

# The National Genomics Research Library v5.1

**Amendment to The National  
Genomics Research and  
Healthcare Knowledgebase  
v5**

**1 April 2020**

Issued and approved by the Chief Scientist for Genomics England

© Genomics England Limited 2018



## About this document

*Procedures for the diagnostic genomic laboratory services and clinical care offered by the NHS Genomic Medicine Service sit outside the scope of this document and will be defined by NHS England under established governance and accountability responsibilities.*

*This document sets out the protocol for the National Genomic Research Library (the Library), and is a development of the 100,000 Genomes Project. It reflects the evolution of the ambition to sequence 100,000 Genomes by 2018, to an ongoing and expanding research resource, the Library. This version supercedes version 5 and includes further details regarding research in the Library.*

*Outlined are the principles and standards established since the 100,000 Genomes Project's inception for inclusion of an individual's data and samples in the National Genomic Research Library. It details the patient and clinical benefits, the scientific and transformational objectives, the implementation strategy, as well as the ethical and governance frameworks required for research.*

*Notably, it encompasses the research offering to patients of the NHS-commissioned Genomic Medicine Service (NHS GMS) and participants of other research cohorts.*

*Not covered by this Protocol are the procedures for the diagnostic genomic laboratory services and clinical care offered by the NHS GMS. These are defined by NHS England under established governance and accountability responsibilities. For other cohorts, where direct benefits to a participant do not fall under clinical governance, they will continue to be covered by this research governance.*

*This is a public facing document, providing transparency around the research activities undertaken in the Library.*

*For more information please visit [www.genomicsengland.co.uk](http://www.genomicsengland.co.uk)*

*Authors:*

*Mark Caulfield, Tom Fowler, Fiona Maleady-Crowe, Tom Billins, Freya Boardman-Pretty, Tom Clark, Peter Counter, Clare Craig, Jim Davies, Martin Dennys, Leila Elbahy, Peter Goddard, Jo Hackett, Atul Hatwal, Shirley Henderson, James Holman, Tim Hubbard, Louise Jones, Luke Jostins, Jeanna Mahon-Pearson, Nick Maltby, Clare McConnell, Gil McVean, Nirupa Murugaesu, Anna Need, Katrina Nevin-Ridley, Matthew Parker, Vivienne Parry, Christine Patch, Augusto Rendon, Laura Riley, Richard Scott, Grant Stapleton, Shereen Tadros, Ellen Thomas, Hayden Thomas, Simon Thompson, Carolyn Tregidgo, Alice Tuff-Lacey, Clare Turnbull, Ryan Weir, Katarzyna Witkowska, Carl Smith, Kerrie Woods and Sue Hill (NHS England).*

*With input from Northern Ireland representatives: Shane McKee, Anne Moffatt, Julie McCarroll*

*With input from Welsh representatives: Iris Egner, Julian Sampson, Sharon Whatley*

## About this document

With international and national peer review from: The Genomics England Rare Disease, Cancer and Sequencing and Annotation Working Groups, the Ethics Advisory Committee, the Science Advisory Committee, and NHS England.

NOTE: Health and Social Care (HSC) in Northern Ireland (HSCNI) is the designation of the publicly funded services providing public health and social care services in Northern Ireland. HSC is delivered by a number of organisations including the Public Health Agency (PHA) and a number of health and social care trusts (HSC Trusts). Where the term “NHS” is used to refer to the National Health Service, where applicable this should be regarded as including HSCNI, unless otherwise stated.

NOTE: National Health Service Wales is the designation of the publicly funded services providing public health and social care services in Wales. Wales NHS is delivered by a number of health boards and the Health and Care Research Wales. Where the term “NHS” is used to refer to the National Health Service, where applicable this should be regarded as including NHS Wales unless otherwise stated.

# Contents

<b>1</b>	<b>Summary and background</b> .....	<b>6</b>
1.1	History of the Project .....	6
1.2	Transition from the 100,000 Genomes Project to the National Genomic Research Library .....	7
1.3	The NHS Genomic Medicine Service (NHS GMS) - 2018 onwards.....	7
1.4	The National Genomic Research Library .....	8
1.4.1	“What is the National Genomic Research Library?” .....	9
1.4.2	“What is the research environment?” .....	9
1.5	Aims of the Library .....	9
1.6	Genomics England and Partners .....	10
1.7	Genomic testing including whole genome sequencing (WGS).....	11
1.8	Key differences from Protocol versions 5 and 5(a).....	11
1.9	NHS GMS-related impacts to potential participants .....	11
<b>2</b>	<b>The structure of the Programme</b> .....	<b>13</b>
2.1	Beyond 100,000 Genomes (2018 onwards) .....	13
2.1.1	The Rare Disease Programme.....	13
2.1.2	The Cancer Programme.....	17
2.2	The Genomics England Infectious Disease partnership with Public Health England.....	22
2.3	Sample size considerations for research in rare disease and cancer.....	22
2.3.1	Declaring significance .....	26
2.3.2	Variation in gene size and mutation rate.....	26
2.3.3	Genetic architecture, and discovering genes vs solving diseases .....	27
<b>3</b>	<b>Opportunities to extend beyond the primary phenotypes</b> .....	<b>28</b>
3.1	Opportunities to extend beyond rare diseases, cancer and infection.....	28
3.1.1	Lifelong electronic health record linkage.....	28
3.1.2	NHS and public health registry data linkage.....	28
3.2	Opportunities beyond whole genome sequencing .....	29
3.3	Opportunities for recruitment beyond the NHS GMS/clinical services.....	29
<b>4</b>	<b>The Genomics England Clinical Interpretation Partnership (GeCIP)</b> .....	<b>32</b>
4.1	The route by which researchers add value to GeCIP .....	33
4.2	The route for those who contribute patients and data from the NHS to add value to GeCIP .....	34
4.3	Establishing and working within the Clinical Interpretation Partnership.....	34
4.4	Accessing data as a Genomics England GeCIP Member .....	37
4.5	The commitment of all data generated to the Genomics England Knowledge Base .....	38
4.6	GeCIP contribution to interpretation, validation and regulation of findings.....	39
4.7	Overarching GeCIP management and governance.....	39
<b>5</b>	<b>Commercial Partnerships</b> .....	<b>40</b>

5.1	Overview and aims of Genomics England commercial activities.....	40
5.2	Commercial research aims.....	40
5.3	Ways of working with commercial partners .....	41
5.3.1	The Discovery Forum – access to a network (Associate membership) .....	41
5.3.2	The Discovery Forum – access to the Library (Full membership) .....	42
5.3.3	Strategic partnerships.....	44
5.4	Commercial research approval framework .....	44
<b>6</b>	<b>Feedback for participants .....</b>	<b>45</b>
6.1	Enhanced interpretation and handling of findings from the research environment .....	45
6.2	Policy for main findings, secondary findings, and incidental findings, and feeding back to clinicians and patients .....	45
<b>7</b>	<b>Programme delivery and management .....</b>	<b>51</b>
7.1	Project management.....	51
<b>8</b>	<b>Informatics architecture .....</b>	<b>52</b>
8.1	Information governance and security .....	52
8.2	Data access, sharing and acceptable uses of information from participants .....	54
8.3	Partnerships to ensure patient recruitment, sample and data acquisition .....	54
8.4	NHS England requirements and specifications – the Library and the NHS GMS .....	55
8.5	NHS England requirements and specifications – Genomic Medicine Service .....	55
8.5.1	Sample and DNA acquisition and logistics for NHS Genomic Medicine Service .....	55
8.5.2	Labelling of samples.....	56
8.5.3	DNA delivery to sequencer (logistics).....	56
8.5.4	Sample and DNA acquisition and logistics for WGS .....	56
8.5.5	Key Performance Indicators (KPIs).....	57
8.5.6	Data from sample acquisition site.....	57
8.5.7	Data from national NHS and other sources .....	57
<b>9</b>	<b>Genomics England data access and acceptable uses.....</b>	<b>58</b>
9.1	Users of the Library data .....	58
9.1.1	Patient access.....	58
9.1.2	NHS healthcare teams.....	58
9.1.3	Academic researchers and trainees.....	58
9.1.4	Commercial researchers and Industry engagement .....	58
9.2	Data access and sharing for the Library .....	58
9.2.1	Data sharing requirement.....	59
9.2.2	Access to service data .....	60
9.2.3	Corporate data sharing.....	60
9.3	Acceptable research uses .....	61
9.3.1	Types of purposes.....	61
9.3.2	Uses that the Library will decline .....	62
<b>10</b>	<b>Genomics England data ownership and intellectual property .....</b>	<b>63</b>

10.1	Genomics England publication policy .....	63
<b>11</b>	<b>Consent .....</b>	<b>64</b>
11.1	The process for consent in the Library .....	64
11.1.1	The consent form and patient information literature .....	66
11.1.2	Giving consent face-to-face .....	67
11.1.3	Giving consent at a requested home visit.....	68
11.1.4	Giving consent by post or email.....	68
11.1.5	Participants aged 16-18 years.....	68
11.1.6	Participants under the age of 16 years.....	69
11.1.7	Seeking consent where participation relates to a health emergency .....	69
11.1.8	Seeking consent to use samples and information regarding a person who is deceased ..	74
11.2	The Process for withdrawing from the Library .....	78
11.3	Personal insurance and the participants in the Library .....	82
<b>12</b>	<b>Communications, stakeholder engagement and patient and public involvement.....</b>	<b>83</b>
12.1	Communications strategy and governance arrangements .....	83
12.2	Communications and stakeholder engagement plan .....	83
12.2.1	The Genomics Conversation .....	83
12.2.2	The Public Dialogue.....	84
12.3	Genomics England research projects – seeking the views of patients and the public .....	84
12.4	The National Participant Panel .....	85
<b>13</b>	<b>Appendices .....</b>	<b>86</b>
	Appendix 1   List of Abbreviations .....	86
	Appendix 2   WGS Sample Requirements .....	87
	Appendix 3   Genomics England Governance and Committees.....	90
	Appendix 5   Letter of Insurance .....	92
	Appendix 6   Consent Process Flows.....	94

# 1 Summary and background

# 1 Summary and background

## 1.1 History of the Project

In December 2012, the then UK Prime Minister (David Cameron) announced as part of the UK Government's Life Sciences Strategy, a programme of whole genome sequencing (WGS) known as the '100,000 Genomes Project' (the Project).

The principal objective of the Project was to sequence 100,000 genomes from patients with cancer, rare disorders, and infectious disease. In addition, the Project's aim is to link the sequenced data to a standardised, extensible account of diagnosis, treatment, and outcomes.

The Project was designed to produce new capability and capacity for the application of whole genome sequencing application in routine clinical care, to transform both the NHS and clinical genomics research.

In 2013, the Chief Medical Officer for England (Professor Dame Sally Davies) established a Strategic Priorities Working Group and recommended three main areas where the introduction of genomic technology would have the greatest benefit for patient health:

- i. Rare and inherited diseases
- ii. Certain cancers
- iii. Infectious diseases

Established by the Department of Health and Social Care, 'Genomics England' was announced as the organisation responsible for delivering the Project by the Secretary of State for Health (Jeremy Hunt) during the NHS 65<sup>th</sup> Anniversary Celebrations in July 2013. The Medical Research Council and devolved nation governments provided funding for the Project to be extended to Scotland, Wales and Northern Ireland.

To identify and enrol participants and manage return of their results, NHS England sponsored and established NHS Genomic Medicine Centres (NHS GMCs). These regionally distributed centres harnessed the capability and capacity of the NHS across England between 2015 and 2018. Their role is now evolving to support the NHS GMS with specific responsibilities around consent pathways, education and training, and implementation.

The Project quickly established a secure infrastructure for the protection and analysis of all clinical and genomic data. This environment was made available for approved academic and industrial research purposes, including contributing members of clinical organisations from the NHS.<sup>i</sup>

The Chief Medical Officer's July 2017 report, 'Generation Genome,' recommended the establishment of a National Genomics Board to oversee the integrated development of genomic services and research in England and across the UK. In preparation for the next phase of the Project, and the application of whole genome sequencing as part of a commissioned NHS service in England, a number of NHS transition groups were set up by NHS England to identify appropriate commissioning strategies for genomic testing, including both WGS and non-WGS genomic testing.

## **1.2 Transition from the 100,000 Genomes Project to the National Genomic Research Library**

In 2018, the Project completed the sequencing target of 100,000 genomes, however analysis and return of results is on an ongoing basis. The Project is transitioning to where the main source of recruitment, sample and data acquisition, and mechanism for return of results will be via the NHS Genomic Medicine Service (NHS GMS).

This will offer patients the dual opportunity of routine clinical care alongside a choice to participate in research spanning across all genomic tests within the NHS, starting with whole genomes. The National Genomic Research Library (the Library) builds on the research infrastructure created by Genomics England for the 100,000 Genomes Project and provides a national standardised genomic research resource.

In October 2018 the Secretary of State for Health and Social Care Matt Hancock said “I’m proud to announce we are expanding our 100,000 Genomes Project so that one million whole genomes will now be sequenced by the NHS and the UK Biobank.

I’m incredibly excited about the potential for this type of technology to improve the diagnosis and treatment for patients to help people live longer, healthier lives – a vital part of our long-term plan for the NHS.

Today’s commitments form part of our bold aspiration to sequence five million genomes in the UK, using ground-breaking technology to do this within an unprecedented five-year period.”

Genomics England has been charged with developing the plan for delivery for this ambition and it is expected that part of this will be delivered under the governance of this Protocol.

## **1.3 The NHS Genomic Medicine Service (NHS GMS) - 2018 onwards**

The NHS has a history of providing genetic and genomic services in a clinical setting. This includes regional genetic laboratories, clinical genetic and cancer services and cancer molecular testing.

Development of the NHS GMS builds on this history, supported by the evidence generated by the 100,000 Genomes Project but extends to other genomic and multi-omic testing.

The future NHS GMS will be led and commissioned by NHS England and will consist of:

- NHS Genomic Medicine Centres (initially established by NHS England as part of the existing 100,000 Genomes Project infrastructure) which may become National Genomic Medicine Alliances
- NHS Genomic Laboratory Hubs (GLHs) that will work as part of a National Genomic testing service. The provisions in this service will be determined by a national genomic test directory that outlines the testing strategies and technology to be employed for rare and inherited disease, cancer and other defined conditions/ applications.
- Clinical Genetic Services
- Cancer services using genomic analysis to guide treatment

In addition to the pseudonymised data held within the research infrastructure created for the 100,000 Genomes Project, a new clinical source of NHS GMS-commissioned WGS and all other

genomic data will be added for those who consent to research. Appropriate governance will determine where data from the clinical service will be made available for research, dependent on patient choice for use of their data and samples.

The NHS GMS will be considered as the primary source of samples, clinical and research data. Other clinical services and research cohorts however, will comprise additional data sources to be used as part of the evolving data infrastructure made available for research in the Library. It is expected that this approach will potentially benefit all participants by allowing greater leverage for genomic data to improve diagnostics.

Participants will continue to receive the same research offer outlined in the 100,000 Genomes Project, including opportunities for feedback to patients via the NHS, recall for research (by the NHS or Genomics England), international data access via the Genomics England Data Centre and opportunities for industry to develop new diagnostics, analytics or therapies as previously approved by the Research Ethics Committee. The research offer for non-NHS GMS cohorts will be able to avail of the full research offer.

The NHS GMS is a route for the provision of data and samples into the Library and similar arrangements will be open to devolved nations and appropriate cohorts.

In the NHS GMS, Genomics England will procure under a partnership with NHS England the NHS WGS requirement as outlined in the National Genomic Test Directory. The NHS will undertake all other genomic and multi-omic testing within the new GLHs. To allow us to undertake research, all tests and clinical data will be concentrated in a single NHS dataset as part of routine care (not part of this protocol). Each patient will be offered the opportunity to consent to take part in research. To fully realise the potential of data from the NHS GMS, the Library will concentrate results from all genomic testing - from a single targeted test to a WGS – to future commissioned omics and other technologies in the research dataset.

#### **1.4 The National Genomic Research Library**

The National Genomic Research Library (the Library) described within this Protocol refers to participants recruited via mechanisms set up for:

- the NHS GMS;
- the 100,000 Genomes Project; and
- other cohorts

A central purpose of these initiatives is to achieve the aims referenced below by bringing together data sets in the Library. The Library encompasses all data collected for research purposes, including data collected under earlier versions of this Protocol. This is to reflect its growing role as a central resource of genomic data. All research data from these cohorts will be stored in the Library.

All programmes included under this Protocol have a commitment to delivering clinical translation and patient benefit. This may be done completely under the research governance outlined here e.g. the 100,000 Genomes Project, or achieved through a hybrid model with external clinical governance e.g. participants recruited via the NHS GMS.

#### 1.4.1 “What is the National Genomic Research Library?”

In the NHS Genomic Medicine Service (GMS), patients will be asked if they want to donate their sample (blood / saliva / tissue, etc.), genome sequence and health data for research. This offer is an integral part of clinical care. In cohorts, the Library provides an equivalent function.

- The Library is a secure national resource of genomic, health data and samples managed by Genomics England in partnership with NHS England

Being able to compare all patient data in one place provides researchers with an opportunity to better understand diseases, develop new treatments and can lead to new discoveries.

Approved researchers can use the samples and data to study diseases and look for new treatments.

#### 1.4.2 “What is the research environment?”

- It is a comprehensive resource that allows researchers to securely access pseudonymised genomic and other associated health data within the Library

It provides a trusted environment to allow researchers to carry out analysis without allowing the download of individual data.

### 1.5 Aims of the Library

The research aims of the Library continue but have evolved to encompass a research opportunity for patients across all genomic tests undertaken by the NHS GMS. These aims are:

- **Patient benefit**
  - Provide clinical diagnoses with research opportunities to participants across all genomic and multi-omic testing in the NHS
  - Discovery of new causes of disease, offer tailored choice of therapies to create the best outcomes, and prime new or more effective treatments for NHS patients.
- **Accelerate the uptake of genomic medicine in the NHS and other healthcare systems**
  - Work with NHS England and other partners to deliver a scalable NHS GMS including all genomic tests and an informatics platform to enable these services to be made widely available for NHS patients.
  - Work with Devolved Nation administrations, where appropriate, to achieve the same aims of the NHS GMS.
  - Through the international coalition of research intellects known as the Genomics England Clinical Interpretation Partnership (GeCIP) and the Discovery Forum, the framework for Genomics England to work with Industry:
    - Create a mechanism for research to continually improve the accuracy and reliability of information fed back to patients
    - Add to knowledge of the genetic basis of disease
    - Increase opportunities for clinical trials
    - Build the evidence base to accelerate the introduction of new technologies into healthcare

- **Increase public knowledge and support for genomic medicine**
  - Deliver an ethical and transparent programme
  - Retain patient and public trust and confidence
  - Work with a range of partners to increase knowledge of genomics
  
- **New scientific insights and discoveries**
  - Create a database of all genomic data, including all genomic and omics tests. This will comprise of, but not be limited to: single gene tests, targeted gene sequencing, coding region sequencing (1% of the genome that codes for proteins –the exome), whole genome sequences and other multi-omic tests. This dataset will be based on the offer of research participation to all individuals being offered a genomic test in the NHS or other cohorts external to the NHS.
  - Information linked to continually-updated, long-term patient health and personal information to aid analysis by researchers.
  
- **Stimulate and enhance UK industry and investment**
  - Provide access to this unique research data resource to industry for the purpose of developing new knowledge, methods of analysis, medicines, diagnostics and devices.

## 1.6 Genomics England and Partners

The Department of Health and Social Care has established Genomics England as a wholly owned, limited company to provide the informatics infrastructure, bioinformatics pipeline and research functions of the evolving Library. Here we focus on the research dataset derived from those who choose to provide consent to the research opportunity from the new NHS GMS, and those who provide their consent for their data to be donated to the Library.

NHS England will act as one of the major data sources for the future research function of the Library. In addition, NHS England will work in partnership with Genomics England to establish the Genomic Medicine Service infrastructure that will hold and process the data and results.

Other partners include:

- Public Health England
- NHS Digital
- Health Education England
- NHS Trusts
- Northern Ireland Department of Health
- HSCNI organisations
- Cardiff & Vale University Health Board (UHB)
- Welsh UHBS
- Welsh Government and Cardiff University (CU)
- Wellcome Trust
- The Wellcome Trust Sanger Institute
- Illumina

Alongside this there will be international partnerships with other nations and health providers e.g. the French genomic initiative [Médecine Génomique 2025](#); [Qatar Genome Programme](#); and the [Million European Genome Alliance](#).

### 1.7 Genomic testing including whole genome sequencing (WGS)

Next generation sequencing using massively parallel sequencing has advanced understanding of the genetic architecture of disease.

Use of WGS offers the opportunity for:

- Detection of variants outside exons and a better representation of those inside exons.
- Improved detection of variants on intron/exon boundaries.
- Detection of an enhanced repertoire of variation including insertions and deletions (indels), copy number variants, and structural changes across the genome.
- Identification of variation that may encode other diseases.

Genomics England will continue to explore providers of WGS, and new technologies, to ensure that the most leading-edge technology can be considered. Where a new technology shows promise of patient benefit, it may be introduced to the Library with non-NHS funding to help build a case for clinical commissioning. A similar approach will also be taken around methods that have the potential to improve sample handling, making the implementation of WGS pathways more feasible in healthcare.

### 1.8 Key differences from Protocol versions 5 and 5.1

- New consent materials have been produced to record the patient's choice in using their data and samples for research when undergoing a clinical genomic test. These have been iterated in response to the developing clinical patient choice materials to enable a hybrid approach to clinical and research consent. This process has included substantial consultation with stakeholders, including clinicians, patient representatives and the 100,000 Genomes Project Participant Panel.
- An electronic method of capturing patient choice has been embedded in the National Genomic Informatics Service.
- Paper documents will continue to be provided to ensure equity of access.
- Greater detail of the Genomics England Discovery Forum.
- Greater clarity on the role of cohorts wholly under the governance of the Protocol for patient benefit.

### 1.9 NHS GMS-related impacts to potential participants

- Potential participants of the research component of the NHS GMS are NHS patients who are undergoing genomic testing based on clinical indications as part of standard care, and as outlined in the National Genomic Test Directory that will be refreshed annually.
- For patients within the NHS GMS, the choice for allowing routinely collected data and samples to be used in research will be offered alongside clinical consent for the test.
- Consent will be sought by the patient's clinical team at the time when the clinical test is being discussed, where appropriate.

- If extra samples and data are required for additional studies, approval will be sought from the appropriate governance forums in line with the Genomics England Access to Data and Samples Policy.
- Where recruitment is in parallel or in partnership with clinical trials, this will only be done where appropriate approvals for interventional aspects are in place. It is recognised that a major benefit for participants is identification of opportunities in relevant trials or layering of genomic information on to existing trials.
- Other participants of the Library will be from existing cohorts. More details about the governance of these cohorts is provided in Section 3.3.
- Results from the Library will not be received by the participant directly, but via the NHS GMS (as in the 100,000 Genomes Project), unless an alternative mechanism is in place, for example in cohorts.

## **2 The structure of the Programme**

## 2 The structure of the Programme

### 2.1 Beyond 100,000 Genomes (2018 onwards)

The Library consists of two main areas of focus based on the genomic or multi-omic testing strategy implemented via the NHS GMS. In addition, to ensure patient opportunities are maximised other disorders or population cohorts may be included (subject to consent):

- i. Rare diseases and other diseases subject to germline genomic testing
- ii. Cancers subject to somatic genomic testing (and accompanying germline testing where indicated)

This may develop over time as new genomic and multi-omic testing strategies are commissioned within the NHS GMS.

Within both programme areas, there are some common aspects :

- Main data and sample source will usually be the NHS GMS
- There will be potential for diagnostic benefit or tailored choice of therapy
- Opportunity to take part in further research
- Inclusion within the research infrastructure based on patient choice and consent
- Co-ordination of the clinical service through a single informatics and data service
- Adherence to the Library governance
- Inclusion of research cohorts sourced outside the NHS GMS, where appropriate
- Linkage to longitudinal data sources
- Collaboration of researchers to drive up clinical utility
- The requirement for clinical data to allow the interpretation of genomic data
- Cross-referencing of the genomic dataset to health datasets to provide improved individual patient diagnosis

As provision of data is based on real-world availability from a clinical service, the following is indicative of the approach to be taken but may be restricted by constraints on the availability of data and samples.

#### 2.1.1 The Rare Disease Programme

The goals of the Rare Disease Programme are:

- To increase discovery of pathogenic variants for rare disease.
- To add value with additional biological insights that build confidence in putative pathogenic variants.
- To enhance the clinical interpretation of WGS in rare disease.
- To develop a programme of functional multi-omics pathways, specifically transcriptomics, epigenetics, micro RNAs and biomarkers.
- To return findings to the NHS for feedback to patients.
- To create a unique dataset for rare diseases that may enable therapeutic innovation.

