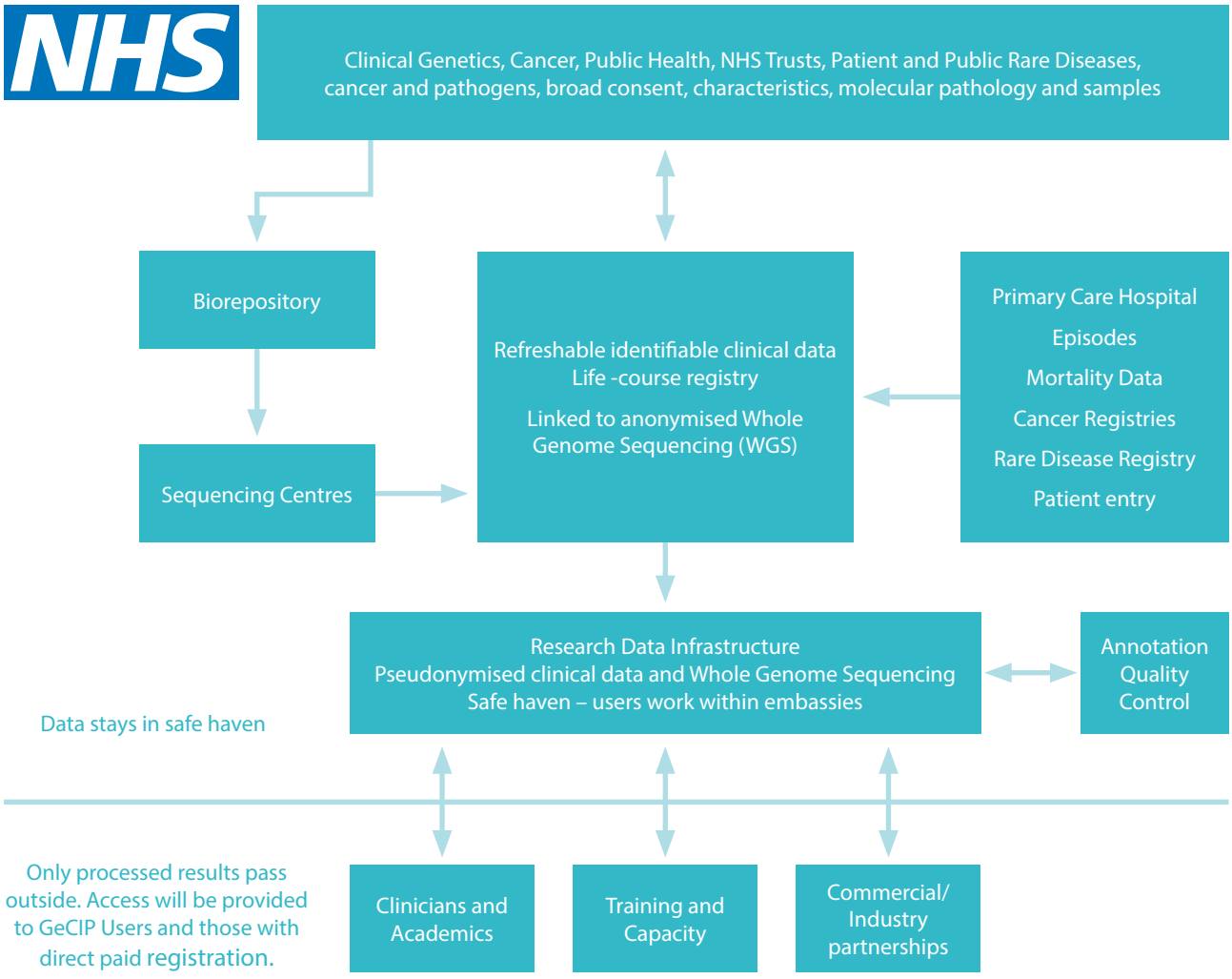
These embassies are intended to be seen as areas in which healthcare professionals, researchers, trainees and pre-competitive industry partners undertake their work. Work undertaken within the embassies is subject to the governance and terms and conditions of Genomics England.

- All GeCIP Domains will need to adhere to Genomics England Data Access, Sharing and Acceptable Uses Policy, accept our Intellectual Property Policy and Publication Policy.
- Breaches of the governance or terms and conditions of use will result in the removal and exclusion of those responsible from the Genomics England data centre.
- All research data will be de-identified in accordance with their REC approved protocols.

Training

Each domain will need a training lead who will ensure that the needs of trainees are looked after and new trainees are welcomed and inducted into the domain. This is especially important as £25m is allocated to extensive training programmes by Health Education England. Genomics England expects each GeCIP to be a rich training environment.

Figure 2: Genomics England Clinical Interpretation Partnership Operational plan



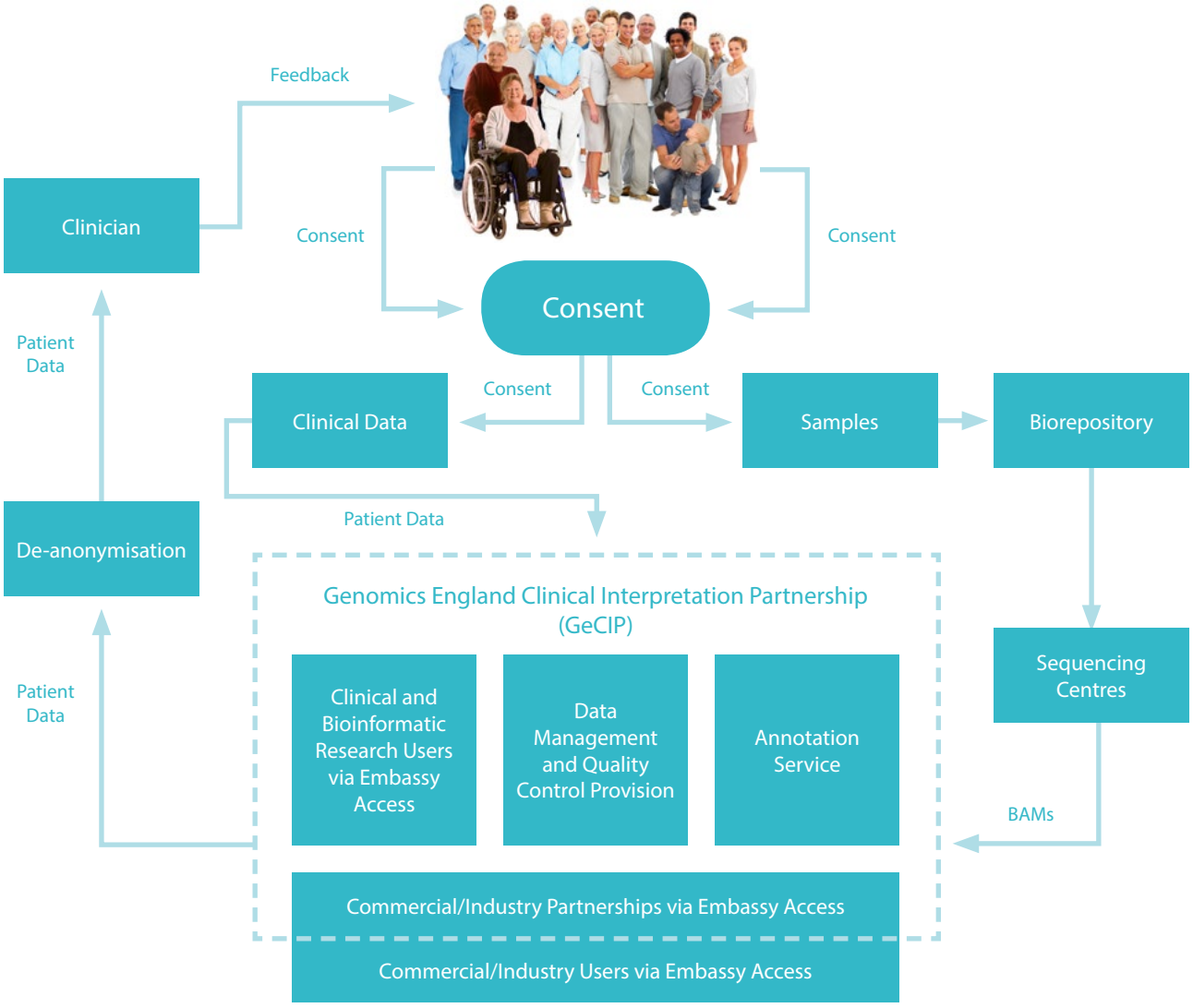
7. Opportunities for researchers within Genomics England Clinical Interpretation Partnership domains