#### Background to rare diseases

There are between 6,000 and 8,000 known rare diseases worldwide. Each disease or syndrome may affect less than 0.1 percent of the UK's population, but cumulatively they affect the lives of 3 million people and are associated with substantial morbidity and mortality. Only 50 percent of the known rare diseases have an existing molecular diagnosis. Of these, 85 percent are associated with a single gene defect, although there may be modifiers elsewhere in the genome that have implications for treatment and outcomes. Recent research has shown that WGS can augment gene discovery by between 25 to 40 percent across a range of rare disease phenotypes (based on the Whole Genome Sequencing 500 study and the 100,000 Genomes Project early findings). Gene discovery in the Library will create significant opportunities for scientific innovation through routine service, our focus on residual unmet need, and our emphasis upon national and international collaborations.<sup>ii,iii,iv,v,vi,vii</sup>

### **National and international collaborations in rare diseases**

Genomics England will work in collaboration with all important rare disease initiatives. In particular, we will actively support the implementation of the UK Rare Disease Strategy. Some key partners include the Deciphering Developmental Disorders (DDD) programme, the NIHR Translational Research Collaboration in Rare Disease (TRC-RD), the Whole Genome Sequencing 500 programme and the NIHR Bioresource in Rare Diseases (BR-RD).

Building on work already undertaken by the Project, this will facilitate the generation of a national data resource of all genomic data, with a focus on whole genome sequence but including all genomic testing. The Library will consist of families with rare diseases, based on family structures appropriate to the provision of the clinical service. More information on family structure guidance is available at [www.genomicsengland.co.uk](http://www.genomicsengland.co.uk). This dataset will be enriched by other studies, including but not limited to: BR-RD, TRC-RD, and DDD through the Genomics England Clinical Interpretation Partnership (GeCIP). This will harness the strength of current UK rare disease programmes and will advance current understanding of rare disease mechanisms. It may also impact on common diseases that share similar phenotypes. It will also offer opportunities for biomarker, clinical, and interventional studies through industrial partnerships. This initiative is particularly timely as Public Health England have initiated a national registry for rare diseases and we are partnering with them to ensure the optimal outcome for patients.

The Library is expected to enable genomically-driven reclassification of rare diseases leading to opportunities to recall patients for deeper phenotyping through TRC-RD. These data will pave the way for functional characterisation of findings thereby adding further value to datasets, improving diagnostic utility and possibly identifying new targets and therapies.

### **The NHS England Genomic Test Directory for rare disease areas**

Patients having any genomic test within the NHS will be offered the opportunity to participate in the Library.

The NHS England Genomic Test Directory will be maintained by the Genomics Unit at NHS England with expert advice from their advisory working group(s) and clinical panel. Where non-WGS testing provided under the Genomic Test Directory has not yielded a diagnosis, there is the potential to

offer whole genome sequencing within this infrastructure, should alternative sources of payment exist.

### **Patient engagement and phenotyping strategy for rare diseases**

Genomics England will work in partnership with NHS England, Wales and DoH NI (operating through its arm's length body BHSCT) to engage, recruit, and characterise patients with specific rare diseases, with residual unmet diagnostic need. A managed network of NHS Genomic Medicine Centres (GMCs) (including NIGMC and the Wales GMC) has been established across England, Wales and Northern Ireland. It is expected that the English NHS GMCs will work within a performance managed network and within the evolving commissioning structure in the new Genomic Medicine Service. Separate arrangements and agreements exist for the establishment, operation and governance of the NIGMC and Wales GMC, based on similar contractual arrangements as English NHS GMCs. It is hoped Scotland will also participate.

NHS healthcare teams will be responsible for offering research participation to patients and family members following agreed Standard Operating Procedures and appropriate consent. They will also be responsible for the provision and recording of core clinical phenotypic and demographic data on patients and relevant family members.

Consent will follow standard requirements for clinical genomic testing, including data requirements. If identified, this data will be collected through a national genomic informatics service as part of a commissioned service. It is intended that this will also include the capacity to digitally record participants' consent decisions. However, paper copies of consent materials will also exist to ensure equity of access.

### **Clinical phenotyping in rare diseases**

For testing carried out under the Genomic Test Directory, clinical data collected will be appropriate to the requested genomic test.

For WGS, within each disease area, a core set of characteristics and data points has been defined in consultation with clinicians and researchers with relevant expertise. In most cases, this will include a number of statements made using a phenotyping ontology. Although the Human Phenotype Ontology (HPO) is currently being used, other ontologies may also be employed. Data will continue to be recorded using controlled, clinical terminologies wherever possible for consistency.

The NHS number / Health and Care (H&C) number (for Northern Ireland participants) (and other core demographic data) will be recorded for each participant, including the biological relatives recruited as part of a trio or duo for patients with rare disease. A family tree and history of diseases will be recorded for parents and all offspring. The NHS/H&C number will be used to identify and link to relevant health and social care records, providing additional context for the core phenotypic data.

A secure, web-based, information system will be provided for the collection of this data, removing the need for any additional, bespoke development within participating NHS organisations. An electronic facility will be provided for transmission of the same data from existing information

systems. Where organisations have developed their own capacity for capturing and managing the same data, to the same standards, there will be no requirement for additional data entry.

Each rare disease may have a different set of phenotypic characteristics ranging across clinical features, laboratory tests and imaging. We will use published literature, ontologies and expert advice to help with development of systems to capture these characteristics from existing health records in increasing numbers of clinical contexts. In cancer, we collected standardised clinical datasets that may include tumour specific features or staging related to tumour type which will typically be provided by routine data feeds to national cancer registries held by PHE.

### **Genomic testing strategy and quality assurance (excluding whole genome sequencing).**

For genomic testing performed by the GLHs, standards will be set across the NHS by NHS England and informed by Genomics England.

### **Whole genome sequencing strategy and quality assurance for rare diseases**

Our initial approach for the main programme was adapted from experience gained in the WGS500 study as well as advice from UK and international experts and information from technology providers. For rare diseases, Genomics England has sequenced whole genomes in all members of the family trio (proband, mother, and father) for most disorders, where available, but has adapted its processes to sequence other family structures as information emerged from the pilot programmes and other efforts worldwide. Family structures in the future will be determined by the route of sample collection, but trios will continue to be sequenced wherever possible.

DNA sequencing for the main programme has been contracted through Illumina. This was informed by the results of an earlier sequence evaluation exercise. Although there is currently only one sequence provider in the programme, we expect that multiple instruments and chemistries will be applied throughout the lifetime of the Library as the technology matures. Sequencing standards will therefore be maintained and enhanced as appropriate through the requirement to meet platform independent metrics (as detailed in the Sample Handling Guidance, latest version to be found at <https://www.genomicsengland.co.uk/information-for-gmc-staff/sample-handling-guidance/>).

The data generated will be transferred encrypted via a dedicated 10G link to the Genomics England Data Centre. For any genomes arriving externally the method will be reviewed and agreed.

Genomic testing supplied under the NHS GMS will be provided as agreed with NHS England.

Genomics England will capture quality metrics for WGS per run and routinely check for the presence of batch effects at the run, flow cell, and lane level (as applicable to the technology) to ensure that sequence quality is maintained and to aid its improvement over time.

### **Alignment and identification of sequence variation**

It is expected that as sequencing technologies evolve, sequencing will continue to align with the clinical service. As part of the sequencing contract, providers will supply sequence data aligned to the reference genome (currently GRCh38) and in a standard file format (currently BAM). They will also identify sequence variants and report them in a standard file format (currently VCF). It is expected that the mapping and calling pipelines will report single nucleotide polymorphisms (SNPs),

short indels (<50bp), copy number variants (CNVs) and structural variants (SVs). It is also expected that our pipelines will incorporate processes to jointly call across family data. This will ensure that missing sequence data is distinguished from that identical to the reference and that putative haplotypes are generated to the extent possible, depending on the underlying sequence fragment lengths. Pipelines should also identify regions of the genome with inadequate coverage for reliable variant calling. Genomics England will define clear Standard Operating Procedures in partnership with NHS England for the WGS NHS-commissioned service and will revise and inform the NHS of any changes as these evolve over the life of the Library and its legacy.

Genomics England will routinely assess the quality of the mapping and variant calling provided and compare this against alternative algorithms to ensure the data has been processed with the best available pipelines.

### **Interpretation of sequence variation**

The interpretation of sequence variation will move to a clinical governance framework and is expected to work in the following manner. To provide timely analysis to support interpretation against specified and clinically agreed turnaround times, we will have a core sequence interpretation service. This may consist of internally-run filtering processes (e.g. a tiering algorithm) and inputs from external interpretation providers through competitive tendering. These are known as decision support providers (DSPs). These annotations will also be available to researchers within the Genomics England Data Centre.

In the NHS GMS as part of clinical service, the output from this pipeline will be presented to clinical scientists in the NHS GLHs in the form of a web-based decision support tool; enabling the final manual interpretation and reporting step that will be carried out within the network of NHS GLHs. WGS entering the Genomics England Data Centre from other sources may also be run through this pipeline and returned by a similar route, where appropriate.

Pipelines will be expected to evaluate all modes of inheritance. They will also identify long regions of homozygosity (LROHs), particularly where consanguinity is suspected, and in such regions, identify potentially pathogenic homozygous variants. In terms of identifying variants most likely to have functional consequences, it is anticipated that pipelines will evaluate variants using various prediction tools (e.g. Ensembl Variant Effect Predictor, Sift, and Polyphen), as well as checking against databases of known pathogenic variants (e.g. ClinVar). Variant annotation will be updated regularly, and snapshots of previous annotation versions will be stored and versioned.

### **2.1.2 The Cancer Programme**

#### **Goals of the cancer programme**

The goals of the cancer whole genome sequencing programme are to:

- Use WGS to identify novel driver mutations for cancer and to understand its evolutionary genetic architecture through primary and secondary malignant disease (by multiple biopsy and WGS).

- Partner stratified healthcare programmes and outcome studies with patients from the NHS in England, to enable understanding of WGS benefits in defining predictors of therapeutic response to cancer therapies.
- To use multi-omic approaches including transcriptomics, proteomics and epigenetics to offer additional biological insights into cancer.
- To utilise WGS to identify new pathways for cancer therapies and improved diagnostic characterisation.

### **The inclusion of specific cancers**

Cancer is fundamentally a genetic disorder where mutations (including copy number aberrations, indels (insertion/deletion variants), complex rearrangements, and non-synonymous substitutions) lead to uncontrolled cellular proliferation. The clinical impact of sequencing technologies has enabled precise definitions of disease, uncovered mechanistic insights into pathogenesis, and identified therapeutic targets based on genetic variation or aberration. Additionally, sequencing approaches have catalogued the complex evolutionary changes that occur in an individual's cancer under the effects of treatment and time. This has demonstrated that there are both expanded clonal populations and the existence of low frequency sub-clones, each with specific genomic architecture.<sup>viii, ix, x, xi, xii, xiii</sup>

Meta-analysis across cancer types has confirmed the importance of at least 200 key genes in driving cancer. However, focusing only on these genes with targeted re-sequencing will be insufficient to significantly impact on the majority of individuals with cancer. Overall response rates to targeted agents with 'actionable' mutations have been relatively modest. The differences in response between patients with mutations in the same gene suggest complex interactions between identified activated pathways and the genomic environment in which they occur. Moreover, power calculations and saturation analysis suggest that many more cancer gene regions remain to be found (see section 2.3 sample size requirements). In many cancers, the 'actionable' gene may not be the dominant driver. Furthermore, recurrent oncogenic mutations are rare and some of the putative driver mutations are caused by copy number aberrations or complex rearrangements that cannot be globally ascertained with gene panels, thus WGS or orthogonal tests are required.

The Library will continue to learn from and collaborate with the Cancer Genome Atlas (TCGA) project and the broader International Cancer Genome Consortium (ICGC), who are producing an inventory of genomic, transcriptomic and epigenomic changes in a wide range of different tumour types.

A list of NHS England-approved cancer genomic tests will be provided in the National Genomic Test Directory. Where there is external funding for the investigation of cancers beyond the NHS GMS (e.g. via the Cancer Fund), the principles for inclusion are described below.

### **The NHS England Genomic Test Directory for the cancer programme**

NHS England evaluated each cancer type for inclusion in the NHS commissioned service. Every proposed disease was assessed by experts in specific transition groups, who then made recommendations. As above, the opportunity exists to add new cancers either to the NHS service or to the research programme following expert advice.

A small number of cancer subtypes are eligible for WGS within the first version of the National Genomic Test Directory. These indications were selected because they fulfilled one of the following criteria:

- For any individual sample, a large range of separate genomic tests are already performed as standard of care that could (pending appropriate validation) be replaced by WGS (e.g. acute leukaemia).
- They represent a broad group of relatively rare cancers with disparate histologies, where although for any individual tumour only a small number of genomic tests may be required- collectively as a sample group a large range of different genomic tests (infrequently performed) is required to be available which could (pending appropriate validation) be replaced by WGS (e.g. paediatric tumours).

Work is ongoing to build a case for commissioning WGS in other cancer types and for now, other forms of genomic tests are specified through the Test Directory.

Comparable to rare diseases, consent will also follow standard requirements for clinical genomic testing, including data requirements. If identified, this data will be collected through a national genomic informatics service as part of an NHS England-commissioned service. It is intended that this will also include the capacity to digitally record participants' consent decisions. However, paper copies of consent materials will also exist to ensure equity of access.

**Inclusion criteria for cancers:**

- A cancer from the NHS England Genomics Test Directory approved list or a Genomics England approved nomination by the NHS, academics, or industry.
- Capability of the NHS in England/HSCNI to recruit, supply samples, and upload data from patients.
- Access to appropriate high-quality DNA from both tumour and germline samples, derived in the overwhelming majority of cases from fresh tissue preparation for WGS, according to the Sample Handling Guidance provided by Genomics England and/or NHS England.
- Existence of consent and upload of the phenotypic data required for diagnostic testing.
- Timelines for recruitment that are in line with Genomics England Standard Operating Procedures.
- The existence of sufficient resources to perform WGS within the Library, or an independent source of funds to complete sequencing in all planned recruits e.g. from a GeCIP funder.

**Exclusion criteria for cancers:**

- The absence of consent, unwillingness to give consent or participate in all aspects of the Library (excludes opt-out of feedback).
- Inadequate phenotyping, or failure to upload phenotypes to the Genomics England Data Centre.
- No access to matched tumour and normal DNA samples.
- Patients not under the care of the NHS in England or HSCNI.\*

- DNA of insufficient quality for WGS, or failure to adhere to the Sample Handling Guidance or other SOPs provided by Genomics England or endorsed/coproduced by Genomics England and issued as part of a clinical service (e.g. that for the NHS GMS by NHS England).

### **Phenotyping of patients participating in the cancer programme**

For each disease area, a set of phenotypic characteristics or data points will be defined in consultation with clinicians and researchers with relevant expertise. This core phenotypic data set will be aligned with the needs for clinical reporting. Additional phenotypic data sets will be aligned data sets for cancer such as, the Cancer Outcomes and Services Dataset, the Systemic Anti-Cancer Therapy Dataset, and the Radiotherapy Data Set. These will also be aligned with the clinical audit data sets collected for specific cancers: colorectal, lung and prostate, to avoid duplication of effort. Data will be recorded using controlled clinical terminologies, and structured ontologies wherever possible.

The NHS/H&C number (and other core, demographic data) will be recorded for each participant. This information will be used to identify and link relevant health and social care records, providing additional context for the core phenotypic data. It will be used also to identify and link registry and audit data, where this is available.

The structures being built for the National Genomic Informatics Service will be the system for both collecting and providing the relevant cancer data for the NHS GMS. This will also be used as a mechanism for acquiring data for cancers included but outside of the NHS GMS.

There is an expectation that the phenotypic data will be provided within clinically-appropriate timelines in order to enable the return of results. Samples will not be sent for sequencing until genomic and somatic DNA of sufficient quality and a minimum data set, have been supplied. Whole genome sequencing will not be undertaken until the dataset, and QA processes are submitted with sample.

### **DNA quality assurance from cancer samples**

National standards for DNA quality for all genomic tests can be found in the sampling handling guidance ([www.genomicsengland.co.uk](http://www.genomicsengland.co.uk)). This may change as the NHS GMS develops. As part of the NHS GMS only fresh frozen samples (FF) will be eligible for WGS except in exceptional circumstances or where there are developments in the technology that allow other approaches.

### **Additional samples for multi-omic functional studies**

We will seek additional funding for appropriate RNA, plasma, and serum allowing the acquisition of expression, methylation, epigenetic data and serum or plasma biomarkers pertaining to the same samples used for WGS. These samples will be collected and stored at the central biorepository or NHS biorepositories and will add the value of functional analysis to the WGS. We have a significant opportunity to apply our expertise in detecting temporal changes in cancer genomes (in cell-free tumour DNA in plasma) to develop non-invasive 'liquid biopsy' based on WGS or broad targeted

panels.<sup>xiv</sup> This could provide a cost-effective means of disease monitoring, to enable innovation in trial design and encourage industrial investment in UK clinical research.

### **Whole genome sequencing strategy and quality assurance for cancer**

Our current approach is being adapted with advice from UK and international experts and information from technology providers. Moving forward, Genomics England will continue to sequence germline samples using the same criteria as for rare diseases samples (see above).

For tumour DNA, we will continue to sequence at least 75X coverage. However, in some cancers, greater depth of coverage may be required and we will continue to modify this as necessary.

WGS will be continue to be performed using the same technology and QC approaches outlined for the rare diseases programme. In addition, we will continue to monitor the frequency of single nucleotide variations (SNVs) potentially caused by DNA alterations created during the FFPE process, (such as G to A, or A to G transitions) due to PCR amplification.

### **Identification of sequence variation**

For the cancer programme, sequence providers will continue to carry out initial steps in read mapping and variant calling to deliver BAM (Binary Alignment Nap) and variant call files (VCF files), for both germline and tumour samples. In addition to the requirements outlined for rare diseases, we expect pipelines to jointly call across tumour-germline DNA pairs, to identify somatic variants and report the purity of the tumour sample. This is specific to cancer.

### **Annotation of sequence variation**

The interpretation process will continue to distinguish between somatic variants, those mutations acquired by the cancer, and cancer susceptibility germline mutations that might have increased the overall lifetime risk of developing that cancer. For germline, annotation will exclude known benign common variants using internal allele frequencies and datasets such as 1000 Genomes, Exome Sequencing Project, and UK10K. The annotation of somatic variants will continue to highlight those found in known cancer genes and identify them as being likely to result in loss of function. To determine the likely importance of variants, it is expected that reference will be made to general cancer databases such as COSMIC, as well as gene specific databases, such as the IARC TP53 mutation database and other resources where appropriate.

A major component in cancer annotation is analysing the consequence of larger scale genomic changes, such as structural variants, copy number aberrations, loss of heterozygosity and other chromosomal mutational events.

As some genomic changes are not unique to the tumour sample, paired approach to the analysis of the two samples will continue to be important to ensure the origin of the variant is clear. Regarding the annotation of individual sequence variants, the consequence of copy number and structural variants will be assessed for their impact on genes implicated in cancer.

The variability in tumour sample purity, both with respect to contamination with normal tissue and the likely presence of multiple tumour sub-clones, is a particular challenge to interpretation pipelines. It is crucially important that in addition to paired analysis, the tumour is also analysed

separately. This will ensure that key somatic lesions are not filtered out by the variant calling algorithm, due to tumour contamination in the germline sample.

The results of interpretation are fed back to patients via an NHS genomic multidisciplinary meeting as described in section 6.

## **2.2 The Genomics England Infectious Disease partnership with Public Health England**

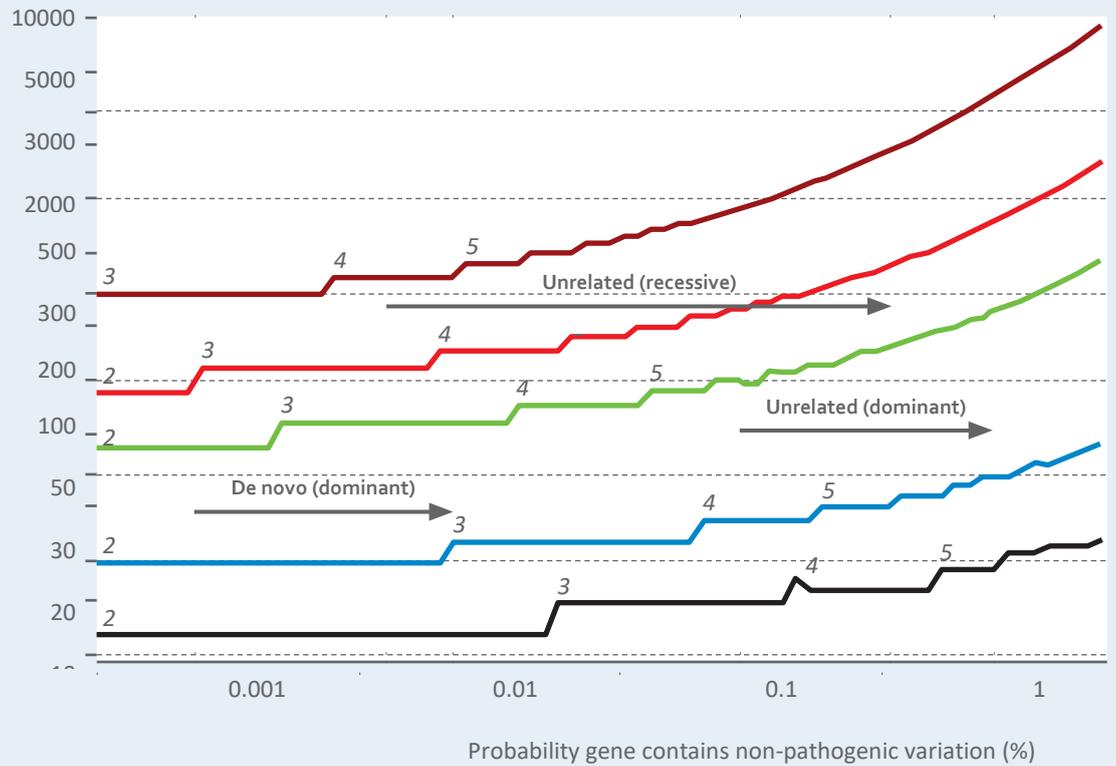
While the Infectious Disease Pathogen Programme formed an important part of the 100,000 Genome Project, it will continue to be implemented by Public Health England. Currently, the data from this programme will not be included in the Library.

## **2.3 Sample size considerations for research in rare disease and cancer**

### **Rare Disease Sample sizes**

The NHS GMS will be obliged to test patients identified through clinical care to obtain a diagnosis for them, as part of a duty of care. Therefore access to testing cannot be governed by power calculations. For research - sample sizes in rare disease are affected by the available number of families, the mode of inheritance, penetrance of the genetic variation and the depth of phenotyping. Different types of variation may be detected more readily by specific tests. Below we depict power simulations for research that allow researchers to gauge under different scenarios the likelihood of detecting a pathogenic mutation.

**Figure 1.** Different rare disease experimental designs are marked by arrows below. The number of detected mutations in cases required to pass genome-wide significance ( $p < 2 \times 10^{-6}$ ) is shown by the italic numbers on the power curves (only numbers five or smaller are shown).



Probability gene contains pathogenic variation    ● 80%    ● 40%    ● 10%    ● 5%    ● 2%

The Library will contain all of the genome data from the 100,000 Genomes Project, together with other WGS cohorts from external collaborators, and will be added to year on year by samples obtained via the NHS GMS; this will result in one of the largest collections of genomic data in the world. For individual research questions, data will be available to allow cohort size to be assessed and power calculations to be applied in planning methods to address research hypotheses. Where the data centre contains insufficient samples to address a particular question, routes to address this will be discussed with the research team.