Data available to the GeCIP domains

To maximise fully the opportunities for research, Genomics England will integrate high-fidelity clinical phenotypes and whole genome sequence data from consented patients with electronic health data, such as, primary care records, hospital episode statistics (HES), data from all relevant national registries and mortality. We envisage that the clinical data, raw read sequencing data, the variant call files and more than one version of annotation will be available to the Clinical Interpretation Partnership. By linking phenotypic and long term outcomes data with whole genome sequence we will generate a uniquely rich and longitudinal dataset for GeCIP available via controlled access to the data for the clinical genomics research community. There will be an independent Access Review Committee who will consider and approve access to data for all Clinical Interpretation Partners.

In a partnership with several universities and institutes (funded by a £24m grant from MRC) we are establishing a shared, secure, high performance data and computing infrastructure (circa 30,000 core processing units and 30,000 petabytes of storage with digital tape back-up), as a platform for innovative clinical genomics research. This MRC funded platform will have the interoperability needed to facilitate UK and international research collaboration, and to support collaborative working between academia and industry. It will provide the home for GeCIP maximising the translational research potential of the 100,000 Genomes Project. Delivery of the new capability presents a unique opportunity for UK clinical research that will enable the discovery of new diagnostics, test complex approaches to stratified medicine, and drive therapeutic innovation.

Figure 3: Genomics England Data Flows



Opportunities for additional whole genome sequencing

It is anticipated that funding for additional whole genome sequences beyond those currently included within the 100,000 Genomes Project will be sought. Anyone hoping to do this must get approval via the Genomics England Nomination Process prior to any application to funders to support WGS of additional samples. It would be necessary for researchers seeking funds for other diseases to secure the funds to enrol, characterise and acquire samples and funds for storage and sequencing as this is not funded as part of the Genomics England Programme.

Research opportunities beyond whole genome sequencing

Genomics England will create a high quality biological sample resource from participants in support of future health research. An essential element of this collection is to provide a resource that supports scale (power) and diversity (of sample types and analytic options) across different 'omics' platforms. The plan is to collect the following from NHS England Genomic Medicine Centres from early 2015:

- 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

Dependent on additional funding, we may collect in the future:

- Cancer cell lines for study or xenotransplantation cancer models (not funded at present)
- Skin biopsies for generation of inducible pluripotent stem cells will be possible under additional consent if funds are raised (not funded at present)

The collection of these samples will ensure that biological samples obtained are suitable for assessing the genome, epigenome, proteome and metabolome in blood and tumour. Analysis of the additional sample assays may not be undertaken immediately after collection but stored until funds are available for future research analysis. The storage and processing of samples will take into account the assays that are anticipated to be the most likely to be undertaken in future, for example whole genome sequencing, transcriptomic, epigenetic, metabolomics and proteomic research.

Lifelong electronic health record linkage

The partnership with the NHS is particularly important to deliver high-quality initial phenotyping data; a flow of electronic health data from primary care, hospital, outcomes, registries and social care records; and an opportunity to work with clinicians and patients to acquire further information on the primary conditions, associated comorbidities, and outcomes over an entire life course.

Research opportunities to extend beyond the primary phenotypes

The evaluation of whole genome sequence data in the context of rich and extended phenotypes and exposures derived from electronic health records. The richness of the 100,000 Genomes Project dataset will allow us to move beyond the primary phenotype of the rare disease, cancer or infectious disease which led to the patient engagement to evaluate the whole genome sequencing in the context of other continuous traits, diseases and response to therapy. To do this it is expected that GeCIP members will learn from Biobank UK, and engage with existing infrastructure, such as the Oxford Big Data Institute and the Farr Institute.