### **Cancer power simulations**

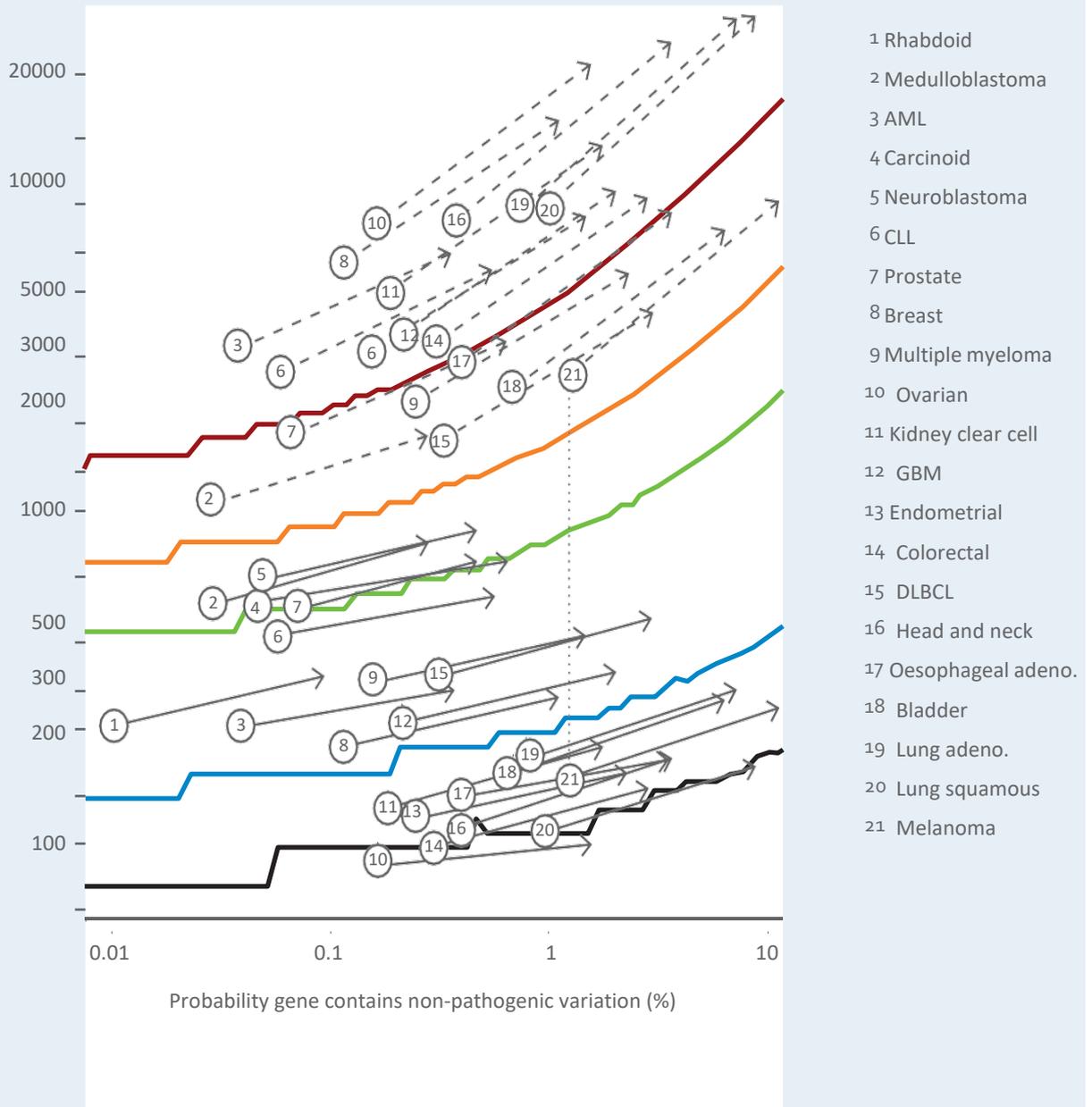
Again, the NHS GMS will be obliged to test patients identified through clinical care to obtain a diagnosis for them, as part of a duty of care. Therefore access to testing cannot be governed by power calculations. To assess the power for research purposes, in cancer studies we have considered the sample size needed to detect a gene containing pathogenic coding variation, we characterise a particular study by the rate of neutral variation in healthy individuals (the x-axis below) and the rate of pathogenic variation (the coloured lines below; i.e. the fraction of cases caused by a mutation in the relevant disorder). The following calculations and plots generalise the results of Lawrence et al 2014.

In the figure below, the solid arrow depicts the range of sample sizes needed to detect the most commonly mutated gene (e.g. TP53 in ovarian and lung cancer, and APC in colorectal cancer). Studies with sample sizes below these solid arrows will find no significant genes at all. The dashed arrow shows the most infrequently mutated known gene. To start finding many new genes in any of these tumour types, you need to have a sample size above this dashed arrow. The arrows are sloped to show the increase in sample size you would need, to find more mutagenic genes at this frequency. Although it is likely new genes will be detected as sample sizes increase above the left-hand side of the dashed arrow, to be well powered across all genes, you will need sample sizes to exceed those indicated on the right-hand side of the dashed arrow.

In cancers where there is only one known gene (rhabdomyosarcoma, carcinoid, and neuroblastoma), and they are expected to be relatively oligogenic, there is only one arrow. This analysis shows us that lung cancers will be the hardest to solve (dashed arrow 19 and 20 on Figure 3).

In lung cancer, significant genes containing driver somatic mutations such as TP53 will be detectable in <30 samples. A study that is powered to detect new lung cancer genes would need to have a very large sample size to get over the high mutation rate, e.g. 5,000 cases at minimum, ranging up to >20,000 to pick up larger, or more mutagenic causal genes (the right-hand side of the dashed arrows 19 and 20 in Figure 3).

**Figure 2.** Sample sizes required for 90 percent power to detect a disease associated gene under different scenarios at a genome-wide significance level of  $P = <2 \times 10^6$ . The x-axis shows the probability that the gene contains a relevant missense mutation/variant in a healthy individual (a function of gene size, mutation rate, and genetic model). The coloured lines show the power required for different frequencies of pathogenic mutations in that gene. We looked at two example genes for each tumour, the most commonly mutated known gene, and the most infrequently mutated known gene found in Lawrence *et al.*<sup>9</sup>



Probability gene contains pathogenic variation

- 80%
- 40%
- 10%
- 5%
- 2%

In the above figure, the solid arrow depicts the range of sample sizes needed to detect the most commonly mutated gene (e.g. TP53 in ovarian and lung cancer, and APC in colorectal cancer). Studies with sample sizes below these solid arrows will find no significant genes at all. The dashed arrow shows the most infrequently mutated known gene. To start finding many new genes in any of these tumour types, you need to have a sample size above this dashed arrow. The arrows are sloped to show the increase in sample size you would need, to find more mutagenic genes at this frequency. Although it is likely new genes will be detected as sample sizes increase above the left-hand side of the dashed arrow, to be well powered across all genes, you will need sample sizes to exceed those indicated on the right-hand side of the dashed arrow.

In cancers where there is only one known gene (rhabdomyosarcoma, carcinoid, and neuroblastoma), and they are expected to be relatively oligogenic, there is only one arrow. This analysis shows us that lung cancers will be the hardest to solve (dashed arrow 19 and 20 on Figure 3).

In lung cancer, significant genes containing driver somatic mutations such as TP53 will be detectable in <30 samples. A study that is powered to detect new lung cancer genes would need to have a very large sample size to get over the high mutation rate, e.g. 5,000 cases at minimum, ranging up to >20,000 to pick up larger, or more mutagenic causal genes (the right-hand side of the dashed arrows 19 and 20 in Figure 3).

The genetic architecture of some cancers such as prostate or medulloblastoma, may require sample sizes of <1,000 to discover many novel lower frequency (1-2 percent) somatic mutations.

Breast cancer has a reasonably low mutation rate, and lots of common driver genes (PIK3CA, PTEN, and P53 are each mutated in about 1/3 of cases). Therefore studies ranging between 30-50 samples will find plenty of genes. However, as this is a very heavily studied cancer, many of the rarer genes have already been found. To significantly add to this set and make new discoveries, a study design will need to have 3,000 to 10,000 samples to be well powered.

### **2.3.1 Declaring significance**

Genome-wide significance, correcting for the number of genes in the human genome, is typically defined as  $p < 2 \times 10^{-6}$ . The number of detected mutations in cases required to pass genome-wide significance is shown by the italic numbers on the power curves in Figure 1 (only numbers five or smaller are shown).

### **2.3.2 Variation in gene size and mutation rate**

The median human gene has 1,293 coding bases and a mutation rate of  $1.2 \times 10^{-8}$  bp/generation. The 90th percentile of gene size is 3218 base pairs, and the 90th percentile of mutation rate is 3.9X, the median rate. In Figure 1, the arrows for rare disease show the range of background variation, from the median gene on the left, and a large mutable gene on the right. For cancer in Figure 3, this is shown by solid or dashed lines.

### 2.3.3 Genetic architecture, and discovering genes vs solving diseases

We model disease heterogeneity as a geometric distribution, i.e. the fraction of cases caused by mutations in the most commonly mutated gene is  $x$ , the fraction explained by the next most commonly mutated gene is  $x(1-x)$ , and the fraction caused by the  $n$ th most common gene is  $x(1-x)^{n-1}$ . Although this is a crude approximation, it is probably conservative.

### **3 Opportunities to extend beyond the primary phenotypes**

## 3 Opportunities to extend beyond the primary phenotypes

### 3.1 Opportunities to extend beyond rare diseases, cancer and infection

#### 3.1.1 Lifelong electronic health record linkage

Genomics England's 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 outcome. This information will further enhance the research dataset.<sup>xv</sup>

The evaluation of WGS data in the context of rich and extended phenotypes derived from electronic health records, such as blood pressure, cholesterol, glucose, and pharmacogenomics, adds significant value. The richness of the Library dataset will allow us to move beyond the primary phenotype of the rare disease, cancer or infectious disease that led to the patient's initial WGS in the context of other continuous traits, diseases and response to therapy including harm. To do this, we will partner with Health Data Research UK as appropriate.

#### 3.1.2 NHS and public health registry data linkage

A key benefit of patient involvement in research concerns the linking of data to wider datasets and the potential direct clinical benefits that may stem from this. As these benefits are determined, we hope to incorporate them in the clinical diagnostic process. We will continue to build on this work throughout the lifetime of the Library.

We provide data to NHS Digital and Public Health England and will do so to any appropriate organisation holding relevant data on our participants. This enables them to extract data from their systems, which they then pass back to us to form the richest possible datasets on which we carry out our research.

NHS Digital, Public Health England and other NHS organisations are responsible for the national and local patient-level registries for several areas relevant to the Library, including cancer and rare diseases. The National Cancer Registration Service collects data on every patient diagnosed with a cancer-registerable condition across England. The data is collected from sources covering the whole pathway from referral, screening, to palliative care. It combines diagnostic, treatment and outcome data, including individual patient reported outcome measures. Data is normalised and quality-assured. There are systems to feedback the data, alongside comparative performance reports, to individual clinical teams and clinicians. This will provide a continual incentive for its improvement.

The National Congenital Anomaly and Rare Disease Registration Service (NCARDRS), part of the UK Rare Disease Strategy, was launched in 2015 and now covers the whole of England.

Similar benefits in terms of access to data are likely to accrue (as seen for the linkage to cancer data) as it continues to develop. As well as being a key source of phenotypic data linkage, this dataset will help inform the genetic makeup of the epidemiology of the diseases (by assessing how representative the sample is of all cancers/rare disease of that type).

There are a number of other health datasets, referred to in Section 9.

### 3.2 Opportunities beyond whole genome sequencing

We propose to continue to create a high quality biological sample resource from participants in support of future health research, where there is available funding for this. 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 National Genomic Test Directory along with the sample handling guidance detail the tissue handling and sample requirements for each test. We intend to continue to collect the following samples in the NHS GMS and the NIGMC as appropriate:

- Serum and plasma for proteomics and metabolomics.
- Cell-free serum for circulating tumour DNA and to assess tumour recurrence.
- Germline RNA for transcriptomics.
- Lymphocyte DNA for epigenetics.
- Tumour for RNA expression profiles, tumour epigenetics and proteomics.

The NHS GMS, and the infrastructure used to support it, will continue to have the ability to collect samples of these types and others, where funding is available from research via Genomics England or any other funding route that has been agreed by Genomics England and NHS England.

We recognise other types of sample collection may be valuable for research but they will need a specific consent. Examples of these may include:

- 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).
- Tissue specific RNA samples may be collected from individuals where a putative splice variant requires assessment.

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. The storage and processing of samples will account for the assays that are considered most likely to be used in the future. For example in WGS, transcriptomic, epigenetic, metabolomics and proteomic research. These samples will bring potential to look at a complete functional pathway - from DNA to the transcription pathway and genetic modifiers to metabolites, and the associated development of companion diagnostics and other technologies.

### 3.3 Opportunities for recruitment beyond the NHS GMS/clinical services

The primary source of data and samples for the Library will be the NHS GMS. However, in a number of cases there are cohorts of patients or population cohorts, who have had their samples and data taken for research or as part of clinical care. These are outside of the funding and structures of the NHS GMS. Individuals who are part of these cohorts, or whose data can be sourced via alternative clinical services to the NHS GMS, would benefit from being included within the Library and their inclusion may help to realise the clinical benefit for those participants who are recruited via the standard diagnostic service structures. Examples of these benefits include:

- A larger set of genome and phenotypic data that allows greater accuracy in, for example, identification of rare variants that will likely improve the diagnostic rate of all individuals who participate.
- Inclusion of trial cohorts of cancer patients where some of the trial data can provide valuable insights that will inform development of cancer genomic clinical reporting functionality, future treatment and diagnosis.

Where patient samples already exist that meet the technical requirements for sequencing and appropriate sample metadata, clinical data and other data is available within existing cohorts, these could be processed via the NHS GMS or Genomics England infrastructure. This is in line with the clinical transformation aspects of the Library, and as would be expected in standard clinical care.

The overriding principle behind this is that where possible it would be inappropriate in a clinical setting to put individuals through 'additional' procedures, such as the collection of blood.

While recruitment via NHS GMS will be the main source of data and samples, it may not always be feasible for clinical, operational, technical, contractual or governance reasons for patients to be recruited for research via this route. For example, where individuals have been recruited through an existing clinical structure it may be in the participant's interest that results are returned via the same structure e.g. this could apply to devolved nations.

To adhere to the above principle and to ensure equity of access to the programme including equity of direct patient benefit for taking part in research activities, the Library has the following approach for participants recruited outside NHS GMS structures.

- A standard operating procedure (SOP) for initial assessment of requirements and systems that need to be put in place for mobilising research cohorts appropriately.
- A governance approach led by Genomics England that allows a legally binding agreement to be put in place with the responsible individual/organisation for the cohort that sits outside the NHS GMS.

The specific responsibilities outlined in the legal agreement will be informed by the assessment of requirements and include, but are not necessarily limited to, responsibility for ensuring the following areas are enacted in a manner equivalent to those in place via NHS GMCs:

- a. recruitment and consenting
  - b. provision of samples and requirements sample quality, including replacement samples if required or duty to inform participants if they become ineligible due to sample quality issues (note, in some cases it may be provision of existing genomic data where sequencing has been undertaken elsewhere)
  - c. provision of relevant sample metadata, clinical data and other data
  - d. analysis of genomes, including the Genomics England bioinformatics analysis pipeline and provision of relevant genomic testing knowledge base
  - e. mechanism of return of results and relevant clinical support or referral for additional findings (where these have been requested).
- Standard resources requirements to project manage the operational aspects of mobilising the cohorts.

- Inclusion in Library governance and oversight structures of ongoing checks that obligations to participants are being met.

While it is envisaged that this mechanism would mainly apply to research cohorts meeting the criteria for inclusion within the rare disease or cancer programme areas, this route may also be particularly appropriate for including cohorts with a focus on complex diseases or those thought to be of polygenic origin into the programme. Research into these disorders is within the programme's remit (e.g. via the linkage and analysis of secondary data sources with genomic data) but presentation of these phenotypes is not currently a direct reason for recruitment via NHS GMCs. There is a growing body of evidence being developed around, for example, polygenic risk scores that suggest genomic data is likely to be of importance to direct clinical care. The longer term transformative aims of the programme include providing the ongoing evidence base for consideration of adoption of these new approaches within the NHS GMS and beyond. In the rare disease programme, it is already recognised that for many participants the benefit in direct care regarding their pertinent disease may only occur after further research and gene discovery, similarly in the cancer programme the main clinical benefit may be for individuals with a recurrence of cancer. As such, expectations of immediate compared to future potential clinical benefit and the possibility that there may not be a direct benefit to the patient are covered in the consenting process. However, for cohorts with complex disorders we can be more certain that the potential direct clinical benefit regarding the individual's pertinent disease is likely to be over a longer time period. In such circumstances the legal agreement outlined above would more explicitly cover responsibility for:

- Ensuring this is understood by participants
- Ensuring that appropriate provision is in place for realising the clinical benefits

Requirements around ensuring the realisation of other benefits to participants, e.g. for those who opt for return of additional findings, would be the same.

Similar considerations apply to cohorts where an existing molecular diagnosis is known but genomic data is identified as important in understanding the presentation and prognosis of the disease.

This approach makes it operationally feasible to include additional cohorts from trials and research studies, in a manner that helps us to achieve major aims of the programme i.e. patient benefit and research.

## **4 The Genomics England Clinical Interpretation Partnership (GeCIP)**

## 4 The Genomics England Clinical Interpretation Partnership (GeCIP)

The overall aim of the Genomics England Clinical Interpretation Partnership (GeCIP) is to create a thriving, sustainable environment for researchers and clinical (NHS) disease experts, trainees, and an industrial engagement approach, including both direct and consortia engagements. This will stimulate and encourage new research endeavour and information exchange. This community will analyse and constantly refine the clinical and scientific interpretation of the Library dataset. The intention is to further improve understanding of findings and their implications for genomic medicine and the clinical setting. GeCIP was launched at the Wellcome Trust on 27 June 2014.

GeCIP will also provide an excellent basis for further research and development in genomic medicine in the UK by academics and industry. All GeCIP users will be required to contribute results and data to the Genomics England Data Centre to enhance the scientific knowledge base. All data is available to all GeCIP members who have access to the Research Environment, however, a nine-month moratorium on presentation and publication is applied from entry of data into the Research Environment to allow GeCIP domains to have priority access to samples that they were responsible for collecting.

The scientific aims of the GeCIP programme are as follows:

- 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. Rare disease will also be informed by other 'omic' investigations.
- Innovative clinical interpretation in cancer, including multi-omic datasets (e.g. transcriptomics, epigenetics, and proteomics), analysis of circulating tumour DNA, sequential biopsy to address the genetic architecture of cancer, validation and characterisation of variants, identification of novel therapeutic targets, or repurposing of existing therapies.
- Expand the programme to include other disease areas with additional funding, to address specific research questions and opportunities to develop stratified approaches. The Genomics England infrastructure will be designed to facilitate expansion and re-use, and individual GeCIP partnerships can be extended to programmes (with funding) outside of the Library and its legacy.
- To capitalise upon electronic health records research, we can build upon and add value to the clinical, laboratory and health records data, linked to variant call data, which is held securely within the proposed data and computing infrastructure.
- Algorithms, models, and tools for clinical genomics research, data quality assurance, and the annotation, interpretation and presentation of genomic, clinical, and laboratory data in combination, may be developed, evaluated, used, and shared within the proposed infrastructure.

All GeCIP partners will adopt the highest ethical standards in accordance with the terms of the Genomics England consent. This will be with the continued support and advice of the Genomics England Ethics Advisory Committee and the GeCIP Board, comprising representatives of the GeCIP community and funders of GeCIP activities (see section 4.7).

The activities of GeCIP will inform clinicians and multidisciplinary teams within the NHS GMS, by providing enhanced data interpretation, additional information on pathogenicity of variants, and functional characterisation. The process of NHS feedback of research findings will be managed by Genomics England and, for the NHS GMS, the Genomics Unit at NHS England. The details for this process are outlined in Section 4.

#### **4.1 The route by which researchers add value to GeCIP**

It is the ambition of Genomics England and the GeCIP funders, to see a really strong research and training programme alongside the evolving Library. This will create opportunities for GeCIP researchers and trainees, which will stimulate and support research programmes funded by GeCIP funding partners, through response mode and specific funding calls. The researchers leading these programmes will have access to appropriate patient groups from the NHS in the UK (see later for details related to the devolved nations). They have formed and now lead appropriate national and international consortia within cognate domains that provide NHS samples, clinical data, and analytical skillsets. This will maximise the value of the dataset and ensure that the clinical interpretation delivered by the Library and its legacy remains at the forefront of genomic medicine. This may be achieved in a variety of ways, including:

- Enhanced clinical interpretation focused on rare disease. Omic data will also inform this.
- Innovative clinical interpretation in cancer, including multi-omic datasets.
- Expanding the Library and its legacy to include other disease areas beyond cancer and rare diseases.
- A researcher or trainee can become a member of the GeCIP provided they work according to the Genomics England policies (including those funded specifically to work within the Library by a GeCIP funding partner).
- If a research proposal or training fellowship is planned, it is recommended that applicants contact the Chief Scientist's team, so they can ensure compatibility with the mission and assist where necessary.
- Delivery of health records research using the linked, longitudinal dataset.
- Delivery of algorithms, models, and tools to facilitate interrogation and analysis of linked, longitudinal datasets.
- Work as part of the clinical validation groups to review data and interpret findings (see below).
- Commitment to provide results or findings to the Genomics England Knowledge Base, in addition to publication to the wider scientific and medical community.

## 4.2 The route for those who contribute patients and data from the NHS to add value to GeCIP

GeCIP will be the route for NHS and other organisations who recruit patients, contribute samples, and/or data to the research programme, or those funded on specific NHS training programmes related to Genomics England (NHS Trusts, NHE England, Public Health England, and Health Education England) to carry out research using the data. This offers the opportunity for those in contributing NHS organisations to participate in the analysis and interpretation of sequencing, to inform release of clinical reports back to the NHS.

Contribution from the NHS could include:

- NHS contributor who adds value to the GeCIP (e.g. samples, data, annotation, tools etc.).
- Delivery of updated clinical, phenotypic, and laboratory data on patients (through existing Genomics England tools and pathways).
- Working as part of the clinical validation groups to review data and interpret findings.
- Committing results/findings to Genomics England knowledge base.

## 4.3 Establishing and working within the Clinical Interpretation Partnership

The GeCIP has been organised into a series of disease-specific domains by an open call - in which UK researchers, trainees and clinicians self-organised and created proposed GeCIP domains. Key selection criteria were published in the GeCIP guidance document accompanying the call. Successful GeCIP proposals clearly articulate the value they will add to the clinical interpretation of the Genomics England dataset.

Researchers, NHS clinicians/healthcare professionals and trainees will work as part of disease-facing or cross-cutting domains (depending on their area of expertise), contributing patients, phenotypes, knowledge, expertise, and undertaking research to add value. This work will be continuous, as the dataset is updated with patient outcomes and detailed clinical, phenotypic and laboratory data, which is collected during routine NHS care. Some clinicians will work in several domains if they have a broader clinical or laboratory remit.

The list of clinical domains described below is representative of the diseases we are now working on in the main programme. It is subject to change at the discretion of Genomics England. In our call for GeCIP domain proposals, we encouraged the applicants to self-organise and make their own domain proposals. This may include domains not mentioned below. They will also be encouraged to propose a UK lead and we will work with applicants to optimise the GeCIP domain as needed.