8. Governance, oversight and management of the Genomics England Clinical Interpretation Partnership

The GeCIP Board

The Genomics England Clinical Interpretation Partnership is overseen by the GeCIP Board, chaired by Professor Dame Kay Davies, who will also represent GeCIP on the Board of Genomics England. The GeCIP Board was constituted on Friday 17 June 2015.

Each GeCIP funder is represented on the GeCIP Board. Members of the GeCIP Board are: Dr David Cox (National Institute for Health Research/Department of Health), Dr Nathan Richardson (Medical Research Council), Dr Michael Dunn (the Wellcome Trust) and Dr Rowena Sharp (Cancer Research UK). Genomics England is represented by its Chief Scientist and Chief Technology Officer. Members of the board include a GeCIP domain lead elected from within each of the cancer, rare diseases, and infectious diseases themes. There will also be patient representatives.

The main roles of the GeCIP Board are:

- To oversee the Clinical Interpretation Programme and act as a focus for coordination
- To sponsor a twice yearly conference of all active GeCIP members
- To receive advice from Genomics England Advisory Committees
- To provide advice to the Genomics England Board

Role of the Office of the Genomics England Chief Scientist

GeCIP is operated by the Office of the Genomics England Chief Scientist which will oversee the following activities:

- Oversee GeCIP overall management and coordination
- Manage appointment process for GeCIP science domain members
- Support the work of the GeCIP science domains; facilitate meetings and knowledge sharing, provide administrative support
- Manage the interface with GeCIP funders
- Manage the interface with NHSE, clinical regulatory bodies, NICE and MHRA
- Manage the interface and relationship with the Royal Colleges
- Enable Health Education England and other funders' training programmes
- Deliver informatics training and support
- Provide input to Genomics England data sharing agreements and contracts with users
- Deliver data embassy (access to anonymised data and Genomics England Knowledge Base for the disease area in question)
- Deliver operational governance oversight (provide assurance to the Genomics England Board on contractual obligations and data security and ethical governance compliance within the GeCIP environment and user community)
- Maintain the list of pathogenic variants that may be fed back to treating clinicians (incidental findings)

The GeCIP Steering Committee

The Genomics England Office of the Chief Scientist will convene a Steering Committee to input into the smooth running and transfer of knowledge across the GeCIP domains. The membership will consist of a representative from each domain and will be chaired by the Chief Scientist for Genomics England. This committee will include the Genomics England Office of the Chief Scientist, representation from the Genomics England Informatics team, representation from the Ethics Advisory Committee and representatives of the public and patients. The GeCIP Steering Committee will directly report to the GeCIP Board.

GeCIP members

All GeCIP researchers and clinicians will adopt the highest ethical standards with regard to research governance, data access and security in accordance with the terms of the Genomics England informed consent.

9. Genomics England Clinical Interpretation Partnership funding and funders

Modes of funding from GeCIP funders are still under discussion but may include:

- Funding calls that explicitly fund research proposals that use the 100,000 Genomes Project dataset and contribute to the Genomics England Knowledge Base
- Strategic initiatives initiated by GeCIP researchers which are directly targeted at enhancing clinical interpretation in partnership with Genomics England
- Response mode funding from applications submitted to funders initiated by GeCIP researchers
- Funding for training as part of either the allied Health Education England programme or other trainee programmes related to the 100,000 Genomes Project

Applications to funders by GeCIP domains

Applications for funding to external funders will need to be registered with the Office of the Chief Scientist so we can manage competing demands fairly and equitably and ensure these accord with Genomics England mission and ethical approvals. Genomics England cannot award grants to GeCIP domains – we are providing the Data Centre, 100,000 whole genome sequences and associated clinical data and a multi-omic sample repository. Instead, GeCIP domains will need to seek funds by application to GeCIP funders or other funders. These applications will be subject to approval via the scientific peer review process of the respective GeCIP funders. This will ensure that whilst incentivising individuals to become GeCIP researchers, the research undertaken will be of the highest quality. We strongly encourage trainees from any relevant field to contact us when considering fellowship or grant applications.

While funding calls will be based on GeCIP funders' own priorities and assessment processes, we expect funding calls to drive collaborations between researchers, both nationally and internationally, as such collaborations are likely to result in stronger scientific researcher proposals.

10. How to become part of the Genomics England Clinical Interpretation Partnership

Forming a GeCIP domain

GeCIP is organised into a series of disease-specific or function-specific domains of UK researchers, trainees, NHS clinicians and healthcare professionals. An open call was put out for groups to self-organise and create proposed GeCIP domains in November 2014. Some clinicians and researchers have proposed to work in several domains if they have a broader clinical, analytical or laboratory remit.
The researchers leading these programmes will have access to appropriate patient groups from the NHS in England (the devolved nations may join later). The GeCIP domains can be led and proposed by researchers, the NHS and those in training can be from anywhere in the UK and are not confined to England.

Below is a list of the current GeCIP domains (as of July 2015). Proposals are invited for additional domains pertaining to other rare disease groups, cancers or cross-cutting themes. Please note that applications for existing domains will **not** be considered. Individuals wishing to join the domains listed below should contact the Office of the Chief Scientist directly.

For a number of the domains, in particular in rare diseases, there will be subdomains as dictated by phenotypic subgroups; it is recommended that the structuring of rare disease domains incorporates expertise in subdomains as appropriate. All GeCIP domains have self-nominated a UK lead. The Genomics England Office of the Chief Scientist will work with applicants to optimize the constituency of each domain needed.

Disease-specific domains		Function-specific domains
Cancers	**Rare Inherited Diseases**	
Breast	Hearing and sight	Electronic records
Lung	Cardiovascular disease	Validation and feedback
Ovarian	Respiratory	Ethics and social science
Colorectal	Endocrine and metabolism	Functional effects
Prostate	Gastroenterology and hepatology	Health economics
Haematological malignancy	Immunology and haematology	Machine learning, quantitative methods and functional genomics
Childhood solid	Neurological	Population genomics
Pan Cancer	Musculoskeletal	Enabling rare disease translational genomics via advanced analytics and international interoperability
	Renal	Functional cross cutting
	Skin	Education and training
	Paediatric Sepsis	
	Inherited cancer predisposition	
	Paediatrics	

Assistance and support with GeCIP domain applications

The Office of the Chief Scientist is committed to making GeCIP a success and are ready to help you with your application by email, phone or face to face. We can do teleconferences or meet co-applicants. If you wish to meet in London we will try to arrange a room at Queen Mary University of London.

Step 1: Expressions of interest

Groups who wish to express an interest in forming a GeCIP domain should complete the **GeCIP Domain Application Form** which can be found at **www.genomicsengland.co.uk.** The online application form asks for the following information:

- The domain name (disease/phenotype/analytical area or other activity).
- Group members, their affiliations, a succinct list of their expertise, the contribution they will make and the value they add (1 page CV each to be uploaded). The constituency will include but is not limited to academics, NHS clinicians, healthcare professionals and trainees.
- A proposed UK leader for the domain (2 page CV to be uploaded).
- Demonstration of appropriate representation of expertise and skills as outlined above.
- A brief outline of plans for how the 100,000 Genomes Project will be utilised for research and discovery within their domain.
- Proposed additional research activity including research activity focused on multi-omics and/or longitudinal data (as appropriate to domain).
- Current funding and plans for procurement of funding for proposed research activities.
- Opportunities for education and training.
- Potential international collaborators (brief letters of collaboration to be uploaded if applicable).
- Mechanisms for pre-competitive interaction with partners from industry (if applicable).

Step 2: Genomics England review and adjudication

The preliminary applications will be reviewed and adjudicated by the Genomics England Chief Scientist, the Science Advisory Committee and Professor Dame Kay Davies. These applications will be objectively evaluated with regard to their demonstration of (i) expertise in clinical reporting and data interpretation (ii) scientific track-record, research plans, existing infrastructure and collaborations (iii) opportunities for training. If there are clear gaps in expertise within an application, it may be recommended that these gaps are rectified via inclusion of additional applicants from outside of the original application. Genomics England retains the right to refashion, reorganise or merge domains to suit primary goals of the programme. If there are multiple applications for a domain, adjudication may be required and we may suggest reorganisation of applying groups to produce an optimal domain. We also retain the right to reject domains.