**Figure 5.** The data flows that enable the Genomics England Clinical Interpretation Partnership to function.

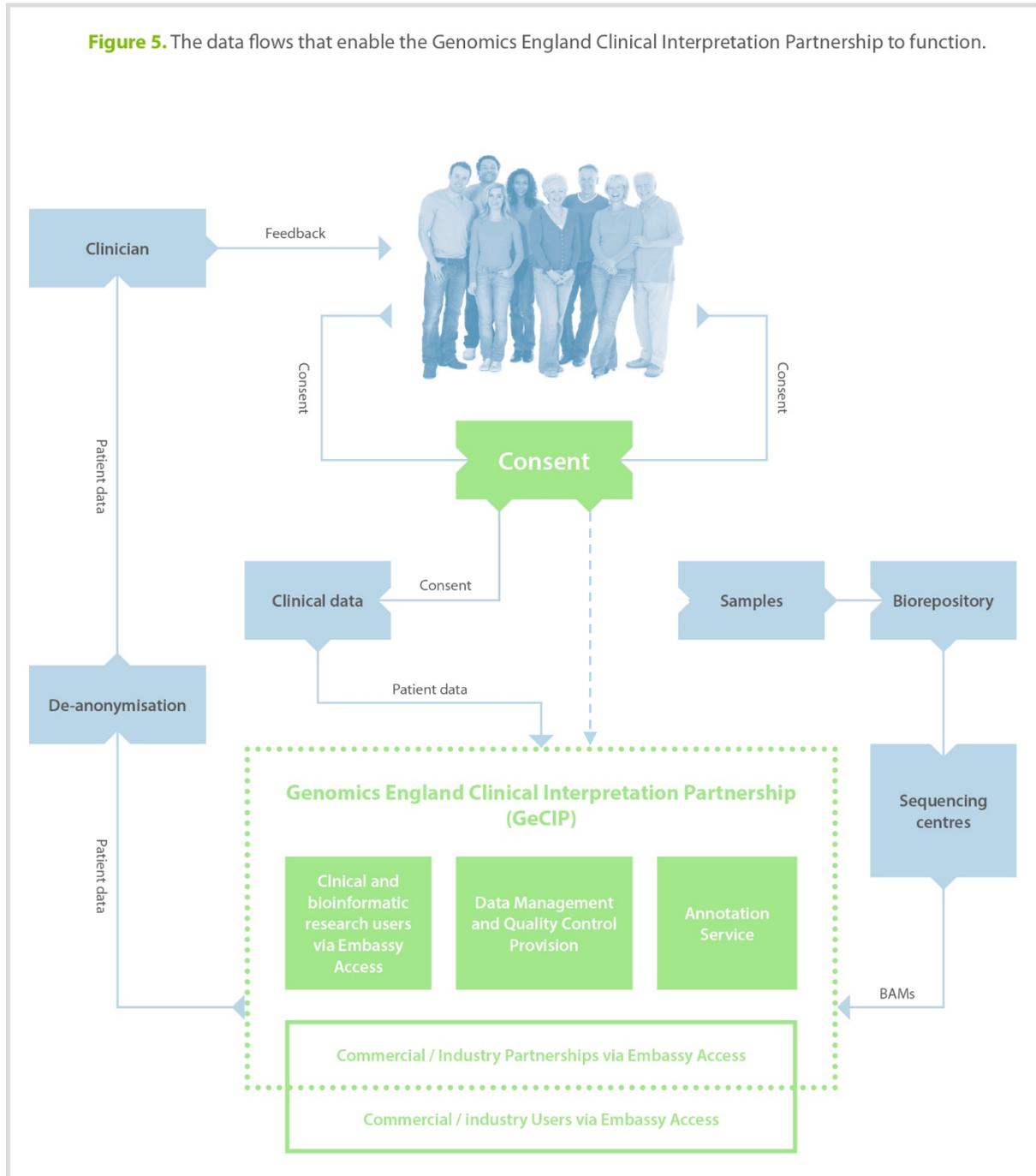


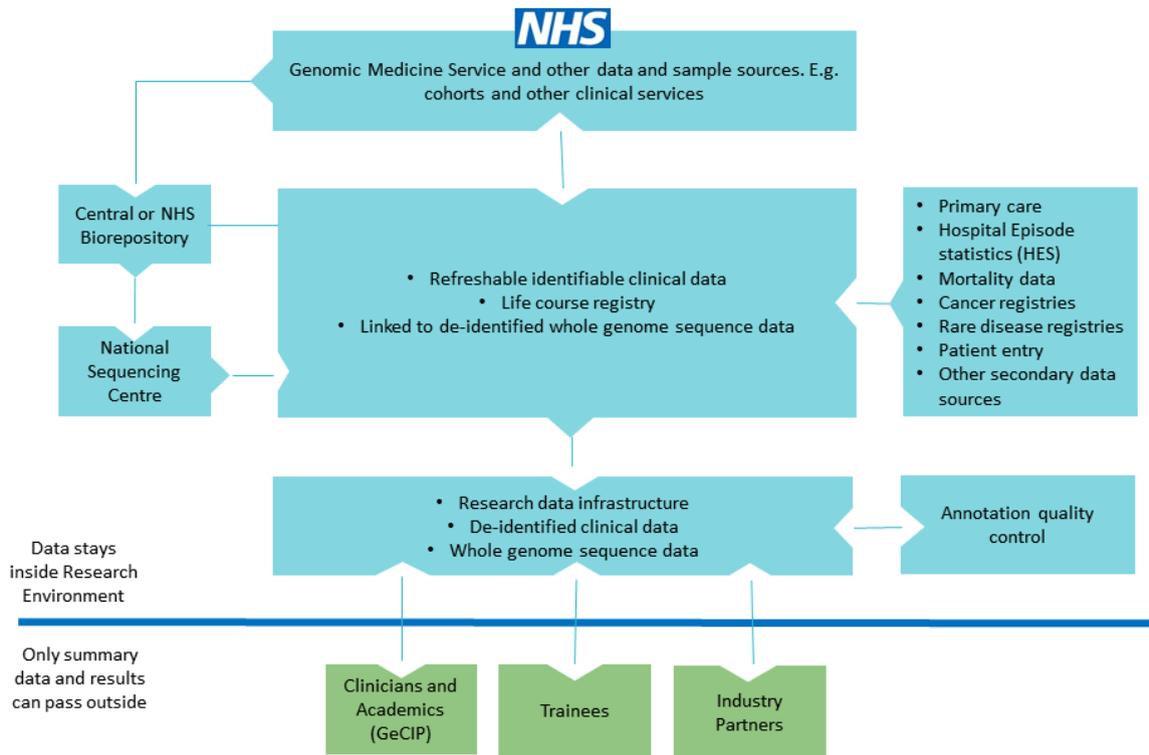
Table 1 - GeCIP domains.

Disease-specific domains		Functional/Cross-cutting domains
Cancers	Rare inherited diseases	
Breast cancer	Hearing and Sight	Education and training
Cancer of unknown primary	Cardiovascular disease	Ethics and Social Science
Childhood solid cancers	Respiratory	Electronic health records
Colorectal cancer	Endocrine and metabolism	Enabling rare disease translational genomics via advanced analytics and international interoperability
Glioma	Gastroenterology and hepatology	Functional crosscutting
Haematological malignancy	Immune disorders	Functional Effects
Head and neck cancer	Neurology	Health economics
Lung cancer	Musculoskeletal	Quantitative methods, machine learning and functional genomics
Melanoma	Skin	Population genomics
Neuroendocrine tumours	Renal	Stratified medicine
Ovarian cancer	Non-malignant haematological and haemostasis disorders	Validation and Feedback
Pan-cancer	Paediatrics	
Prostate cancer	Inherited Cancer Predisposition	
Renal cell carcinoma	Paediatric Sepsis	
Sarcoma		
Testicular cancer		
Upper gastrointestinal cancer		

This list is correct as of May 2018, see [www.genomicsengland.co.uk/about-gecip/gecip-domains/](http://www.genomicsengland.co.uk/about-gecip/gecip-domains/) for the most up-to-date information.

The GeCIP domains can be led and proposed by researchers, NHS clinicians/healthcare professionals and those training anywhere in the UK. They are not confined to England.

## Genomics England Clinical Interpretation Operational Plan



Patients outside the UK cannot be sequenced in the Library, unless full funding is provided by a funder and they join a GeCIP domain and agree to Genomics England Terms and Conditions.

### 4.4 Accessing data as a Genomics England GeCIP Member

All research data will be pseudonymised and managed in accordance with the latest guidance.

GeCIP users will be granted access to all data and knowledge held within the Library dataset. Access to the secure virtual Research Environment will be governed via a data access registration and approval process. Each GeCIP domain will have access to its own private shared area of the research environment for data storage and collaboration.

The secure virtual desktop infrastructure will provide the 'workspace' for clinical teams, research groups and trainees undertake their work. Work undertaken within the domains is subject to the governance and terms and conditions of Genomics England (see below). Breaches of the governance, or terms and conditions of use, will result in the removal and exclusion of those responsible from the Genomics England Research Environment.

The table below illustrates the structure of a domain.

**Table 2.**

<b>Groups Who May Access the Research Environment</b>	
<b>Activity/Group</b>	<b>Key functions and outputs</b>
Genomics England Chief Scientist's Team	<ul style="list-style-type: none"> <li>• Oversight, informatics and logistics</li> </ul>
GeCIP Steering Committee	<ul style="list-style-type: none"> <li>• Coordination and management of domains</li> </ul>
Clinical Interpretation including NHS clinicians	<ul style="list-style-type: none"> <li>• Highest fidelity dynamic reporting system</li> <li>• Contribution to Interpretation, Validation and Feedback domain</li> </ul>
Genomics researchers	<ul style="list-style-type: none"> <li>• Novel genomic discoveries</li> <li>• Engagement of international collaborators</li> </ul>
Multiple phenotypic sub-groups	<ul style="list-style-type: none"> <li>• Deeper phenotyping and sub-analyses</li> </ul>
Analysts and Bio-informaticians	<ul style="list-style-type: none"> <li>• Novel analytic approaches</li> <li>• Development of analytical and reference resources</li> </ul>
Functional characterisation Multi-omics	<ul style="list-style-type: none"> <li>• Single cell or model functional studies</li> <li>• RNA, epigenetics, proteomics</li> </ul>
Trainees and training director	<ul style="list-style-type: none"> <li>• Research projects and higher degrees through the Genomic Medicine Academy</li> </ul>
Precompetitive industry partners	<ul style="list-style-type: none"> <li>• Academic-industry collaboration to accelerate application of new findings</li> </ul>
Industry partners	<ul style="list-style-type: none"> <li>• Industry partnerships to accelerate application of new findings</li> </ul>

Note: It is anticipated that precompetitive industry partners will work within the GeCIP domains. All data generated in a GeCIP domain remains available to all.

In certain situations, there may be direct clinical benefit in allowing researchers to access data which could be considered identifiable (for example the use of full postcodes to identify geographical health-related risk factors). In such special circumstances, if the Access Review Committee agrees that access to this data would be appropriate and scientifically justified, this may be made available where appropriate legal agreements and additional safeguards are in place. Explicit provision for this, within this protocol, has been requested by patient representatives on the Access Review Committee.

#### **4.5 The commitment of all data generated to the Genomics England Knowledge Base**

Genomics England will encourage and promote an environment of open collaboration. GeCIP members will be expected to comply with this approach to enhance and maximise the value of the Genomics England knowledge base.

In all cases the research results, along with raw data and analytical steps undertaken within the GeCIP domains, must be provided to the Chief Scientist's team, for commitment to the knowledge base.

Genomics England is committed to delivering a managed research environment and will negotiate with GeCIP members regarding sharing of intermediate research results within the knowledge base. While it is understood that this is a sensitive area, it is hoped that such sharing will serve to accelerate scientific progress, delivering further patient benefits by enhancing feedback to the NHS.

#### **4.6 GeCIP contribution to interpretation, validation and regulation of findings**

The GeCIP domains and committees are expected to collaborate with regulators, including E-Quality Management System, National Institute of Health and Clinical Excellence (NICE), Medicines and Healthcare Regulatory Products Agency (MHRA) and NHS England to develop standardised approaches to validate findings from the Library and its legacy. In addition, this collaboration is expected to support the development of systematic approaches to verification of these scientific findings, evaluation of proposed applications on diagnosis and treatment options, and management of impact on commissioning strategies and costs. This joined-up-approach will ensure swift translation of clinically validated research findings into adoption for patient benefit.

#### **4.7 Overarching GeCIP management and governance**

The GeCIP Board, chaired by a Genomics England Board member, oversees the governance of GeCIP. Membership will also include a GeCIP lead for the NHS contributors relating to cancer, rare disease and cross cutting domains, as well representatives of major funders, and a representative(s) from the National Participant Panel and from NHS England. Genomics England will be represented by its Chief Scientist. Further details on the governance structure are available at [www.genomicsengland.co.uk/about-genomics-england/the-board/clinical-interpretation-partnership-board/](http://www.genomicsengland.co.uk/about-genomics-england/the-board/clinical-interpretation-partnership-board/).

A GeCIP Steering Committee will be established to ensure the smooth running and transfer of knowledge across the GeCIP domains. The membership will consist of leadership from each disease-facing domain. It will be chaired by the Chief Scientist for Genomics England. The GeCIP Steering Committee will directly report to the GeCIP Board. This committee will include representatives of the public, patients and medical research charities.

## **5 Commercial partnerships**

## 5 Commercial Partnerships

### 5.1 Overview and aims of Genomics England commercial activities

Commercial partnerships are crucial to achieving the aims of the Library. Genomics England supports use of the Library to stimulate the Lifesciences Industry, to develop new therapies and technologies that will improve healthcare and to recover inward investment made by the UK government. This also includes bringing together relevant industry stakeholders to grow the the genomics medicine 'UK plc'; attract inward investment from life science companies; and with an aim of increasing opportunities of access to medicines that would otherwise be unavailable to UK patients.

Genomics England engages with commercial partners across a wide industry ecosystem. These include:

- Biopharmaceutical
- Biotechnology
- Molecular Diagnostics
- Genetic Testing
- Informatics and Sequencing
- Data Infrastructure
- Healthcare SME
- Clinical Research Organisations
- Consultancies and investment companies
- Clinical Research Organisations

A wide network of industry partners enables Genomics England to create awareness of the potential that the application of genomics in healthcare can bring. In turn, this allows Genomics England to engage with relevant stakeholders to describe unmet healthcare needs and to provide a resource that allows industry to engage with addressing these issues.

Fundamentally the framework and governance structures that enable commercial research activities within the Library mirror those that have been put in place for the GeCIP. All commercial partners accessing the data in the Library need to have appropriate ethical and scientific approvals, legal agreements and information governance training in place before gaining access to the pseudonymised dataset. The Library enables the same data and tools to be used for research within a secure environment by academic and commercial researchers and is designed to promote collaboration across the research groups.

### 5.2 Commercial research aims

The most common collaboration between Genomics England and a commercial partner is through accessing the Library. As with academic research led by GeCIP, commercial research aims to bring benefit to the patients and, through the use of the data, inform development of platforms and tools for future diagnostic discovery. Commercial research can be broadly categorised into four themes that answer different questions along the typical Research and Discovery Biopharmaceutical Pipeline. At a high level they are divided into:

- Diagnostic discovery
- Pre-clinical research
- Clinical Trials Referral
- Real World Evidence / Market Access

All research activities undertaken in the Library aim to enrich the existing dataset via one or multiple routes:

- Identification of diagnoses originally missed by the standardised pipeline
- Feedback of new diagnoses to patients
- Import of additional datasets to enable further research
- Mobilising multiomics samples which can help to identify diagnoses that were missed through analyses of WGS alone

Other forms of collaboration include direct partnerships to develop tools to improve systems and services.

### 5.3 Ways of working with commercial partners

#### 5.3.1 The Discovery Forum – access to a network (Associate membership)

The Discovery Forum has been established by Genomics England to enable commercial partners to join a collaborative network and to enable Genomics England to engage with industry.

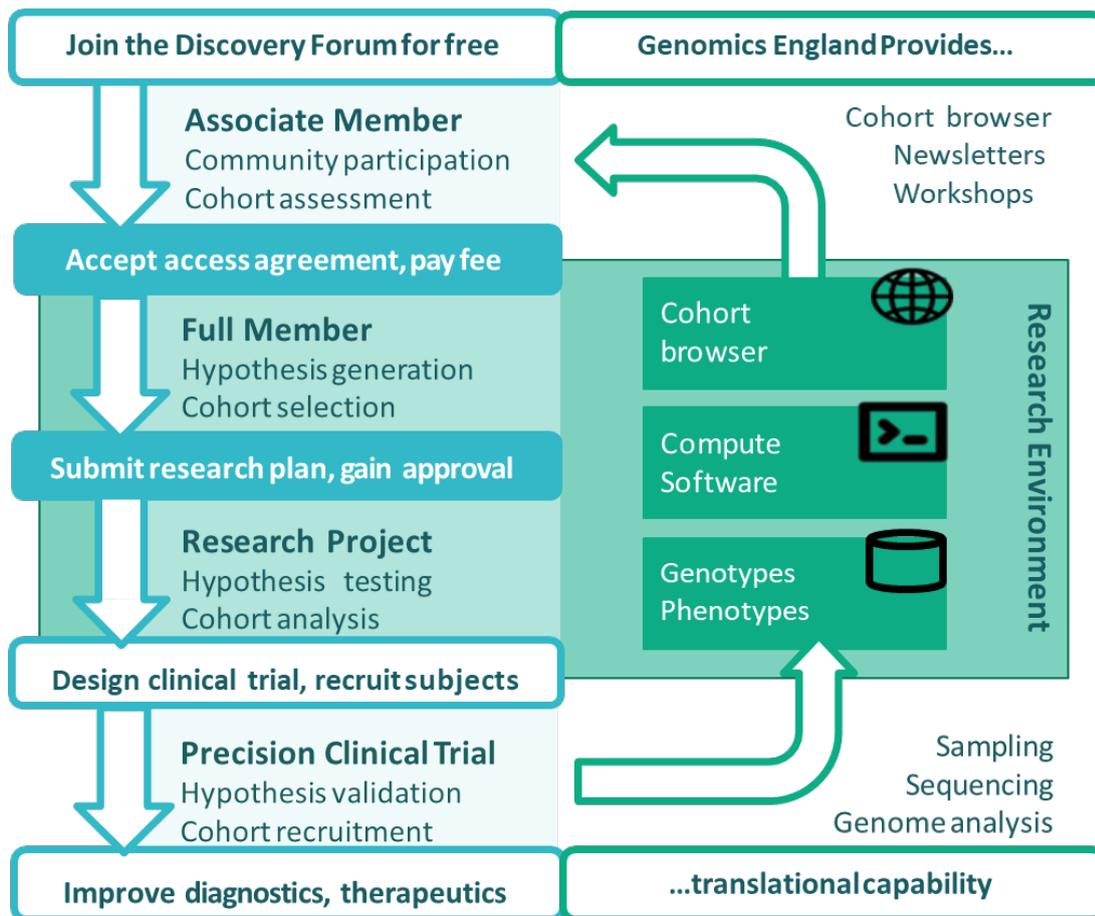
There are two types of membership: Associate and Full. All members of the Discovery Forum must sign terms of reference to join.

Associate members are non-paying members entitled to:

- Attend Discovery Forum meetings;
- Receive the Industry Newsletter, which details the latest developments within Genomics England
- Support from other members to facilitate next steps for engagement with Genomics England and its various partners

This network provides the following benefits:

- Technical input into platform and data structure development (research user base)
- Input into strategy planning
- Trusted community for partnership building



Any commercial partner wishing to engage with Genomics England must become an Associate member and undergo an approval process.

### 5.3.2 The Discovery Forum – access to the Library (Full membership)

Discovery Forum Full members are fee-paying and have access to the Library in line with the acceptable uses described in Section 9. Full members engage with the data to achieve commercial research aims, as described above.

Full members undertake two types of activity within the Library:

#### Pre-research:

- Allows researchers to define the scope of their proposed activity.
- Provides assurance that the Library is fit for purpose.
- Allows researchers to define cohorts based on known molecular, phenotypic and/or clinical attributes and/or to reproduce known phenotypic and/or genotypic associations.
  - These known attributes need to be in the public domain i.e. to have been published in scientific literature or other materials.

A pre-approved umbrella application (as approved by the Genomics England Access Review Committee) enables all companies to undertake pre-research after fulfilling legal and governance requirements. Pre-research is primarily for the development of business and scientific cases to

enable more detailed proposals. Pre-research does not allow for the results of association testing to be exported, or to be published.

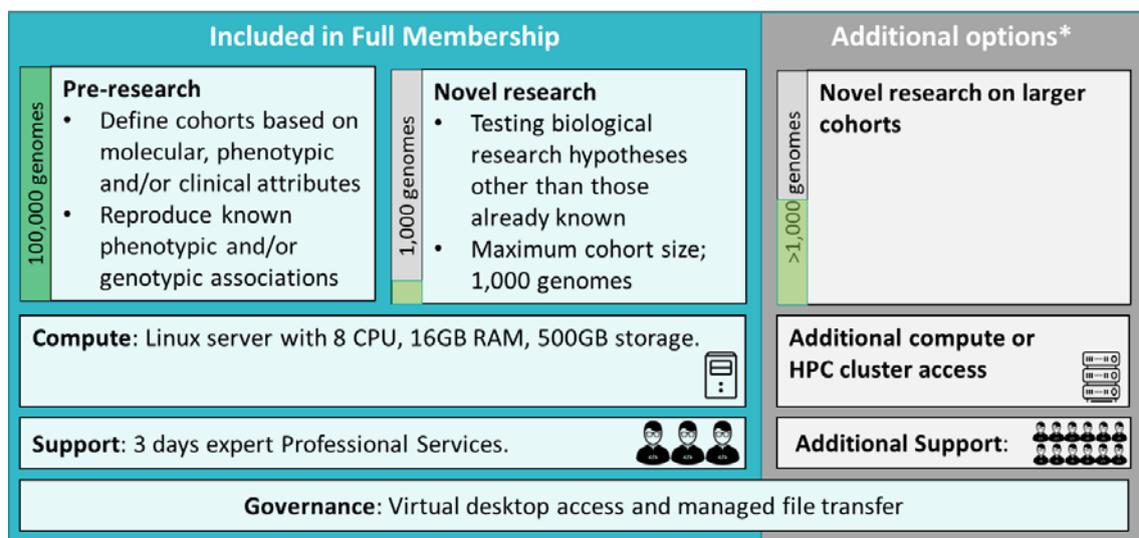
**Novel research and specific projects:**

- Analysis that sets out to generate never before published outputs.
  - Includes discovery and scientific validation of previously undescribed association between genetic and clinical data.

All novel research must be reviewed and approved in accordance to the established framework. Only after this has been completed, commercial researchers can export summary level data from the Library in accordance with the Airlock Process.

Details of activities and support provided to commercial partnerships are included in the diagram below. To accommodate Library growth and the evolution of platform analytical capabilities, Discovery Forum members have an option to engage with Genomics England at different tier levels. The tiering model allows Genomics England to:

- Demonstrate the ever growing value of the dataset to potential clients
- Enable building of trust
- Evolution of partnerships that is adaptable to commercial needs
- Access to data for all types of industry partners ie fostering growth of small start-ups and strategic partnering with major pharmaceutical companies



### 5.3.3 Strategic partnerships

Depending on the strategic aims of the company, Genomics England may need to facilitate other partnerships that enable collaboration beyond that of research on the Library dataset and ability to refer patients onto clinical trials. Strategic partnerships enable Genomics England to enhance its capabilities to in turn, ensure the data and discoveries stemming from it are translated into the healthcare setting. Some examples of such functions are:

- Generating data from patient samples
- Providing technological support to improve platform functionality and/or clinical reporting
- Facilitating patient recruitment and trial facilitation

### 5.4 Commercial research approval framework

All commercial researchers are expected to adopt the highest ethical standards in accordance with the terms of the Genomics England consent. Commercial partners are required to submit a countersigned and date-marked Data Access Agreement committing them to Genomics England policies. Commercial entities intending to carry out work in breach of section 9.3.2 (Uses that will be declined) are immediately rejected. Research proposals are subject to the same governance as outlined in section 4.5 (Accessing data as a GeCIP member), whereby a commercial research plan is assigned to an existing GeCIP domain. The exception being that commercial researchers are not subject to an institutional GeCIP Participation Agreement, but are instead verified against a defined set of exclusion criteria to ensure only reputable companies and appropriate researchers within those are eligible to access the Library.

As with GeCIP, PPI Representatives have an important role to play in Genomics England commercial initiatives. In involving PPI Representatives in the work of GeCIP, we aim to pursue a common purpose of providing patient and public perspectives throughout the research process. This is particularly important with commercial research to ensure transparency is upheld, and the interest of those whose data is being used is always being respected.

## **6 Feedback for participants**

## 6 Feedback for participants

Feedback of findings as a result of a diagnostic test request will be to the NHS GLH network (or equivalent for other cohorts where appropriate), through web-based tools developed or procured by Genomics England, as well as summary findings being available in a patient Genomic Record developed as part of the Test Order and Management System (TOMS). Each GLH will work with NHS GMS genomic multidisciplinary teams who will make recommendations to the patient's clinician based on the whole genome sequencing and other genomic test results. This will fall under usual processes for clinical governance.

### 6.1 Enhanced interpretation and handling of findings from the research environment

We have converted the GeCIP domain on 'Validation and Feedback' to 'Enhanced Interpretation' and they are developing a plan connecting the GeCIP research endeavour into the clinical service. Each GeCIP domain will be connected via an Enhanced Interpretation Domain member. The GeCIP team, working with the GeCIP domains, will alert the Enhanced Interpretation Domain to new findings that may be of potential clinical utility and benefit to participants. Where this is the case, this will be assessed and these findings may contribute to reporting from the Genomics England Interpretation Pipeline and the generation of a new whole genome analysis for the NHS GMS.

The decision on whether this analysis is taken forward to validation of key diagnostic findings, regardless of their origin, will be led by the new GLHs (including the Wales and NI GMC) and the new NHS GMS. Results of validation tests will continue to be shared with Genomics England to enhance the value of the Data Centre.

It is our goal to use this mechanism to improve the relevance of whole genome sequencing findings for healthcare during the life of the Library. The Library/GeCIP is committed to regularly publishing peer-reviewed updates on the current state of knowledge in this area. In preparing the Genomics England Feedback Policy, we took full account of all relevant legal and ethical principles including contemporary research and the advice of patients. This has also been informed by Professor Mike Parker's letter to the Chief Medical Officer from the 100,000 Genomes Project Ethics Advisory Working Group (see [www.genomicsengland.co.uk](http://www.genomicsengland.co.uk)).

GeCIP will be operated by the Office of the Genomics England Chief Scientist, with two major primary functions:

- (i) its role with regard to research and the research community (described above) and
- (ii) responsibility for supporting clinical validation of genetic variants of expected pathogenicity.

Diagnostic capability to detect pathogenic variants is expected to evolve. Any results generated within the research environment could be fed back to diagnostic clinical scientists and others for consideration for diagnostic reporting.

### 6.2 Policy for main findings, secondary findings, and incidental findings, and feeding back to clinicians and patients

Information on feeding back of clinically relevant findings to clinicians and patients is available at [www.genomicsengland.co.uk/taking-part/results](http://www.genomicsengland.co.uk/taking-part/results). The Library will take account of the evolving

policy and academic debates in the area and has been influenced by these in respect of its current formation.<sup>xvi</sup>

As in the main programme, types of results sought by the Genomics England bioinformatics pipeline will continue to include:

**Known or pathogenic variants or driver mutations** directly connected to the main disease that led the patient to take part in the Library and that are clearly actionable.

**Additional findings.** If the clinical test involves WGS, a very limited number of **looked-for known, pathogenic mutations** of high clinical relevance,<sup>xvii</sup> confined to a very limited list. This will be more limited than that used by the American College of Medical Genetics.<sup>xviii</sup> It will be possible for patients to opt-out of the feedback of findings of high clinical relevance. The list of findings to be sought in children will not include conditions where no action is indicated before reaching adulthood, to allow children to take their own decision on receiving these findings once they attain capacity to do so.<sup>xix</sup> Prospectively, this will be implemented as part of an evaluation framework within the context of the new NHS GMS.

**Parental carrier status** in cases of rare disease due to bi-allelic germline mutations. It is feasible that this programme will identify double parental carrier status or X-linked maternal carrier with high-risk of a reproductive consequence. These results will be assessed for suitability for diagnostic reporting by an NHS clinical scientist or equivalent.

Type of finding*	Description	Nature of the information to be fed back	How feedback will happen	Approach to consent or refusal
<b>Pertinent finding</b> (also known as a primary finding). Found in the research environment	A pertinent (primary) finding is relevant to the explanation, main diagnosis or treatment of the disease for which the patient was selected for testing.	To be reported as a pertinent (primary) finding, <b>mutations must be confirmed by a clinical service.</b>  Information provided to the clinical service will include categorising results as <b>1) 'Known to be Pathogenic' (KP)</b> i.e. sequence variation is previously reported and is a <i>recognised</i> cause of the disorder.  <b>2) Mutations that are 'Expected to be Pathogenic' (EP)</b> i.e. sequence variation is previously <i>unreported</i> and is of the type which is <i>expected</i> to cause the disorder.	Pertinent (primary) findings will be given to the referring clinician or clinical team by the 100,000 Genomes Project for discussion with the patient after the NHS Genomic medicine Service has validated these findings.	Where the data and sample source are a clinical service, the service is expected to define the return of results approach under clinical governance. For example, in the NHS GMS it is currently expected that Somatic cancer findings will be reported back without delay but pertinent germline findings will require consent. There may also be revaluation of negative results under clinical consent.  Where reporting is not under clinical governance the continuation of existing research-based return of results is expected to apply, thus ensuring continuity of the offer to participants in both setting.  Where testing is under a clinical service, patients may refuse to allow their data to enter into the research environment.

\* Classification of results is adapted by Genomics England taking into account the 'Bioethics Commission's Classification of Individualized Results of Medical Tests', P.27 table 1.2, of the Presidential Commission for the study of Bioethical Issues report, 'Anticipate and Communicate: Ethical Management of Incidental and Secondary Findings in the Clinical, Research, and Direct-to-Consumer Contexts', December 2013. <http://bioethics.gov/node/3183>. The utilisation of terms by Genomics England has been chosen to provide the greatest possible clarity for Genomics England's activity. Broadly speaking we identify two groups of results; primary/ pertinent and secondary/ non-pertinent. Within the secondary findings group we envisage that some findings will need to be (a) Actively looked for because they reach an evidence/ utility threshold/ on a predefined list, for which consent is sought at the time of testing and that (b) a second group, expected to occur rarely, are of health-related findings not actively sought because they fail to meet the evidence/ utility threshold, but where disclosure may still need to be considered, should they happen to be found. Findings in this second group may have personal utility. Those wanting to access their raw codes might derive this, but Genomics England will not routinely filter for these findings and disclose them.

**Incidental findings** are those that occur unexpectedly in the course of analysis, which could have implications for a patient’s health or care. Incidental findings identified during the course of research will not be fed back to patients or clinicians unless there is an exceptional reason for doing so. These might be considered by the Genomics England Science Advisory Committee or Ethics Advisory Group if advice is required to direct decisions about feedback. Incidental findings that occur in the context of a diagnostic analysis will be managed according to standard clinical practice initially.

**Table 3.** Summary of Genomics England’s policy on feedback of findings:

Type of finding*	Description	Nature of the information to be fed back	How feedback will happen	Approach to consent or refusal
<p><b>Secondary finding</b> (additional looked-for findings of healthcare importance). In patients having Genome Sequencing</p>	<p>A secondary finding is an additional, looked-for health related finding, that is <b>not</b> pertinent to (or a primary cause of) the main condition.</p> <p>It may be found <i>in addition to (or in the absence of) any</i> pertinent finding.</p>	<p>NHS England will evaluate the impact of looking for genomic findings that are known to cause serious conditions for which there is good evidence that knowing about them could influence healthcare. They will build on the limited list of relatively rare conditions, in the 100,000 Genomes Project and adapt this list as more evidence accumulates. The current list will be available on the website</p> <p>Please note: The Genomics England approach to findings relating to adult onset conditions in children will be in accordance with the BSGM policy on genetic testing of children** i.e. only the mutations in genes that are known to cause childhood-onset disease will be looked for in the case of minor participants.</p>	<p>Participants will be asked whether they would like these secondary findings to be actively sought.</p> <p>If they consent to this, information relating to secondary findings will be looked for and given to the referring clinician or clinical team for discussion with the participant.</p> <p>(Without this consent these secondary findings will <b>not</b> be actively sought.)</p> <p>Following positive or negative confirmation regarding these secondary findings, a health professional will feed back to the participant on these results.</p> <p>Some variants in this gene list may not yet have sufficient evidence for their clinical impact to be known, or clearly predicted. These variants of uncertain significance will not be reported.</p> <p>An important evaluation aim of the 100,000 Genome Project and the NHS Genomic Medicine Service and beyond is to generate evidence about the clinical relevance of new findings. Our understanding of the clinical significance of a secondary finding will change over time as more evidence is gathered.</p>	<p>Consent for the feedback of these secondary findings will be by opt-in to feedback.</p> <p>Patients who do not wish to receive information about these findings are free to refuse to consent.</p> <p>Participants can opt-in to consent to the list(s) of identified conditions to be looked-for at the time of consent, plus, if they wish, to findings made as that list extends. This will permit results regarding other conditions that meet this criterion in future to be looked-for and fed back.</p>

\*\* This list is based upon that proposed by ACMG plus subsequent expert review at Genomics England.

Green RC *et al.* American College of Medical Genetics and Genomics recommendations for reporting of incidental findings in clinical exams and genome sequencing. *Genetics in Medicine* 2013; 15(4): 565-574. Available at: [https://www.acmg.net/docs/IF\\_Statement\\_Final\\_7.24.13.pdf](https://www.acmg.net/docs/IF_Statement_Final_7.24.13.pdf).

See also ACMG Updates Recommendation on "opt Out" for Genome Sequencing Return of Results, April 2014. Available at: [https://www.acmg.net/docs/Release\\_ACMGUpdatesRecommendations\\_final.pdf](https://www.acmg.net/docs/Release_ACMGUpdatesRecommendations_final.pdf).

\*\*\* British Society for Genetic Medicine 2010. Genetic Testing of Children. Available at: [http://www.bsgm.org.uk/media/764635/wgs\\_discussion-paper.pdf](http://www.bsgm.org.uk/media/764635/wgs_discussion-paper.pdf).

Type of finding*	Description	Nature of the information to be fed back	How feedback will happen	Approach to consent or refusal
<b>Carrier status</b> Under certain circumstances carrier status may affect future children.	Autosomal recessive.  Where both parents participate, we will be able to identify a limited list of double carrier states where knowledge might affect future family planning.  X-linked disorders in women carriers.	NHS England will evaluate the impact of looking for genomic findings that are known to cause serious conditions for which there is good evidence that knowing about them could influence healthcare. They will build on the limited list of relatively rare conditions, in the 100,000 Genome Project and adapt this list as more evidence accumulates. The current list will be available on the website	Participants will be asked whether they would like these carrier states to be actively sought.  If they consent to this, information relating to carrier states will be looked for and given to the referring clinician or clinical team for discussion with the participant.  (Without this consent these carrier statuses will <b>not</b> be actively sought.)  When either positive or negative results are confirmed regarding these looked-for findings, a health professional will feed back to the participants.	A list of conditions for opt-in looked-for carrier status.  For double carrier of recessive feedback will be undertaken for those who opt- in, provided <b>both parties</b> in the couple consent to seeking and return of these results as a <b>couple</b> .  Or a woman carrying an X-linked disorder consents to opt-in to feedback.

Type of finding†	Description	Nature of the information to be fed back	How feedback will happen	Approach to consent or refusal
<b>Incidental findings</b> A sub-category of additional findings that are not actively sought. Also described as 'unsolicited' findings****	This is any secondary health-related finding that has not been included in the gene list above.	Incidental findings found in the research resource will only be utilised in the clinical setting once they are assigned to the list of secondary findings above.	<i>Incidental findings</i> will usually not be sent to clinicians or patients by Genomics England.	The policy on not feeding back incidental findings will be explained to the patient at the time of consent.

\*\*\*\* Anneke L notes: Although these terms have no clear definitions and their meanings may be more or less relevant to the different parties involved. For example, a patient may find a secondary finding unexpected, but a laboratory worker analysing a whole genome should expect to possibly find anything.

### **Consent process to receive additional/secondary looked-for findings and carrier status\***

Individual adult participants (or their consultees or equivalent in the form of advice), or a suitable adult (in regards to child participants) will be asked to give their explicit consent to permit additional findings and to receive the results of these. Within the NHS GMS, these will be discussed with the patient or their nominated representative by the clinician requesting the test, or a suitably trained colleague acting on their behalf. For other cohorts, these will be discussed at the point of consenting or when a consultee's advice is sought.

With regards to the lists of secondary and carrier status findings, consent will be given to 'any other conditions that meet these criteria over the life of the Library as these lists evolve.

These lists will be regularly reviewed by NHS England and updated by experts to guide this policy. Changes to the shortlist of conditions would require the support of NHS England and Wales who would need to fund clinical support to patients.

### **The feedback timeframe**

The new NHS GMS will be responsible for determining the timeframe for return of results within the clinical service. For results being returned outside of the NHS GMS the default expectation is comparable feedback times to those for the NHS GMS.

# 7 Programme delivery and management

## 7 Programme delivery and management

### 7.1 Project management

For the NHS GMS and broader genomic developments involving the NHS, a Partnership Board has been established to ensure visibility and co-operation between NHS England, Genomics England and NHS Digital at the highest level. In addition, a Data and Informatics Board convenes monthly to specifically oversee matters relating to the informatics delivery.

The Partnership Board is informed and takes in account other organisation specific governance arrangements.

## 8 Informatics architecture

## 8 Informatics architecture

Identifiable clinical, laboratory and the health data flowing from NHS and other organisations into Genomics England will be pseudonymised prior to storage within the Research Environment. The pseudonymisation process is in line with the latest guidance from NHS Digital and the Information Commissioner's Office.

In the NHS GMS, standard NHS approaches to identifiers will be adhered to. For the research environment, these will be linked to identifiers that enable us to ensure that we continue to return relevant findings from research activity.

### 8.1 Information governance and security

Patient confidentiality and protection is a key cornerstone of the Project and its legacy, the Library. We wish to promote an active Clinical Interpretation Partnership and encourage appropriate and qualified users into the data infrastructure.

A key feature of the Genomics England programme is that individual level data will not be 'released', but will instead be analysed within a secure, monitored environment akin to a reading library. That is, where books cannot be taken away but must be read in situ in a monitored environment. The Genomics England 'Airlock Policy' ensures that the movement of data in, and summary results out of the Research Environment occurs in a controlled and supervised manner, with the goal of facilitating research and discovery whilst maintaining the security of the datacentre.

To protect patient confidentiality, access to this environment will be granted only for specific, approved purposes in accordance with informed consent. Any attempted use beyond the specified purpose may lead to exclusion and possible legal action, where appropriate.

Genomics England will require:

- Appropriate legal gateways for analysis:
  - Under the common law duty of confidentiality, consent to use the data will be gained by a fully informed discussion with the patient as to the use of their confidential data, including additional support materials for them to read and inform their choice (termed 'Patient Choice'); and
  - Under the EU General Data Protection Regulation(GDPR), for the processing of personal data, Genomics England will rely on Legitimate Interests under Article 6(1)(f), and will rely on the research exception in Article 9(2)(j) for the use of special category data (health and genetic data).
- Data sharing agreements which are legally binding will be in place between all parties accessing and using data as part of the Genomics England pipeline. Genomics England will have a simple electronic agreement system which all users will sign who wish to access the data.
- This agreement will include a description of our rigorous data access, sharing and acceptable uses procedure and approval to ensure compliance with governance requirements across the pipeline, and safeguard patient confidentiality and data privacy.

The list below outlines the activities, checks and balances required to ensure Genomics England is processing patient data in an appropriate way.

- Delivery of the service model, whereby customers cannot export raw data (only summary results) and data delivered for research use is pseudonymised.
- Delivery of processes relating to use of patient data (agreement forms and accompanying literature) ensuring that patients understand how we will be using their data and can be confident that we are respecting their data protection rights.
- Delivery of processes to honour the choice of the patient should they wish to dissent to the use of their data for research
- Including seeking REC approval for the protocol, record of discussion forms and patient information sheets.
- Delivery and active management of signed data sharing agreements between Genomics England and all its data suppliers, and Genomics England and its users.
- Continued submission of the NHS Data Security and Protection Toolkit.
- Delivery of data protection impact assessments and equality impact assessment, and accompanying action plans.
- Delivery of information governance continuous improvement plan, and training and awareness for Genomics England staff.
- Delivery of staff contracts with explicit clauses relating to employee responsibilities to safeguard data integrity and security.
- Where appropriate, policies and Standard Operating Procedures (SOPs), based on guidance from the Information Commissioner's Office (ICO) that set out the ways in data may be accessed, handled and disclosed within the secure Data Centre.
- Data access request log and process, and data transfer logs relating to all identifiable data – reviewed regularly to ensure adherence to SOPs
- Delivery of clear data models for use and plans for effective data management.
- Within the research-accessible part of the infrastructure, the ability to link back data. This is to enable updates of secondary data or further information and to ensure the ability to inform patients of relevant findings is necessary.
- Tools for enhanced audit logging and processes to verify customer use of the data and take action if non-compliance is detected.
- Standard OGC ITIL compliant processes for service management, including management and control of live service incidents.

Policies on data access, data integrity and security controls (including encryption of both data and mobile devices, backup, firewalls, anti-virus protection), audit and verification procedures; and formal change management of all security documentation will continue to be kept under review and be developed in line with NHS standards of good practice.

All data transfers and data access requests will be logged and subject to regular scrutiny by the Service Delivery Team. Escalation procedures will be in place to manage non-compliance if detected, based on the data sharing agreements signed by users (or internal staff/supplier contracts in the case of an internal breach).

The Research Environment will provide clinical researchers with managed, audited access to data and computing resources. Activity logs will be analysed on a regular basis, with any exceptions flagged to the Service Delivery Team for further investigation and escalation as required. Additionally, the Genomics England Board will ensure that the organisation establishes appropriate procedures for documentation, tracking, monitoring and reporting for overall process of data sharing, including overview and approval by Caldicott Guardian and Senior Information Risk Owner respectively. Genomics England has also appointed a Data Protection Officer as required by the General Data Protection Regulation

As part of its assurance and approvals plan, Genomics England will seek ethics approval for the research database itself, complete the NHS Data Security and Protection Toolkit on an annual basis and maintain a continuous improvement plan for information governance. In addition to the above IG standards, Genomics England has adopted a wider approach including gaining ISO 15189 for the clinical bioinformatics service.

## **8.2 Data access, sharing and acceptable uses of information from participants**

Patient literature outlines the plans around data access, sharing and acceptable uses in addition to security, third party and participant access requests to data. This will be accessed in a pseudonymised format by clinicians, academics and industry in the UK and overseas. Genomics England is committed to understanding the importance of balancing the responsibility to store and use data safely, with the responsibility to make data available in a way which is usable to drive research and treatment development.

Genomics England takes data security and patient privacy extremely seriously. It will take action where a breach, or near breach, is proven. Penalties for non-compliance include Genomics England's revocation of user access, organisation access and reporting the offending activity to the appropriate regulatory authority. We will where possible, notify individual participants affected if we discover data breach notifications.

All policies, procedures and literature relating to data security and privacy shall be formally placed under change control. Formal policies relating to purposes and uses will be subject to discussion and consultation with stakeholders.

Genomics England welcomes applications for use of its datasets from the commercial, academic and clinical communities, and is keen to create an environment through the GeCIP to enable different kinds of users to work together to share and build new knowledge about the data that can be translated into new diagnostic and treatment options for patients.

The 100,000 Genomes Project commissioned participant, patient group and health professional surveys and interviews around issues of data sharing and access. A number of workshops were held pertaining to the collection, use and sharing of data, privacy, and consent. Genomics England continues to keep participants involved in decisions around data use and sharing by having members of the Participant Panel on the Access Review Committee.

## **8.3 Partnerships to ensure patient recruitment, sample and data acquisition**

Samples and data will be provided primarily through the NHS GMS (details to be established by NHS England) and through cohorts (details to be negotiated with each cohort owner).

#### **8.4 NHS England requirements and specifications – the Library and the NHS GMS**

There are a number of activities required to enable NHS England to deliver aspects of the pipeline described above. These are described or specified in further detail below. Delivery of the Library in Northern Ireland, Wales and Scotland (which is under a separate protocol) is managed through separate Memorandum of Understanding, Services Agreement and other supplementary agreements.

#### **8.5 NHS England requirements and specifications – Genomic Medicine Service**

In 2018, NHS England is procuring (through standard NHS processes) a network of GLHs as one of the key elements of the evolution of the 100,000 Genomes Project to the complete NHS GMS. These organisations will be charged with the delivery of genomic testing services to their relevant region and/or nationally, including the provision of data and samples to Genomics England, operating under a partnership agreement to procure and provide the diagnostic whole genome sequencing requirements.

The GLHs will be responsible for the provision of sufficiently high-quality samples and data to enable whole genome sequencing in accordance with Genomics England's ISO-accredited standards and processes (when achieved) and in line with agreed NHS England contractual obligations.

NHS GMCs will also evolve to become part of the NHS GMS, along with clinical genetics services and other activities around mainstreaming of clinical genetics and will include relevant cancer services using genomic analyses to direct clinical management decisions.

The Central Biorepository function will potentially be provided in several places; by the NHS GLHs/GMCs for the clinical diagnostic service; and by the Genomics England central biorepository for other Genomics England activities.

Throughout the course of the Library, there has been ongoing work to reduce the requirements for DNA in order to improve the viability of WGS in a clinical service. This has been communicated to NHS GMCs and will continue to be communicated within the NHS GMS through sample handling guidance documents, available at [www.genomicsengland.co.uk](http://www.genomicsengland.co.uk) and through other mechanisms being put in place by NHS England.

##### **8.5.1 Sample and DNA acquisition and logistics for NHS Genomic Medicine Service**

Sample acquisition and DNA extraction specifications for the NHS Genomic Medicine Service are depicted below.

After a test request is made on NGIS, primary sample tissue and blood is collected according to existing NHS standards and practices and despatched to the GLH. The receiving GLH will validate the test request and received sample type, and extract DNA to the specifications of the NHS GMS guidance. DNA extracted for all elements of a test (e.g. tumour and germline for diagnostic cancer whole genome sequence) will be stored locally within a GLH/GMC NHS Biorepositories or the Genomics England Central Biorepository.

### 8.5.2 Labelling of samples

Each biological sample sent to the NHS Biorepositories or the Genomics England Central Biorepository from a GLH should be preceded by a sample despatch message identifying the sample as labelled. It should also include additional relevant sample-specific data such as volume, concentration etc.

### 8.5.3 DNA delivery to sequencer (logistics)

DNA samples extracted in GLHs may be plated and/or despatched in tubes directly to the sequencer after approval of the plating/sample proposal, to assure appropriate sample placement and/or provision. Equivalent processes will be in place for other sample and data sources than the NHS GMS.

### 8.5.4 Sample and DNA acquisition and logistics for WGS

Sample acquisition and DNA extraction specifications for cancer and rare disease are depicted below.

DNA extracted from blood and tumour must be sent together to a specified biorepository, following confirmation that clinical data is available that meets Genomics England and/or NHS England requirements as laid out in relevant Genomics England and NHS GLH SOPs.

Throughout the Library, it is expected that Genomics England and NHS England will work in partnership to identify and implement changes to the diagnostic pathways inclusive of molecular pathology within the accredited GLH/GMC centres as part of the NHS GMS to enhance sample quality, in order to optimise the Genomics England pipeline. It should be noted that some of the requirements may evolve.

Genomics England will work with the NHS GMS inclusive of GMCs/GLHs GMCs to finalise the following dataset and collection tool:

- DNA source – germline/tumour
- Standard Operating Procedure version(s) used for collection, extraction, QC and logistics.
- Tumour type (if cancer– primary/recurrent/metastatic)
- Topography and Morphology (if cancer)
- Tumour cellularity percentage (if cancer)
- DNA concentration (nanogrammes/microliter)
- DNA volume (microliters)
- DNA QC metrics (to be finalised with sequencer)
- Genomics England identifier for data subject (site ID, local patient ID (NHS Number), sample ID)
- Gender and ethnicity
- Date of birth
- Pedigree structure and affectation status (rare disease)
- Version of Consent and Patient Information Sheet used
- Syndrome or disease name
- Named site contact and email (to verify receipt of samples by collection hubs, for queries etc.)

### **8.5.5 Key Performance Indicators (KPIs)**

Key performance indicators will be recorded to monitor the process and also to enable the identification of variables that affect sequence quality. These will be agreed within the NHS GMS.

### **8.5.6 Data from sample acquisition site**

Local sample metadata will be required to be submitted in advance and alongside the DNA samples, to enable sequencing and annotation. Data will be submitted in accordance with the NHS GMS requirements or equivalent.

### **8.5.7 Data from national NHS and other sources**

Genomics England expects to collect other health data on its consented patients into its data centre, to link outcomes data to the WGS and clinical data already held. This may involve ongoing data delivery from other organisations, including but not restricted to: NHS Digital, the CPRD or other sources of primary care data, Public Health England, disease registries, screening programmes and patient communities.

All organisations operating within the pipeline need to provide monitoring and reporting functionality to facilitate the tracking of samples and data end-to-end.

## **9 Genomics England data access and acceptable uses**

## 9 Genomics England data access and acceptable uses

Policies for data access, sharing and acceptable uses have been developed in line with the Medical Research Council (MRC) Policy and Guidance on sharing research data from population and patient studies. These will be based on similar policies from the Wellcome Trust Centre for Human Genetics, and input from the Information Commissioner's Office and the Genomics England Ethics Advisory Committee and patient groups.

### 9.1 Users of the Library data

The data access and acceptable uses criteria have been summarised in this section.

Access to data will be controlled according to the provisions of the Data Access policy and all users will sign a Genomics England data access agreement.

#### 9.1.1 Patient access

Genomics England has published on its website a privacy notice that explains how it will respect the rights and privacy of all data subjects.

#### 9.1.2 NHS healthcare teams

Findings will be fed back securely to the identified clinician and/or team responsible for caring for the patients via the NHS GMS/GLHs. These findings will contain patient identifiable data (see above).

#### 9.1.3 Academic researchers and trainees

They will either be a Genomics England Clinical Interpretation Partnership (GeCIP) user or if not, they will pay to access the data.

#### 9.1.4 Commercial researchers and Industry engagement

The Library also covers commercial research. The same rigorous criteria will be applied to data access for industry.

Industry engagement is managed through collaborative fora and direct engagement with companies. This provides a platform for collaboration and engagement between Genomics England, industry partners, academia, the NHS and the wider UK genomics landscape and is open to a range of companies world-wide. The main aim is to bring together multidisciplinary expertise to maximise the potential of the Library dataset and to assure that the main aim of bringing benefit to patients is perpetuated well beyond the Library end date.