Step 3: Approval of GeCIP domains for introductory period

We will inaugurate the most functional proposals as domains for an introductory period of 12 months during which they must demonstrate effectiveness in improving clinical interpretation and generating new research opportunities.

Step 4: Submission of a full research plan

The elected domain will be required to submit a structured detailed document detailing allocations of activities, plans for research, plans for collaborations, timelines for delivery, timelines for acquisition of funding within three months of inauguration.

GeCIP domain monitoring

The activities and progress of each domain will be reviewed every six months via the GeCIP Steering Committee with oversight from the GeCIP Board. If a GeCIP domain is failing to deliver or failing to progress in its programme of research and training, the domain may be required to restructure or dissolve.

Applications for new GeCIP domains

Applications for domains that are not established within the first application process will be welcomed and will be reviewed by the GeCIP Steering Committee with oversight from the GeCIP Board.

Timelines for GeCIP domain applications

July 2015:	Second call for expressions of interest to form GeCIP domains. For further information and the application form, please see **www.genomicsengland.co.uk**
October 2015:	Deadline for second wave for submissions of GeCIP domains on Friday 30th October at 5pm (UK time).
Nov 2015:	Review and evaluation of submissions by the Office of the Chief Scientist.
Dec 2015:	Establishment of second wave of GeCIP domains.

Becoming a GeCIP clinician/healthcare professional

NHS clinicians and healthcare professionals who have recruited patients and/or collected clinical data are automatically eligible to become GeCIP members. These clinician/healthcare professionals can apply for access to the Genomics England Knowledge Base and part of the 100,000 Genomes Project dataset that relate to their patients via the access provided to the recruiting NHS Genomic Medicine Centre.

Becoming a 100,000 Genomes Project researcher outside of a GeCIP domain

The invitation to register expressions of interest in domain formation is not the route for applications to access the 100,000 genomes dataset outside of a GeCIP domain group. We will announce a separate process for applications to access the dataset for researchers outside GeCIP in due course. A fee may be levied if that researcher has not otherwise contributed to GeCIP via commitment of samples or clinical data to the 100,000 Genomes Project dataset or via commitment of expertise to the Genomics England knowledge base.

How to contact us

For further information and guidance please email us at **chiefscientist@genomicsengland.co.uk**

11. Genomics England Data Ownership, Intellectual Property and Publication Policy

Genomics England Data Ownership and Intellectual Property

Genomics England has received legal advice regarding intellectual property and it was recommended that Genomics England owns the combination of the whole genome sequence and the clinical data for the entire dataset from the 100,000 Genomes Project. In addition Genomics England owns any new intellectual property generated from the data but we will license this to third parties the opportunity to commercialise opportunities on favourable terms.

There are very clear reasons why this is essential:

- It ensures that GeCIP investigators can collaborate in academic/NHS partnerships and academic industry partnerships without concern for the intellectual property being generated.
- The ready licensing with the capability to include all inventors offers a fair approach to potential intellectual property.

Genomics England Publication Policy

Genomics England encourages publication. All publications will be on behalf of Genomics England as a banner heading. The usual rules of authorship will apply and all co-authors will be named. The position of authors on all papers will be based on work done with the application of starred authorship to recognise that there may be multiple authors. The Genomics England team will also be co-authors and will typically share key authorship and corresponding author positions reflecting work done. The decision on authorships will usually be made within GeCIP domains and the GeCIP will advise the GeCIP Board whose decision will be final in the event of disagreement. There may be academic users or industry users who are not GeCIP members and pay to access the data. The same approach will apply to authorship and publication with the GeCIP and Board having oversight.

Acknowledgements in all publications will recognise the contribution of the Department of Health, NIHR and any other GeCIP funder with the following form of words.

"Genomics England is a wholly owned company of the Department of Health and this programme was made possible by the National Institute for Health Research, NHS England, Public Health England and Health Education England."

Other GeCIP funders must also be acknowledged.

Genomics
england

Genomics England
Queen Mary University of London
Dawson Hall
Charterhouse Square
London
EC1M 6BQ

www.genomicsengland.co.uk