In the first instance activities will focus on research to improve diagnoses for the patients (via research discoveries or tool improvement); to increase availability of potentially life-saving clinical trials; and to drive drug delivery from development pipelines to the healthcare.

### 9.2 Data access and sharing for the Library

The purpose of the infrastructure is to provide maximal access to researchers, the NHS and those in training, enabling managed access to the data assets created by the Library. The infrastructure will

be designed to facilitate the sharing and re-use of data and compute resources, and to promote scientific collaboration in clinical genomics research. Software infrastructure – models, interfaces, and tools – will be provided for the generation of ‘analysis ready’ datasets.

All requests for data access will be subject to the following considerations:

- Protection of data subjects (honouring commitments made to them, acting within the scope of consent and according to conditions of Research Ethics Committee approval).
- Compliance with legal and regulatory requirements General Data Protection Regulation 2018, Data Protection Bill 2017, Freedom of Information Act 2000, NHS Act 2006, Health and Social Care Act 2012, the Common Law Duty of Confidentiality, Human Tissue Act 2004 and applicable requirements from organisations affiliated with the Health Research Authority, including Research Ethics Committees and the Confidentiality Advisory Group (CAG).
- Provision of a signed Genomics England data access agreement to the Access Review Committee (see Information Governance above).
- Prioritisation of access according to resource availability.
- Facilitation of high quality research.

Access to the data and compute infrastructure will be subject to approval by the Genomics England Access Review Committee.

### **9.2.1 Data sharing requirement**

For the purposes of this document, data held shall be defined as either:

- i) The data held within the Genomics England Data Centre about participants within the Library (service data), or
- ii) The data held by Genomics England relating to its business, its customers and users, its staff etc. (corporate data).

It is the core business of Genomics England to share service data, within an agreed governance framework, in order to deliver access to treating clinicians which may be identifiable for NHS use (see above). It is also part of our core mission to provide access to researchers and scientists. Such data will be pseudonymised, unless there is a legal basis in place for the disclosure of identifiable information.

Furthermore, Genomics England is bound, through various legislative and regulatory instruments, to deliver data or access to specific requestors for specific purposes e.g. to share data requested under the General Data Protection Regulation, or to deliver the results of financial audits to regulators. For these sorts of requirements, both corporate data and service data might be shared, anonymised or aggregated unless there is a legal basis in place for the disclosure of identifiable information.

As previously noted, sharing of data for prescribed purposes is key to the Genomics England business model. However, decisions on data sharing must always be assessed against the risk to

data confidentiality and the privacy of data subjects. All decisions relating to data sharing should be made considering the policy provisions enshrined here.

Where risks to data sharing fall outside of the tolerances set out on the agreed SOPs, Genomics England shall work with its legal counsel and information governance and data security leads, to deliver a transparent decision-making process and associated assurance, validation and reporting.

### **9.2.2 Access to service data**

All data held within the Genomics England Data Centre that relates to the Library data subjects is considered to be service data.

Service data will be made available within an agreed governance framework in two ways:

- i) With internal staff (under substantive contract of employment) within Genomics England, in order to check data quality and integrity; to undertake agreed cleansing activities; to deliver initial bioinformatics analytics on the data; and to undertake other monitoring, audit and service improvement activity. It is noted that internal access to all data held shall be controlled, monitored and reported upon. Access to identifiable data, which is held separately, must be explicitly logged and approved by the nominated Genomics England Information Governance Team.
- ii) With external users and consumers, as follows:
  - a. In the form of proactive publication of pseudonymised aggregated information as summary metadata.
  - b. Delivery of access (within the Genomics England Data Centre) to agreed datasets for approved users, including scientists undertaking research into whole genome data. Data will be pseudonymised unless a legal gateway exists for access to identifiable information.

The aggregation of rare and cancer genomic data at international levels gives scope not only to resolve disparities in clinical interpretation<sup>xx</sup> but also to place patients' data in the correct phenotypic and genomic contexts, including with respect to ethnicity.<sup>xxi</sup> The Library presents a unique opportunity to optimise patient benefits, while minimising the risks of sharing potentially identifiable personal data, as is commensurate with global initiatives for data availability and under the standards laid out by the Global Alliance for Genomics and Health.<sup>xxii</sup>

Summary-level data, including the publication of allele frequencies, must be made available to clinical and research communities to aid genomic analyses. Individual-level data can be shared where there is the potential for individual clinical benefit. Further details can be found in Airlock documentation.

In the area of case-solving, routine clinical practice and governance applies.

### **9.2.3 Corporate data sharing**

Corporate data is defined as data that is held by Genomics England that relates to the work and operations of the business, including its organisation, its people, its performance and its decision-making. Certain corporate data may be shared, within an agreed governance framework, in order

to meet statutory and regulatory requirements. This includes the proactive publication and maintenance of information about the operations of the business, in line with the provisions of the Freedom of Information Act.

### 9.3 Acceptable research uses

All users and uses must be consistent with the terms of informed consent policy, procedures and literature pending REC approval.

#### 9.3.1 Types of purposes

The following list of purposes will be used by Genomics England to classify the type of work its users want to undertake. Under the terms of ethics approval, Genomics England plans to operate an internal process to assess applications and grant access to data that relates to any of the purposes below.

Definition	Explanation
Clinical care	Use of data to directly diagnose / treat a patient.
Hypothesis driven research and development in health and social care - observational	Hypothesis driven review and analysis of data collected within the Genomics England data centre.
Clinical trials feasibility	Queries on the broad location and numbers of patients eligible to participate in interventional research project.
Hypothesis driven research and development in health and social care - interventional	Hypothesis driven review and analysis of agreed cohorts of patients within a research study.
Non-hypothesis driven R&D - health	Review and analysis to identify associations between data and variables - to improve our understanding of the causes of disease.
Non-hypothesis driven R&D - non-health	Use of the data by data scientists to test the effectiveness of tools and models for data analysis.
Public health purposes	Work to gain a population level understanding of the prevalence of disease and corresponding health protection actions where possible.
Other health use - clinical audit	Use of the data to assess whether clinical standards are met.
Other health use - commissioning / commissioning policy	Use of Genomics England data to inform decision-making by any relevant commissioning body / organisation developing commissioning policy. Trust use to undertake local planning and administrative tasks.
Subject access request	Request by the data subject regarding the data Genomics England holds about them.

<b>Tool and test evaluation and improvement</b>	Use of the data to validate, improve and deliver new analysis and interpretation tools and tests.
<b>Interpretation and validation of the Genomics England Knowledge Base</b>	Use of the data to interpret annotation findings and validate results for use in a clinical context.
<b>Education and training of health and public health professionals</b>	NHS professionals and others learning how to analyse, interpret, and apply genomic medicine.
<b>Deeper phenotyping</b>	Gaining further information on patients with particular genetic results.

### 9.3.2 Uses that the Library will decline

These types of purposes might be proposed by participants or third parties. They will typically be refused outright following consideration. This will be by the Genomics England Informatics Team and the Genomics England Access Review Committee who will decide upon approval or rejection on a case-by-case basis.

Additionally, all requests that are required by court order will be notified to the Genomics England legal counsel and the Executive team.

The Genomics England decision-making process on all occasions shall be documented and transparent and involve external bodies (e.g. REC) where necessary.

- Paternity testing and level of relatedness testing (though participants can request their data and, as such, if several family members make this request and take the data to a third party to analyse, it is conceivable that the data can be used in this way).
- Parental request for child's data once the child reaches 16, where either the participant assents or dissents.
- Requests in the form of Court Orders will be referred to Genomics England's Legal Counsel as promptly as possible, so that all representations may be made to the court, for example to limit the information requested.
- Uses requiring linkage to employment records, tax records, benefits records or personal life insurance records for non-scientific or non-healthcare related purposes.
- Uses requiring linkage for personal insurance or forensic purposes.
- Purposes that might lead to discrimination against a person or a group of people, based on genetic/ genomic characteristics.

# 10 Genomics England data ownership and intellectual property



## 10 Genomics England data ownership and intellectual property

Genomics England has received legal advice regarding intellectual property. This recommends that Genomics England owns the combination of the whole genome sequence and the clinical data for the entire dataset from the Library. In addition, Genomics England owns any new intellectual property generated from the data, but we will license this to third parties on favourable terms.

There are clear reasons why this is essential:

- It ensures that Genomics England Clinical Interpretation Partnership (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.

### 10.1 Genomics England publication policy

Genomics England encourages publication through our publication policy and process.

Acknowledgements in all publications will recognise the contribution of the Department of Health and Social Care, 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 Social Care 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.

# 11 Consent

## 11 Consent

Achieving informed consent in the context of clinical genomics is recognised to present important practical challenges and requires careful ethical consideration. This is also true for the design of the Library and NHS GMS, an initiative combining standard clinical care with research and translational medicine aims.<sup>xxiii,xxiv</sup> The Library will follow established legal and regulatory standards for seeking the informed consent of its participants, or those who could consent on their behalf. It will follow applicable legal and regulatory standards around the potential or actual involvement of adults who lack capacity, including seeking the advice of appropriate consultees regarding their participation, including in regards to participants who join the Library and thereafter are deemed to have lost capacity.<sup>xxv</sup>

The earlier stages of the Project are now generally recognised as good-practice approaches to obtaining consent. Building on this and the latest guidance for the NHS GMS we will continue to iterate and develop our approach.

This has also been informed by the Chief Medical Officer's Report, 'Generation Genome,' (see <https://www.gov.uk/government/publications/chief-medical-officer-annual-report-2016-generation-genome>), published in July 2017. Professor Mike Parker's letter to the Chief Medical Officer from the 100,000 Genomes Project Ethics Advisory Working Group (now renamed the Genomics England Ethics Advisory Committee) sets the standards which we continue to uphold. This can be found on our website: [www.genomicsengland.co.uk](http://www.genomicsengland.co.uk). Professors Mike Parker and Anneke Lucassen have also provided guidance in the form of a letter outlining principles on which to base the consent model for the NHS GMS.

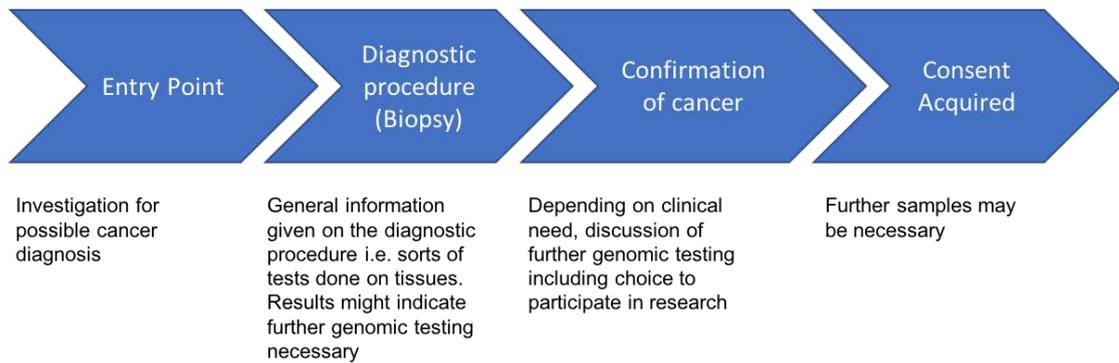
### 11.1 The process for consent in the Library

In the NHS GMS, the patient will be given the choice of germline genomic testing which includes the opportunity for them to contribute their data and samples to a research resource at the point at which they have the test. For patients having somatic cancer testing it is envisaged that consent for research may happen before or after this testing to best suit the needs of the patient. In conventional terms, as in the current 100,000 Genomes Project, a broad consent model is proposed whereby permission is given for future research, the exact nature of which may, as yet, be unknown.

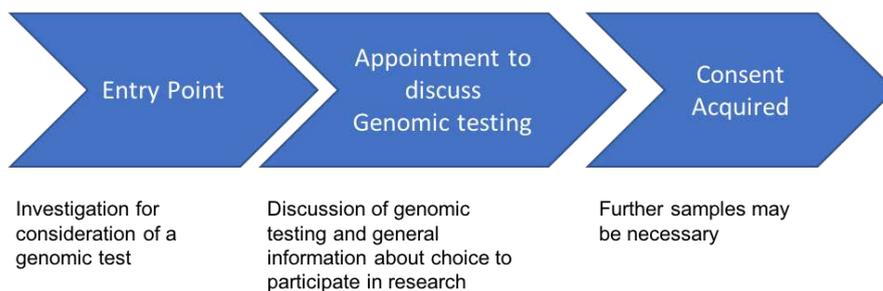
The materials supporting the consent process that will be presented to the patient will be layered; a minimum level of critical information will be made available with access to other more specific and detailed information as required by the patient.

#### **Process for consenting a patient:**

## CANCER



## RARE DISEASE



In circumstances where it is not appropriate to offer the choice to participate in research (for example unconscious patients, urgent tests required in a seriously ill child etc.) the offer of a genomic test will be a clinical decision and the choice to take part in research will be offered later when appropriate. The offer of the choice to take part in research is integral to the NHS GMS and will be audited at the hub level.

More detailed process maps are available in Appendix 6.

Patients under the care of the NHS in England will be referred by their treating clinician for invitation to join the Library because:

- a. they have a (suspected) rare disorder or certain forms of (suspected) cancer and joining the programme may reveal important information about their health. Or
- b. they are undergoing any genomic test found in the NHS England Genomic Test Directory
- c. they are part of a cohort who have had samples and data taken for research or as part of clinical care, but outside of the funding and structures of the NHS Genomic Medicine Centres (see section 3.3). In such cases, they may also be consented under a separate consent model, but this should be aligned with that of the Library (e.g. the Scottish arm of the 100,000 Genomes Project, which is under a separate protocol)

Eligible relatives of probands in the above categories may also be invited to join the Library to inform research regarding the health of the proband. Deceased participants may be included with the support of their treating clinicians (or their relatives), where appropriate phenotypic information and DNA samples are available and appropriate consents are given. His or her relatives may also join the Library in order to gain information about the deceased proband's health condition.

Research participants agree to:

- Deposit data (and samples) for use in approved research according to acceptable uses in linked, pseudonymised format.
- Re-contact to invite to further research (including re-contact by genotype) to ask for further data or samples, or offer to invite to further research including clinical trials.
- The return of clinically applicable research results – i.e. potential individual benefit from research participation.

As in previous phases of the Project, patients can change their choices over time (discussed below).

Examples of recontact include clinical trials opportunities. Currently, participants eligible for potential clinical trials are identified through the Library. The schematic below outlines the current model workflow for delivery. This process is currently in the pilot phase.



**Clinical trials referral workflow**



**11.1.1 The consent form and patient information literature**

Drawing on relevant guidance and exemplars of good practice,<sup>xxvi</sup> the participant materials (i.e. both digital and hard copy versions of letters of invitation, Patient Information Sheets, and assent and consent forms) have been developed in accordance with the key principles in the Key Ethical Principles document and Ethical Governance Framework, and with the advice of the Ethics Advisory Committee, with comments from the Science Committee, and other stakeholders and the appropriate guidance.<sup>xxvii</sup>

The consent discussion with participants will draw attention to key elements of the Library that may be regarded as having particular ethical, legal or social implications. This discussion will include seeking explicit consent to:

- Linked access to participants’ health records in perpetuity.

- Access to deidentified participant information and samples by third parties with the appropriate permissions (who may be based in the UK and overseas) and which includes commercial ('for-profit') companies.<sup>xxviii</sup>
- Agreement to recontact by their clinical team (including recontact by genotype), in order to permit invitations to participate in further studies (including relevant drug trials, or ethics and social research to do with participating in the Library).

The patient literature is also written so that the discussion can draw potential participants' awareness (or that of an adult consenting on behalf of a child, or that of a consultee advising in respect of an adult participant who lacks capacity) to key issues or risks from joining the Library, including that:

- For patient groups needing particularly time-sensitive treatments, the results of genomic analysis may not be able to be returned to their clinician in a timeframe that can influence clinical decision-making around treatment for their pertinent condition.
- In general, at this stage in scientific understanding, the significance for health and disease of information revealed by WGS may be relatively uncertain (particularly in the absence of a relevant family history).

This knowledge has improved and is expected to continue to do so, but findings of unknown significance are likely to be generated. These will not be returned to the clinician for discussion with the patient unless there are exceptional circumstances.

- Should a specific and serious risk to the health of a participant's close family member(s) become apparent via the participant's pertinent or additional findings (whether these are unexpectedly detected or purposely 'looked-for') and where preventive or ameliorative action is possible in regards to these close relatives, the clinical team may seek to contact the relatives to inform them of this in line with standard clinical practice. This may include risks identified after the participant has died.<sup>xxix</sup>
- Whilst measures will be in place to keep participants' identity and information confidential, there is an unavoidable, though remote, risk of re-identification through participation.<sup>xxx</sup>
- The ways in which data security, privacy, access and sharing access to samples and the datasets will be appropriately restricted and monitored.
- Participants can request to see the information held about themselves by the Library and to know the types of third parties who have accessed their data (taking into account all legal obligations under Freedom of Information or Data Protection).
- The process for withdrawing consent to remain in the Library.
- Given the long-term nature of the donation, in future, the samples and information donated may be subject to the application of as-yet unanticipated analyses and investigation because of new knowledge.

### **11.1.2 Giving consent face-to-face**

Generally, consent to participate in the Library will be part of the clinical conversation (in most instances) by trained members of staff. This discussion will conform to established good consent practice guidelines.<sup>xxxi</sup> Where possible this should be timed to fit in with scheduled appointments. During the consent discussion regarding research, potential participants and those accompanying

them will have the opportunity to ask questions. All potential participants (or people consenting on their behalf) will have been given appropriate time to read and consider the Library's patient information literature. Literature may be translated by the NHS Genomic Medicine Centres (undertaken forward and back).

In order to record their consent, the potential participant (or the person consenting on their behalf, or the person advising about someone else's participation) will need to follow the instructions on the consent/ consultee declaration form, tick the boxes, and sign and date the final section, adding their own date of birth as an additional confirmation of their identity.

#### **11.1.3 Giving consent at a requested home visit**

It is ideal if all consent and procedures take place in a suitable healthcare environment. If a participant (or the person who might consent on their behalf) requests that the consent discussion takes place in their home (or in another appropriate healthcare setting) it will be a decision for the NHS GMS as to whether this can be provided as an option. If possible, an appropriately trained team member(s) will schedule an appointment convenient to the potential participant(s) and will travel to the agreed location or telephone the patient.

#### **11.1.4 Giving consent by post or email**

In appropriate cases, some potential participants may be offered the opportunity to give consent to be part of the Library via the return of a completed consent form to the recruiting clinician or appropriate alternative in the post/via emailed document. A telephone discussion with a suitably trained person will be offered to them as part of this, to take place at a time suited to the potential participant.

This route to seeking consent will only be used where the clinical team has reason to think that seeking postal consent is an appropriate method for the individual concerned and that consenting face-to-face would inconvenience the patient.

#### **11.1.5 Participants aged 16-18 years**

At ages between 16 and 18, it is legally presumed that young people have the ability, or competence to make decisions about their own health care or treatment. When a young person is believed to be competent, consent from those with parental responsibility is not legally required.

On reaching 16 years old, all existing participants will be asked to give their own consent to remain in the programme (unless it is deemed by their medical team that they do not have the capacity to do so at that time). This also applies to those being invited to join the programme at the age of 16.

In most circumstances, young people consenting on their own behalf will be encouraged to involve the person(s) with parental responsibility for them in their decision-making.<sup>xxxii</sup> Some participants in the Library who join at younger ages, will then reach the age of 16 as participants and at that point will be deemed not to have capacity to consent (or to refuse), their continued participation in the Library on their own behalf. In such cases, the advice of a personal or legal consultee will be sought regarding their continued involvement, in accordance with the provisions of the Mental Capacity Act 2005 for people in research studies in England and Wales.

### **11.1.6 Participants under the age of 16 years**

Where a potential participant is aged under 16 and has not been deemed by their clinical team to have attained the capacity to consent to participation on their own behalf, at least one person with parental responsibility for them, or a guardian, will need to provide consent in order for the minor to participate in the Library.<sup>xxxiii</sup> The Library will follow appropriate guidance as to the involvement of children in research.<sup>xxxiv</sup>

Patient information will be available on an age-appropriate basis for children and young people.

All children and young people will be given verbal information about the Library by the person conducting the consenting discussion (who should have experience of working with children/young people, in accordance with the minor's age and individual circumstances) and will be given opportunities to ask questions.

The person taking the consent should seek the assent of the child as part of the discussion regarding their participation. If the child or young person is capable of assessing the information provided, they will consider their explicit wishes, including their refusal to take part, or desire to withdraw from the study.<sup>xxxv</sup>

Seeking assent implies that children could have an understanding of the research process and can be informed about what is involved and what they are expected to do. This can be an opportunity for some children to express their opinions and concerns surrounding participation in research, providing them with a formal means to be included or excluded. Where appropriate, the person taking the consent may wish to offer the child the opportunity to evidence their assent via the completion of an Assent Form. These are provided for the use of 6-15 year olds in the Library.

At the time this discussion takes place, the child or young person may not be in a position to, or may not wish to, complete this form. During any discussion around assent, (regardless of whether an Assent Form is offered or completed), a completed Assent Form will not be considered by the person taking the consent to be the sole or overriding signifier of a child's willingness to participate in the Library.

If, at an age below 16, a young participant has been deemed competent by a medical team to consent in their own right to remain in the Library, they will be invited to give this consent. The same will apply to young people being invited to join the programme under the age of 16, who have been deemed competent by a medical team to consent in their own right to joining the Library.<sup>xxxvi</sup> Young people consenting on their own behalf under the age of 18 will be encouraged to involve their parent or guardian in their decision, as a matter of good practice.

### **11.1.7 Seeking consent where participation relates to a health emergency**

In specific scenarios and in line with the appropriate guidance, consent may be sought to join a person as a proband into the Library while they are in a health emergency, such as in the case of an extreme response to sepsis.<sup>xxxvii</sup>

#### **Adults not able to consent for themselves due to an emergency situation**

In a health emergency situation, the law in England and Wales allows adults unable to consent for themselves to be recruited into a project without prior advice from a consultee, specifically in

regards to that emergency situation. Therefore, and on the advice of the patient's clinical teams, the Library will accept as participants eligible adults who lack capacity due to an emergency situation.

This will include where the timeframe for recruitment does not permit the appointment of a consultee before including the patient into the Library as a participant. This exceptional route to participation is proposed because the potential or actual clinical benefit to be gained by the patient if he or she is included in the Library is likely to depend on their prompt recruitment. Participation may provide clinically relevant information to inform the patient's care and may also inform the care offered to the patient's close contacts (for example if the health emergency was caused by an infectious disease).

Recruitment would only take place prior to the appointment of a personal or nominated consultee appointed under the principles of the Mental Capacity Act 2005:

- If it is not reasonably practicable to seek advice from a personal or nominated consultee, and
- The procedure has been approved by an NHS Research Ethics Committee; and
- Provided a consultee is consulted as soon as possible to seek advice on the participant's likely views and feelings on being included in the Library.<sup>34, xxxviii, xxxix</sup>

When the patient's consultee is appointed and consulted, they will be given sufficient information about the Library in an understandable form, to permit their informed advice. They are free to decide to provide this advice or not. Advice given by consultees will be recorded by the person seeking the recruitment on a Consultee Declaration form (not a consent form). If the adult then recovers their capacity, as soon as possible they will be asked for their consent to participate in the Library on their own behalf.

The clinical team caring for the participant will also offer the participant themselves with information, appropriate to their current state, about the Library and its risks and benefits.<sup>xi</sup>

### **Children who are unable to consent for themselves and in an emergency situation**

In an emergency situation, in England and Wales, the law would allow children and young people to be recruited into the Library specifically in regards to that emergency situation, without prior consent from at least one person with parental responsibility or a guardian. These exceptional circumstances for recruitment would include:

- Prior NHS Research Ethics Committee approval for the Library.
- The same research question could not be assessed by recruiting from a non-emergency environment, and that the research is of potential benefit to the child/young person themselves.<sup>xii</sup>
- Someone with parental responsibility for the child/young person is informed about the research as soon as possible.
- Their consent (and the child/young person's assent if they are able to give this) is sought as soon as possible.
- The person seeking the recruitment (who must have previous experience of working with children/ young people) makes it clear to the child/young person or their parent (if the

child/young person is not competent) that the child/young person can withdraw (or be withdrawn by their parent) at any time and without needing to give any reason.

Adults consenting on behalf of children/young people during an emergency will be informed at the time they give consent, that if and when, the child/young person attains/regains capacity, their consent will be sought for participation as would normally be the case.

### **Participants who regain capacity during their participation in the Library**

In all circumstances, any person who attains (or regains) the capacity to decide whether or not to become a participant in the Library will be asked to give us this consent in their own right as soon as is practically possible.

Those acting as consultees for people whilst they were lacking capacity will be informed at the time they gave their advice, that should the person become a participant in the Library and then regain their capacity, that their consent will be sought for participation in their own right.<sup>xiii</sup>

### **Recruiting relatives of the proband in emergency situations**

Adult or child relatives of the proband would not be recruited if they themselves were in a health emergency situation, prior to the appointment of a personal or nominated consultee, unless in exceptional circumstances. Consent would instead be sought for their participation (or advice from a consultee as appropriate) in the normal way, after the health emergency has passed.

### **Participants in the Library who lose capacity to consent**

Participants who have joined the Library by giving consent for this in their own right, may afterwards lose their capacity to give their informed consent to take part in treatment or research, or this capacity may become fluctuating.

The consent they gave to become part of the Library, at the time when they had capacity fails at the point it is reported to Genomics England that they cease to have this capacity. Once they become aware of a loss of capacity, the NHS clinical care team should seek to appoint a personal or (if unavailable) a nominated consultee to advise on the patient's wishes, in line with the Mental Capacity Act Code of Practice published by the Office of the Public Guardian.

In order to safeguard the interests of adult patients who join on the basis of their own consent, Genomics England will ask the NHS GMS to implement a specific check on each participant's capacity (in cases where a participant has consented to join the Library in their own right), around five years after their consent was first given to join the Library. This routine check should be done via appropriately trained staff on the clinical team at the closest appropriate clinical contact time to that five-year interval, so as not to cause undue burden to the participant or to their clinical team.

Then, for as long as the person remains in the Library under their own consent, the NHS GMS will then be sent a prompt to instigate further routine capacity checks via the clinical team at (approximately) five-yearly intervals. This instruction to re-check from the NHS GMCs will be prompted by a reminder issued from the organisation running the Library. This obligation on Genomics England would also pass to any successor organisation with the responsibility for the

ongoing collection of data for the Library. An obligation to participate in this process has also been included in the commissioning of NHS GMS.

Under the Mental Capacity Act as applied in the research setting, we note that there is no legal obligation on researchers to exhaustively ascertain whether or not a participant has lost capacity. For example, if a patient becomes untraceable and falls out of contact with their clinical care team whilst they are also a participant in this programme, it would not be required that Genomics England should require the participant's clinical team to take measures to seek to restore contact with the patient for the purposes of making this capacity check.

In accordance with the Mental Capacity Act, Genomics England will continue to assume that an individual participant retains capacity, even where a capacity check is inconclusive in relation to that particular individual. Examples of this might be that the participant does not respond to the invitation to this check, or he or she does not wish to take part in this discussion, or if a participant becomes uncontactable to the clinical team.

Capacity may also be assumed where the health care professional tasked to conduct the check by the NHS GMC does not return any results from the check to the NHS GMC (to then return to Genomics England).

However, where results have been ascertained by the NHS GMC as to the capacity of any participant, whether it is via the five yearly capacity check or as determined in the course of routine clinical care, the NHS GMC must pass this information on to Genomics England. The NHS GMC has discretion as to whether or not to pass this information either directly to Genomics England upon receipt, or at the five yearly point that it is requested.

If a clinician wishes to notify the NHS GMC of a participant's loss of capacity or fluctuating capacity at any intervening point they may of course do so. In normal circumstances this would then prompt the seeking of a personal (or if necessary, nominated) consultee in line with the Mental Capacity Act Code of Practice, published by the Office of the Public Guardian. At the point that Genomics England is informed that the participant has lost capacity, the individual will immediately be removed from the Library as a participant.

If an adult participant is discovered to have lost capacity, (and where no personal or nominated consultee has been appointed) clinical teams must advise their NHS GMC of this fact. The Genomic Medicine Centre will then inform Genomics England of this, in order for Genomics England to remove that individual as a participant in the Library forthwith. For example, this means that prospective data collection in relation to that individual will cease once this notification is received. However, data already held by Genomics England under prior consent can still be used in accordance with the terms of the consent. In accordance with the principles of the Mental Capacity Act 2005, as applied in the research setting, the affirmative advice of the personal or nominated consultee is then required before the person lacking capacity is able to re-join the Library. If the consultee advises against this, clearly the person may not re-join the Library as a participant.

The advice of the consultee that the person lacking capacity should re-join the Library as a participant will, for example, trigger the resumption of prospective data collection regarding that

individual. As with any other participant, this will also permit data collection from recorded information covering any period during which the person was not a participant in the Library.

The advice of the consultee is also required in each instance in relation to any Library activities that both:

- involve an adult participant who is deemed to lack capacity, and are
- physically intrusive, such as the taking of a blood sample (unless this activity is taking place within a health emergency situation, at which point usual principles of clinical care would prevail).

Where the consultee advises that in this instance the participant would not have wanted to undergo the physically intrusive activity that is proposed, then that activity must not be carried out.

In line with the provisions of the Mental Capacity Act as applied in the research setting, should a consultee advise the participant's clinical team that the participant would have wanted to be withdrawn from the Library (or should the participant recover their capacity and withdraw themselves), this will begin the process of withdrawal without further delay.

A participant's family may sometimes have a different opinion about their participant relative's continued participation in the Library after their relative has lost capacity. This will be handled sensitively by the participant's clinical team, with relatives being encouraged to respect their loved one's wishes, which were given at the time when they were freely able to act upon these. However, we note that Genomics England has a legal obligation to carefully consider a request made by a representative to withdraw a participant.

Details for how to withdraw are clearly noted on the Patient Information Sheet and will be publicly available on the Library's website.

If a participant becomes aware that their capacity to give full and informed consent is likely to lessen, or that their lifespan is likely to be shortened, and they would like to withdraw from the Library whilst they have the capacity to make this decision for themselves, they can withdraw at any time, without being required to give a reason. Participants may or may not also choose to take this opportunity to also inform the clinical team of the names and contact details of individual(s) whom they would like the clinical team to approach as personal consultee(s) in the event that they lose mental capacity. If a participant wishes to 'lodge' this information with their care team in circumstances where they are not considering their imminent or future withdrawal, this is also acceptable.

In Wales, the 100,000 Genomes Project is initially set up as a service agreement between Genomics England, Cardiff University, Cardiff & Vale UHB and the Welsh Government. Cardiff & Vale UHB, supported by Cardiff University who secured MRC funding, is leading the operationalisation of the Library and will support any capabilities checks and resulting re-consenting procedures if necessary. Ongoing capability will be assured by the referring clinical geneticist whenever new samples or data for the Library are provided or when due to changes to the scope of the Library, continuing consent has to be ensured. The final structure of the NHS GMC will be determined during the term of the service agreement to allow long term support of the Library, this will be outlined in the service agreement.

### **11.1.8 Seeking consent to use samples and information regarding a person who is deceased**

#### **Library participants who die**

Some participants will join the Library giving consent for this in their own right and will then die at a later date, whilst they are still participants in the Library. In line with the Human Tissue Act 2004 (in England, Wales and Northern Ireland) and HTA Code of Practice on Consent, (v14.0, July 2014) the participant's consent to join the Library would remain valid even after their death. This provision is explicitly outlined to potential participants at the time of consenting. It is in place to maximise medical researchers' access to deceased participants' samples and health and personal information, thus maximising the potential for increased medical knowledge far beyond the participant's own lifespan.

Although the participant's consent extends beyond their death (e.g. to the collection of recorded health data that emerges after their death or to continued access by approved researchers to the appropriate parts of the participant's information held by the Library), the participant's relatives (including those in qualifying relationships) may sometimes have a different opinion after the participant has died. This view will be handled sensitively by the clinical team of the deceased person with relatives being encouraged to respect their relative's wishes (or in certain cases, to respect the decisions given by the deceased person's nominated representative/nominee).

We note that the decisions made by a nominee of this kind, appointed by the (now deceased) participant to take decisions after the participant's death, cannot be overridden by others, even those in qualifying relationships with respect to the deceased person. We note the Genomics England's legal obligation to consider a request made by a representative to withdraw a deceased participant from inclusion in the Library.<sup>xliii</sup>

It is also explicitly outlined at the time of consenting, that should a specific and serious risk to the health of their close family member(s) become apparent via the participant's pertinent findings, or additional findings and where treatment exists, then the clinical team are likely to seek to contact these family members, to inform them of this. This may include risks that are identified after the participant has died.

#### **People who die before giving consent to join the Library: legacy collections**

There may be specific situations where we would like to include samples and available health information from people who have already died, but who are not already participants in this Library. This will generally be in situations where the equivalent information is difficult or impossible to obtain from a living person within the anticipated time frame of the Library. This may include where a deceased person affected by one of the disease groups included in the Library has left biobanked samples or health information as a legacy for use in further medical research.

Legacy collection samples may benefit living patients via inclusion in our Library for example by increasing the number of samples available to researchers within a particular disease group. For example, where a child participant in our Library has a particularly rare form of cancer, other

participants with the same condition are unlikely to be found in the time frame of the Library. Biobanked samples and information can therefore be particularly precious to researchers in these situations. We note also that the deceased patients who donated these samples were clearly keen that they should be used to potentially help others after they themselves have died.

The inclusion of relevant legacy samples in the Library will only be considered where likely to add substantially to knowledge and potentially offer benefit to patients. The aims and uses to which the Library would put these samples and information would need to be agreed as being scientifically and ethically in accordance with the broad consent given by the patient who donated them to be bio-banked. This would be agreed by the relevant parties holding the samples and information and by the Genomics England Science Advisory Committee, the Ethics Advisory Committee and the Genomics England Board.

### **Adults who die before giving consent to join the Library: seeking consent where a person has (perhaps very recently) died**

Where a person with health problems that may relate to the disease groups in the Library, has died without receiving a genetic diagnosis for their condition, and their living relatives might benefit from knowing this, we will consider the inclusion of an appropriate sample and their information into the Library, on the advice of their clinical team.

Where the adult concerned has not left specific directions to cover this eventuality, consent from the appropriate person(s) who are in a 'qualifying relationship'<sup>xliv</sup> to the recently deceased person at the time of their death will be sought, in order to allow us to include the deceased person in the Library in accordance with the Human Tissue Act.<sup>xlv</sup> This consent will be sought where:

- The deceased person had not previously specifically consented to join the Library before they died (and also, they had not specifically refused, or withdrawn from participation).
- The deceased person's clinical team would like to include their information and tissue in the Library which will require the use or storage of tissue removed from their body after death.
- Genomics England supports the inclusion of the deceased person's samples and information into the Library. The consent given in regards to the deceased adult's participation in this Library should be recorded using the Consent form in respect of a deceased adult and corresponding Participant Information Sheet in respect of a deceased adult.

An adult who has died in England, Wales and Northern Ireland, may have already nominated someone to represent them after their death and to give consent on their behalf. Those close to the deceased will be asked whether this is the case. If there is such a nominated representative, he or she will be asked by an appropriate person from the deceased person's clinical team to give consent in regard to the deceased person joining the Library. This nominated representative's consent must not be overridden by other individuals, including those who were in a 'qualifying relationship' with the person immediately before their death<sup>xlvi</sup>

In cases where an adult has recently died, their nominee or the person(s) in a qualifying relationship to the deceased person would be approached in a sensitive and timely manner by the healthcare professional seeking their consent. This trained person will use their clinical judgement

to decide whether it is appropriate to make this request. They will explain why the deceased person's samples and health information might produce pertinent information which may benefit others in similar situations. Research use of the samples taken from their loved one will be discussed, as well as whether these samples were taken as part of their standard clinical care while they were alive, or if necessary, will be gathered by removing a sample (blood or tissue) from their loved one's body after death, for use in this Library.

Additional looked-for findings will not be offered where the adult in question is deceased.

The healthcare professional seeking this consent is likely to be from the deceased person's clinical care team, or to be a public health professional who is already communicating with the family around necessary public health actions e.g. after a death from invasive meningococcal infections, identifying 'close contacts' and organising appropriate measures for them (such as antibiotics).

**Children who die before giving consent to join the Library: seeking consent where a child has (perhaps very recently) died or after a stillbirth, miscarriage or termination of pregnancy at any gestation**

If in connection with a (suspected) rare genetic disease, a child<sup>xlvii</sup> or a baby has died or he or she was stillborn (i.e. born showing no signs of life and after more than 24 weeks gestation) or where a pregnancy has ended in miscarriage (i.e. at less than 24 weeks gestation) or a termination of pregnancy; it can be very important to the bereaved parents that the reasons for their loss can be sought by doctors and scientists and that they can be informed of any pertinent results.

Some parents may therefore wish for their child's or foetus' samples and information to be included into this Library, in order to look for a diagnosis of the condition which has affected them, and perhaps also to inform the parent's future reproductive decision-making, or that their close family members'. We use 'child' as the descriptor in this Protocol and in the patient literature to refer to a deceased child, baby or foetus at any gestation, although the clinical team should take care to use the same terms used by the woman or couple in conversation with them.

Provided that the foetus or child's health problems may relate to the disease groups in the Library and the clinical team proposing this inclusion feel that the living relatives might benefit from being able to gain further information in connection with their loss, and where an appropriate sample(s) can be gathered, we will consider the inclusion of the sample(s) and related information into the Library, taking into account the advice of the child's or parent's clinical team.

The mother or parents may also wish for carrier status testing to be carried out as part of their own participation in the Library e.g. as part of a 'trio' in which the parents enter the Library together with the samples and information of the foetus, baby or child who has died. Consent to this carrier testing should be discussed separately as part of the consent process for the parents regarding their own participation and is to be recorded in their own (adult) consent form.

Even where the miscarriage or termination took place when the baby or foetus (depending on the terms the parent wishes to use) was at under 24 weeks' gestation, the Consent Form for parents of a (deceased) child with a rare genetic disease should be used to record the consent of the mother with legal parental responsibility for their baby or foetus' samples and information to be included in

the Library. The mother giving the consent should be asked to read the corresponding Participant Information Sheet for parents of a (deceased) child with a rare genetic disease.

While making a record of the woman's consent for the inclusion of her baby or foetus (who died at under 24 weeks' gestation) into the Library would be legally sufficient as an entry into her medical notes,<sup>xlviii</sup> unless for some reason the clinical team find entering her consent into her medical notes to be the only appropriate course of action, we would expect the Consent Form for parents of a (deceased) child with a rare genetic disease to be used to evidence her consent. Taking this separate consent is important because it fits in with existing NHS clinical practice in respect of prenatal diagnosis more generally, and also because it makes it explicit that the woman has consented for the analysis of the foetus or baby's sample.

Having the two signed Consent Forms (one for the woman herself and one for the foetus or baby) also allows each sample to be tracked individually with its consent which is a standard requirement of this Library, including more straightforwardly facilitating, e.g. withdrawal of the foetus or baby's samples and information from the Library if requested.

The law requires that where a foetus at above 24 weeks' gestation, or a baby or child aged under 18 years has died, their remains should be handled in accordance with provisions for gaining consent for the use of the tissue of the deceased. The consent required here needs to be given by an adult with legal parental responsibility for the child who has died. Consent to permit the deceased child's inclusion in the Library should therefore to be evidenced by this adult using the Consent Form for parents of a (deceased) child with a rare genetic disease and after reading the corresponding Participant Information Sheet for parents of a (deceased) child with a rare genetic disease.

The consent of one person with parental responsibility is required for a child to be included as a participant in the Library after the child had died (above the age of 24 weeks of gestation), although the consent discussion with the clinical team may involve other family members as appropriate. If no one with parental responsibility for the child is available, then consent will be sought from someone in a qualifying relationship with the child.<sup>xlix</sup>

In all cases where a child or baby has recently died, or the pregnancy has recently ended in the miscarriage, stillbirth or termination of pregnancy, the appropriate family member(s) will be approached by the person seeking their consent in a sensitive and timely manner. This healthcare professional should use their clinical judgement to decide when and whether it is appropriate to make this request.

The healthcare professional should explain why the child or foetus' (blood or tissue) samples and related health information might produce pertinent information which may benefit their family members or may benefit others in similar situations. Ongoing research use of the samples taken will be discussed, as well as whether the samples that will be used by the Library have already been taken as part of standard clinical care, or if necessary, if they will need to be taken by removing a sample from their child's or baby's body after his or her death (or by taking a sample from the foetal remains). Additional looked-for findings (about child onset conditions) will not be offered where the child in question is deceased.

If the parents make any specific request in relation to including their child's samples or information in the Library, that the person taking consent is not sure about, the child's clinical team or Genomics England should be contacted about this to resolve any outstanding issues, before the consent is taken.

### **Approaching relatives of the (living) proband to seek consent to participate in the Library**

Up-to-date information about the diagnosis and clinical status of other family members can be vital in (for example) determining risks for a proband, particularly when, for example, there appears to be an inherited predisposition to cancer. Accessing this is discussed in 'Consent and confidentiality in genetic practice' (2006, Joint Committee on Genomics in Medicine), the principles of which we have referred to in creating the relevant processes.<sup>1</sup> We note that access to information about other family members is governed by the Data Protection Act 1998 and the Access to Health Records Act 1990.

If the (relative) potential participant is an adult, the clinical care team may give the recruited adult participant (who may be the proband or may not be e.g. a proband child's parent) a recruitment pack to pass on to their relative, if they are willing to take on this role of passing on information. The recruitment pack will contain a letter of introduction, Patient Information Sheet, Consent Form and a postage-prepaid envelope.

The recruited participant may also be provided with an electronic copy of the Patient Information Sheet and Consent Form for forwarding by email to their relative. To indicate their interest in participating in the Library, the relative can return the slip to say:

- They do NOT wish to be part of the Library.
- They agree to be contacted by a team member to find out more about the Library.
- They would like to be part of the Library and give written consent to participate.

The recruitment pack will contain contact details for the proband's clinical care team and the Library (via Genomics England). Relatives considering joining the Library will have the option to telephone or email for further information. Relatives can then make appointments with the proband's clinical care team to give blood or saliva samples.

If the potential participant is aged under 16 and lacks capacity to consent (dependent on their circumstances) their parents, guardian or at least one person with parental responsibility for the minor will be provided with the relevant Patient Information Sheet and be asked to consent on behalf of the minor.

### **11.2 The Process for withdrawing from the Library**

The Library will be more successful as a resource if few people withdraw once they have joined. However, participants can withdraw at any time without giving a reason. Participants will be reminded that they are welcome to discuss concerns with their clinical team at any time, or to contact Genomics England directly.

Full patient literature including how to withdraw, is available on the Genomics England website.

This lists the various options for withdrawal and provides the contact details for this process. Numbers withdrawing will be monitored by Genomics England, because withdrawal may indicate problems in the consenting or participatory experience.

Participants can withdraw at one of two levels:

‘No further contact’: this means that the Library would no longer contact the participant via their clinical team but would have their permission to retain and use information and samples provided previously, and to obtain and use further information from health records. This level of withdrawal leaves the Library dataset intact and will allow researchers to continue to study disease with the goal of improving knowledge and health.

‘No further use’: in addition to no longer contacting the participant or obtaining further information, any information and samples collected previously would no longer be available to researchers. The Library would destroy samples or return them to the diagnostic archive (although it may not be possible to trace all distributed sample remnants) and would only hold information for archival audit purposes. Such a withdrawal would prevent information about the participant from contributing to further research, but it would not be possible to remove data from research that had already taken place.

In order to allow the clinical team and the Library to action the withdrawal, the participant or those who give consent on their behalf, or another appropriate person (in the case of adults lacking capacity to consent of their own behalf), would need to evidence their wish to withdraw from participation in the Library, by requesting and completing the appropriate Withdrawal Form. This will be either the Full Withdrawal from Participation Form, or Partial Withdrawal from Participation Form.

These forms are available via the participant’s clinical team. The relevant person will be asked how they would like (or would like the participant) to withdraw. The differences between the withdrawal options available will be explained by the person actioning the withdrawal with the patient via the use of a withdrawal form.

As expressed in the withdrawal information, they can choose from either:

Option 1) ‘No further contact’: Partial Withdrawal from participation

Not to be contacted directly by Genomics England, any further.\*

However, previously collected samples from the participant can still be used and collected information will continue to be updated into the Library from the participant’s records as usual.

It is important to note that a participant’s clinician might want to discuss with them information found in their information or samples, however, which would still be possible under this option.

OR

Option 2) ‘No further use’: Full Withdrawal from participation

Not to be a participant in the Library any further.

Genomics England will:

- Not contact the participant again.
- Put beyond any further use any samples we hold from the participant.

(It may not be possible to trace all divided parts of their samples that have been distributed to approved third party research projects. Any surplus from these samples is routinely destroyed after the third party's research is completed.)

- Put beyond any further use the information we hold about the participant (aside from what is required for audit purposes - we need to retain a record that they were once part of the Library and then withdrew).
  - Not retrieve anything further from the participant's health or other records.
- \* In both options, 'any further' means from the point that Genomics England confirms to the participant's clinician that the withdrawal form has been received by us.

A copy of this form would then be given by the clinical team to:

- 1) The participant who withdraws.
- 2) The participant's clinical team – to be retained locally in the participant's medical notes.
- 3) The Library - the clinical team will make Genomics England aware of the withdrawal and send a copy of the withdrawal form electronically for storage.

### **Participant withdrawal by patient group**

Different arrangements are likely to be required by different participant groups to facilitate their withdrawal from the Library.

#### **Withdrawal: Adults who can consent on their own behalf**

This group can withdraw on their own behalf by completing the withdrawal form, which their clinical team will then process appropriately. Their withdrawal from the Library is considered to have been actioned from the point that Genomics England confirms to the participant's clinician that the withdrawal form has been received by us.

#### **Withdrawal: Under 16-year olds deemed to have capacity to consent on their own behalf**

A decision from a participant in this patient group to withdraw from participation in the Library, will be treated the same as that of an adult (aged over 16 years old), who is able to consent for themselves.

In normal circumstances, those with parental responsibility<sup>li</sup> or the appropriate legally authorised representative will be encouraged to be involved in this discussion, but the final decision about withdrawal rests with the Gillick competent young person.<sup>lii</sup>

The decision to withdraw may be felt to be against the young person's best interests, but unless the Court of Protection decides that the decision could lead to death or severe permanent injury, which is unlikely to apply here, the young person's withdrawal of consent must be respected.

### **Withdrawal: Under 16-year olds who lack capacity to consent on their own behalf**

Decisions around withdrawal from this Library on the basis of the child or young person's dissent will rely on individual circumstances to some degree. The usual practice around participation in research is not always appropriate in the context of this Library. It is the view of the Library that here, it is more appropriate to follow accepted practice in clinical care in relation to refusal of treatment, because the Library offers potential direct clinical benefits to the proband children and young people who participate. The General Medical Council and Royal College of Paediatrics and Child Health advise in relation to children's participation in medical research, that if a child or young person 'objects or appears to object in either words or actions',<sup>liii</sup> they will not be required to participate in research.<sup>liv</sup>

Decision-making around the clinical care of children or young people who are not Gillick competent should be therefore guided by an assessment of their best interests. It follows from this that it would not be legally determinative if such a child or young person withdrew their assent to participate in the Library (or declined to join the Library). Adult/s with parental responsibility retain the legal power to withdraw (or give) their consent to a child's participation in the Library.<sup>lv</sup>

Children and young people are made aware via the patient information literature for the Library that their views will be taken seriously and if they don't want to join, or don't want to continue within the Library, that this dissent will 'usually' be accepted.

In practice, this acceptance may depend on the reason that the child or young person gives for seeking to withdraw. Actioning a withdrawal solely on the basis of reluctance expressed to give a blood sample for example, would be likely to entail different considerations than if, say, the child expressed dissent for other reasons.

If a child or young person expresses dissent for reasons which are perhaps not able to be mitigated straightforwardly, and there is no overriding contradictory factor relating to their best interests - especially where direct clinical benefit seems unlikely, then they should be withdrawn from the Library in accordance with their wishes. This should happen even where consent continues from the adult or authority with parental responsibility for that child or young person.

In the case of children or young people who join the Library in order to contribute their information in relation to a proband (such as siblings), the potential for direct clinical benefit to them may be less obvious. In this case, if such a child or young person expresses their wish to withdraw from the Library, this should usually be actioned, on the assumption that benefit may be minimal and withdrawal is not likely to put the health of the sibling child or young person at risk of significant harm.

### **Withdrawal: Adults who lack capacity to consent on their own behalf**

Adults lacking capacity may be recruited to the Library, or they may have joined it under their own consent and then have lost their capacity. In such cases, if the consultee of an adult participant

advises that they would wish to be withdrawn, this withdrawal will be actioned, using the 'Withdrawal Form for a consultee to advise on an adult lacking capacity', without delay. This is in accordance of the requirements of the Mental Capacity Act (MCA) in a research setting.

If an adult participant who lacks capacity implies or expresses their dissent directly, they will be withdrawn from the Library immediately in line with the MCA and in accordance with General Medical Council guidance which provides specific considerations to be made about the care and treatment of patients who lack capacity.<sup>lvi</sup>

### **11.3 Personal insurance and the participants in the Library**

Under an open-ended agreement between the Department of Health and the Association of British Insurers the results of whole-genome sequencing carried out in the Library are not disclosable to personal life insurers. This is because they are part of an NHS transformation programme with a significant research element. However, insurers have the normal expectation that patients will disclose relevant family history, other non-genetic diagnostic test results and GPs' reports when applying for new insurance. Under the Genetics and Insurance Concordat and Moratorium, the results of other genetic tests do not need to be disclosed unless the test is for Huntington's Disease for life insurance and the insured sum is over £500,000. The Concordat and Moratorium has recently been extended from November 2017 until the end of 2019 (see attached letter) from the Department of Health and Association of British Insurers website. We will continue to adhere to the Concordat and any subsequent agreements and will include in any FAQ or other materials this information for participants.

## **12 Communications, stakeholder engagement and patient and public involvement**

## 12 Communications, stakeholder engagement and patient and public involvement

### 12.1 Communications strategy and governance arrangements

Gaining and retaining public trust and confidence is a key element of the Library, both in terms of continued recruitment and wider societal confidence in the use of genomics in medicine. When the NHS Genomic Medicine Service goes live, this will continue to be critical to confidence and take up. Particular issues are privacy and security in relation to data held by Genomics England, access to data services by commercial organisations, such as pharmaceutical companies. The public are particularly concerned about access by insurance companies. Some societal groups have concerns about access by state agencies including police and border agencies.

A communications and stakeholder engagement plan has been developed with communications colleagues in DH and with input from communications leads in NHS England, PHE and HEE. A fortnightly call is held to ensure all partners have sight of, and input into, communications about genomics. Development of a specific engagement plan for the NHS GMS and the GLHs is underway and led by NHS England.

The minister-led National Genomics Board, which consists of specialists from industry, patient support groups, and other key stakeholders will also convene an engagement subgroup to ensure that public and patient involvement remains at the forefront of the future service.

### 12.2 Communications and stakeholder engagement plan

A communications and stakeholder engagement plan has been developed. Whilst some patient and public involvement (PPI) is specific to Genomics England (for instance, in relation to policies around consent for research and the development of participant literature), there are broader societal concerns about the use of genomics, these relate in particular to commercial use of data and insurance but also how much feedback NHS patients should expect from genomic testing.

These issues require extensive dialogue with the public as genomic medicine is being introduced into routine use by the NHS and continue as a running theme throughout implementation. As part of the future NHS GMS at both a national and local level, Patient and Public involvement and engagement groups and initiatives will play an important and key role in shaping the future NHS service. Genomics England will play a leading role in a coalition of key organisations, to develop, initiate and deliver, this essential public dialogue – acting on the Chief Medical Officer for England's recommendations in the 2017 'Generation Genome' report.

#### 12.2.1 The Genomics Conversation

The Genomics Conversation was a programme of activities, beginning in 2016 and continued in 2017, led by Genomics England. The aim was to engage the general public and relevant stakeholders in key topics relating to genomic medicine.

The purpose was to begin a dialogue with the wider public and other relevant stakeholders to raise awareness of genomics and to better understand public attitudes to it. It was also important for Genomics England to learn more about potential barriers to embedding genomics into mainstream healthcare. The Genomics Conversation rolled out a broad range of activities, including debates, discussions, presentations, and outreach through social and traditional media.

### **12.2.2 The Public Dialogue**

Following on from the Genomics Conversation programme, it was clear that more work was needed to understand the public's view of mainstreaming genomic medicine, and in particular to address the Generation Genome report's recommendation of a re-thinking of the wider 'social contract'. In 2018 Genomics England, with support from Sciencewise, commissioned Ipsos MORI to undertake a series of public dialogue workshops to explore the public's views around how genomic medicine might affect the social contract in healthcare. This included the principles and red lines they felt should be in place for any social contract between the public and providers of genomic medicine. The findings from this work would then be used to inform how genomics should be implemented in the NHS in England, as well as to inform policy in healthcare in Scotland and the future of genomics generally.

A rapid literature review of what is already known to the public was undertaken in August 2018, the findings of which were shared with an Oversight Group (providing guidance) and a Stakeholder Group (providing greater diversity of views in genomics). Workshops were held across October and November 2018 in Coventry, Edinburgh, Leeds and London, which were attended by 97 members of the public and 30 experts, culminating in a final summit event to present the findings in London in April 2019.

### **12.3 Genomics England research projects – seeking the views of patients and the public**

In the early stages of the Library, Genomics England undertook a range of work to ensure that potential participant's views were included in the formulation of the ethical policies submitted for research ethics approval and in the development of patient information. The views of different groups of potential participants (those affected by cancer, rare disease, and those from BAME communities) in relation to ethical issues raised by the 100,000 Genomes Project were sought and findings were published on the Genomics England website (See all reports under 'patient and public involvement - <https://www.genomicsengland.co.uk/library-and-resources/> and the Genomics England Engagement Strategy). Genomics England will continue to engage with these stakeholders. Further to this, each of the 13 currently recruiting NHS Genomic Medicine Centres had dedicated Patient and Public Involvement leads (PPI) who are responsible for engaging with and involving local potential participant groups from diverse backgrounds. It is expected that the future NHS GMS will continue these local PPI activities to shape and inform the service.

#### **Patient and public involvement (PPI) and public engagement**

Patient and public involvement is an integral and vital part of the Library. Potential participant views were sought by Genomics England, on key ethical policies relating to consent for rare disease and cancer using a range of experienced, independent specialist market research companies (e.g. Solutions for BAME work, GfK for cancer work, Genetic Alliance UK for rare disease). This feedback was then used in a separate project to inform the design and content of literature. This literature was then extensively tested with potential participants before submission to REC. Literature supporting consent was also tested with professionals involved in the consent process, in order to ensure that there was no divergence in understanding between professionals and patients. Genomics England continues to evaluate and revise its literature, and patient and public materials, for instance including a wide variety of animations and infographics about results.

Bearing in mind the need for equity of access, and mindful particularly of those communities disproportionately affected by rare disease, a programme of work with BAME groups was initiated to understand how best to involve these groups in the project. The rare disease work is complete but work in relation to cancer, remains in progress.

Having established the trust with patients and the public, Genomics England has also a vehicle to support affiliated and partner organisations, such as Health Education England and Genetic Alliance UK, in their PPI work.

#### **12.4 The National Participant Panel**

A Participant Panel has been established. Now in their stride, this 30-strong group has provided invaluable advice on a range of topics for instance in shaping how analysis is monitored, how results are returned and how advice and support should be framed. Participant Panel members have either donated samples to the Library themselves or are carers of participants. They take part in a wide variety of consultative groups, such as the Genomics England Ethics Advisory Committee but most importantly are guardians of the dataset, with representatives on the Access Review Committee. Participants play an important part in every decision made about access to data. This will continue.

## 13 Appendices

## 13 Appendices

### Appendix 1 | List of Abbreviations

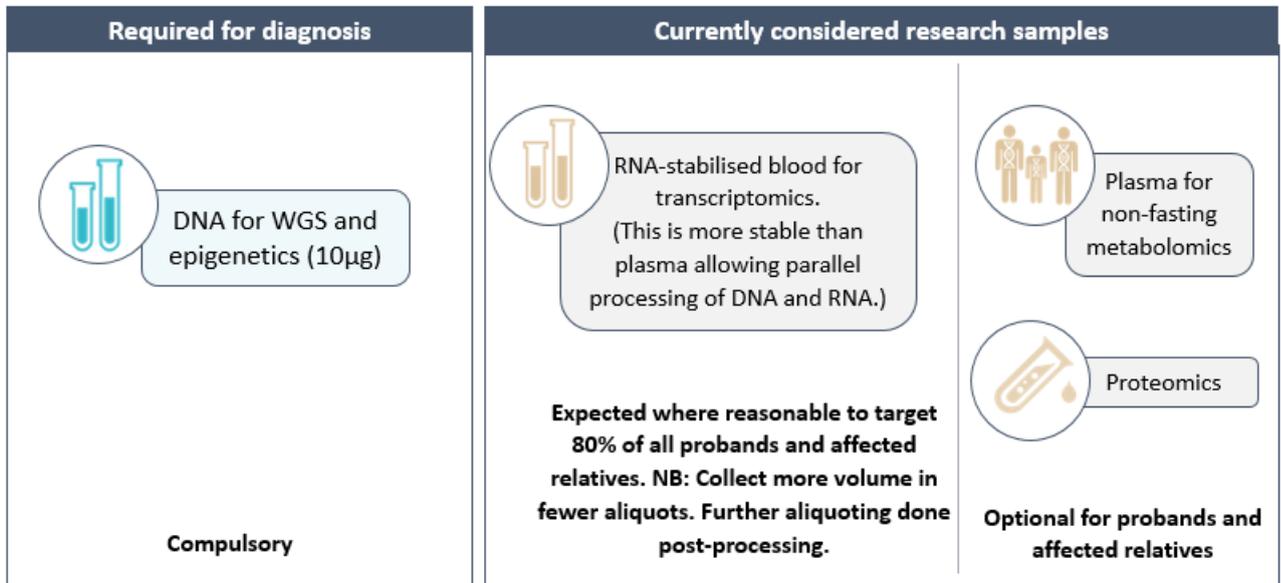
Abbreviation	Expansion
CPRD	Clinical Practice Research Datalink
EAC	(Genomics England) Ethics Advisory Committee
HEE	Health Education England
HRA	Health Research Authority
HSCIC	Health and Social Care Information Centre
NDPH	Nuffield Department of Population Health
NGS	Next Generation Sequencing
NHSE	NHS England
PHE	Public Health England
QA	Quality Assurance
QMUL	Queen Mary University, London
REC	Research Ethics Committee
WGS	whole genome sequencing
CLL	chronic lymphocytic leukaemia
DPA	Data Protection Act
FOI	Freedom of Information Act
ICO	Information Commissioners Office
FF	Fresh Frozen
FFPE	Formalin-fixed, paraffin-embedded
SOP	Standard Operating Procedures
SMS	Sample Messaging System
NHS GMS	NHS Genomic Medicine Service

## Appendix 2 | WGS Sample Requirements

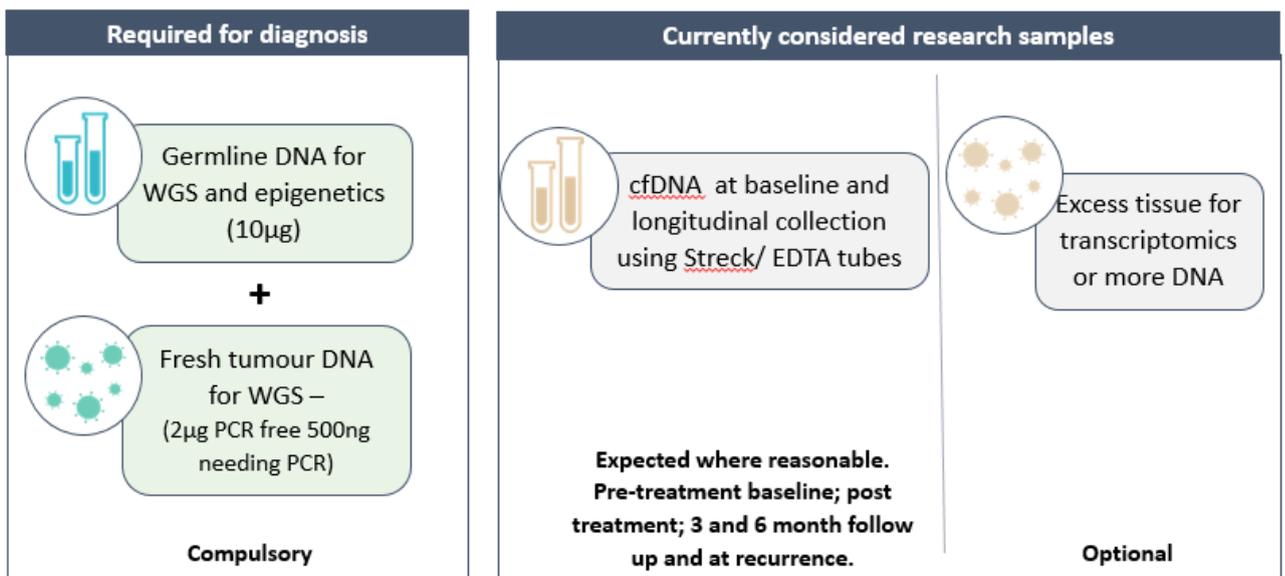
The sample requirements for WGS are detailed in the Sample Handling Guidance and the sample requirements for the NHS GMS will be an integral part of the test directory.

The research sample requirements are summarised below.

Samples required for participation in Library as a rare disease patient and additional optional research samples.



Samples required for participation in Library as a cancer patient and additional optional research samples.



N.B. 5mm x 5mm x 2mm of tumour tissue or a 15mm x 2mm core biopsy needle for tumour DNA extraction is advised to achieve a 2µg sample of DNA but this is dependent on the nature of the tumour sample.

**Blood sample volumes – Mandatory for participation**

Sample Type	Purpose	Collection Strategy	Required tubes	Blood volumes
<b>Blood for DNA extraction*</b>	WGS & excess for storage	Required for all participants	EDTA	2 x 3-5ml*

\* For patients who have had a bone marrow transplant, pre-bone marrow transplant stored DNA extracted from blood or DNA extracted from cultured fibroblasts may be used in place of the Blood EDTA DNA sample. In exceptional circumstances, where considered clinically appropriate and alternative is available DNA extracted from saliva samples may be used.

**Blood sample volumes – Indicative optional samples**

Sample Type	Purpose	Collection Strategy	Required tubes	Blood volumes
<b>RNA - stabilised blood</b>	Transcriptomics	Expected where reasonable for probands & affected relatives. Optional for all unaffected relatives.	PAXgene® blood RNA	2.5ml
<b>Plasma collected for ctDNA</b>	ctDNA	Optional for all cancer patients at diagnosis; 2-6 weeks post-surgery; at regular intervals post treatment and/ or at recurrence.	Streck® or alternative	1-12 x 5mls
<b>Blood for Plasma</b>	Metabolomics	Optional for all probands & affected relatives. Not required for unaffected relatives.	PST	8ml
<b>Blood for Serum</b>	Proteomics	Optional for all probands & affected relatives. Not required for unaffected relatives.	SST	8.5ml

	<b>Mandatory for rare disease and cancer</b>	<b>Optional***</b>			
	<b>EDTA</b>	<b>PAXgene® Blood RNA</b>	<b>Streck® or equivalent</b>	<b>PST</b>	<b>SST</b>
	<b>DNA</b>	<b>RNA</b>	<b>ctDNA</b>	<b>Plasma</b>	<b>Serum</b>
<b>Adult (14yrs+)</b>	3-5ml x 2*	>2.5ml	>10mls	>8ml	>8.5ml
<b>3-14 years**</b>	>3ml x 2*	>2.5ml	>3ml	>3ml	>2.5ml
<b>0-3 years**</b>	1-3ml	>2.5ml	>2ml	>1ml	>1ml

\*Sufficient blood needs to be taken to ensure the required quantities of DNA can be extracted see quantities above 2.0). Quantities can be modified based on local evidence.

\*\*Volumes given for children and adolescents are minimum volumes, wherever possible full collection tubes (vacutainers or paediatric collection tubes) should be obtained of an appropriate size, rather than a partially filled larger tube.

\*\*\*For optional samples, larger volumes are encouraged. This is because there is a potential diagnostic impact of this testing and transcriptomic tests require repeat runs for technical and sampling validation.

## Genomics England Governance and Committees



## Appendix 4 | Genomics England Disease List Nomination Process

Please refer to [www.genomicsengland.co.uk](http://www.genomicsengland.co.uk) for further details.



15th October 2014

Professor Mark Caulfield, Chief  
Scientist, Genomics England Ltd  
Charterhouse Square London

Dear Professor Caulfield

**1. 100,000 Genomes Project – disclosure of genome results for insurance purposes**

I am writing in my capacity as the genomics policy lead in the Department of Health to clarify the current policy regarding the use of genetic and genomic information by insurance companies. I am aware that this is potentially a concern for participants in the 100,000 Genomes Project, especially given the novel nature of the project and the long-term storage, analysis and feedback that is envisioned for the project.

The use of genetic test information by insurers is the subject of an agreement between the Government and the Association of British Insurers (ABI). It is set out in the Concordat and Moratorium on Genetics and Insurance last updated and published in 2011. For the purposes of this project the key feature is that the insurers will not require the disclosure of:

“a predictive or diagnostic test result acquired as part of clinical research. To avoid doubt, customers may be asked to disclose details of any symptoms, diagnosis or treatment received outside of the clinical research programme, even if those relate to a condition they found out about through the research programme.”

We have agreement from the ABI that the wording used in the Patient Information Sheet is in line with their interpretation of the current policy. The ABI are supportive of the aims of clinical research generally and of the 100,000 Genomes Project particularly. I am sure that they are willing to work with the Department and Genomics England to give specific guidance given the high-profile nature of this project. The initial patient engagement suggests that further detailed Q&A for patients, clinicians and insurance companies may be necessary around the interpretation of the Concordat and Moratorium in relation to the feedback of secondary findings.

The key issues that the agreement covers are the uses of results from clinical genetic tests for highly penetrant late onset disorders. At present, the results of such tests do not need to be disclosed unless they are approved by an independent panel and are for policies over the financial threshold (£500,000 for life insurance). The only approved test is for Huntington's disease for Life Insurance. Even if the insurance companies were minded to make applications for the other conditions that may be returned as pertinent or incidental findings the timescale is likely to be many months which would give ample notice of a possible need for an amendment of the Patient Information Sheet.

Insurers do expect to be told about symptom, diagnoses or treatment if they request information at the application stage or via a medical report. They may ask for information about family history but many will be willing to consider favourable genetic test results even if they are not approved by the independent panel.

Turning to the longer term, the currently published Concordat and Moratorium states that the agreement will be reviewed in 2014 and will expire in 2017 if not renewed. I am pleased to advise you that we have reached

agreement with the ABI that the Concordat and Moratorium will be extended until 2019 with a review in 2016. We are in the process of finalising the announcement and intend for this to be made public in the autumn and before recruitment begins via the NHS Genomic Medicine Centres.

Looking beyond 2019 is difficult as the 2016 review will need to be conducted in light of the priorities of the Government of the day. However, it has been a long-standing position in successive administrations that people should not be deterred from accessing healthcare or research opportunities because of concerns about insurance. In the most recent major review of discrimination legislation, in 2008 the Government decided not to extend protection against discrimination on the ground of genetic predisposition because “the existing arrangements for a voluntary moratorium on insurers’ use of predictive genetic test results ... along with continued monitoring of the use of genetic testing in the UK should provide sufficient reassurance”.

It has also been the long-standing policy to ensure that the timing of the periodic reviews of the current position always allow at least 3 years before the end of the Concordat and Moratorium in order that the Government has the opportunity to consider alternatives if either party does not wish to continue with the current voluntary agreements.

I hope that this reassures you that the Department is alert to the concerns about the long-term position on genetics and insurance and that we will work with Genomics England and the ABI to ensure that participants concerns are addressed. In doing this, I would remind you that the original policy announcement establishing the 100,000 Genomes Project stressed the importance of ensuring public trust and confidence and said: “To ensure public confidence in matters of confidentiality and access, this work will be monitored by the Chief Medical Officer for England.” As Dame Sally Davies is a Director of Genomics England we will be considering further how she should fulfil this role and how she can address any specific concerns about insurance or other uses of the data in her role advising the Government on medical and scientific matters.

I am copying this letter to Dame Sally Davies and to Felicity Harvey who is the Senior Responsible Officer for the 100,000 Genomes Project.

Please do not hesitate to contact me if I can be of further assistance.

Yours sincerely

A handwritten signature in black ink that reads "Mark Bee". The signature is written in a cursive, slightly slanted style.

Deputy Director: Genomics, Science and Emerging Technologies Health Science & Bioethics Division

## Appendix 6 | Consent Process Flows

Detailed consent process maps are included here for reference:



Consent flows.zip

---

## Appendix 7 | References

<sup>i</sup> [www.genomicsengland.co.uk](http://www.genomicsengland.co.uk)

<sup>ii</sup> Department of Health 'UK Strategy for Rare Diseases' 2013. Available at [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/260562/UK\\_Strategy\\_for\\_Rare\\_Diseases.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/260562/UK_Strategy_for_Rare_Diseases.pdf) (accessed 7 May 2018)

<sup>iii</sup> Boycott K, Vanstone M, Bulman D, MacKenzie A. Rare-disease genetics in the era of next-generation sequencing: discovery to translation. *Nat Rev Genet.* 2013 Oct;14(10):681-91.

<sup>iv</sup> Bragin E *et al.* DECIPHER: database for the interpretation of phenotype-linked plausibly pathogenic sequence and copy-number variation. *Nucleic Acids Res.* 2014 Jan 1;42(1):D993-D1000.

<sup>v</sup> <http://phenotips.org/>

<sup>vi</sup> International Rare Diseases Research Consortium 'Policies and Guidelines, Long Version.' April 2013. Available at [http://www.irdirc.org/wp-content/uploads/2017/10/IRDIRC\\_Policies\\_Longversion\\_24May2013.pdf](http://www.irdirc.org/wp-content/uploads/2017/10/IRDIRC_Policies_Longversion_24May2013.pdf) (accessed 7 May 2018)

<sup>vii</sup> <http://www.orpha.net/consor/cgi-bin/index.php>

<sup>viii</sup> Vogelstein B *et al.* Cancer genome landscapes. *Science.* 2013 Mar 29;339(6127):1546-58. doi:10.1126/science.1235122.

<sup>ix</sup> Lawrence M *et al.* Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature.* 2014 Jan 23;505(7484):495-501.

<sup>x</sup> International network of cancer genome projects. International Cancer Genome Consortium, Hudson T *et al.* *Nature.* 2010 Apr 15;464(7291):993-8.

<sup>xi</sup> Jacobsen A *et al.* Analysis of microRNA-target interactions across diverse cancer types. *Nat Struct Mol Biol.* 2013 Nov;20(11):1325-32.

<sup>xii</sup> Omberg L *et al.* Enabling transparent and collaborative computational analysis of 12 tumor types within The Cancer Genome Atlas. *Nat Genet.* 2013 Oct;45(10):1121-6.

<sup>xiii</sup> Weinstein J *et al.* Cancer Genome Atlas Research Network, The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet.* 2013 Oct;45(10):1113-20.

<sup>xiv</sup> Murtaza M *et al.* Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. *Nature.* 2013 2; 497(7447):108-12.

<sup>xv</sup> Kohane I. Using electronic health records to drive discovery in disease genomics. *Nat Rev Genet.* 2011 Jun; 12(6):417-28.

<sup>xvi</sup> MRC/ Wellcome Trust 'Framework on the feedback of health-related findings in research', March 2014. Available at <https://mrc.ukri.org/documents/pdf/mrc-wellcome-trust-framework-on-the-feedback-of-health-related-findings-in-researchpdf/> (accessed 7 May 2018)

---

<sup>xvii</sup> Criteria wording from The Public Health Genomics Foundation. 'Managing incidental and pertinent findings from WGS in the 100,000 Genomes Project. 2013' Recommendations, para 8.5, p. 18. Available at [http://www.phgfoundation.org/documents/326\\_1369298828.pdf](http://www.phgfoundation.org/documents/326_1369298828.pdf) (accessed 7 May 2018)

<sup>xviii</sup> Kalia et al Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): a policy statement of the American College of Medical Genetics and Genomics American College of Medical Genetics and Genomics Genetics in Medicine 2017; 19: 249–255

<sup>xix</sup> British Society for Genetic Medicine 'Genetic Testing of Children' 2010. Available at [http://www.bsgm.org.uk/media/678741/gtoc\\_booklet\\_final\\_new.pdf](http://www.bsgm.org.uk/media/678741/gtoc_booklet_final_new.pdf)

<sup>xx</sup> Harrison S et al. Clinical laboratories collaborate to resolve differences in variant interpretations submitted to ClinVar. *Genet Med* 2017 19(10):1096-1104

<sup>xxi</sup> Manrai *et al.* Genetic misdiagnoses and the potential for health disparities. *N Engl J Med* 2016; 375:655-65

<sup>xxii</sup> The Global Alliance for Genomics and Health, 'Constitution,' Jan 2018 (v.2.0). Available at <https://www.ga4gh.org/docs/Constitution-FINAL.pdf> (Accessed 7 May 2018)

<sup>xxiii</sup> See The Presidential Commission for the Study of Bioethical Issues, 'Privacy and Progress in Whole Genome Sequencing,' 2012. Available at: [https://bioethicsarchive.georgetown.edu/pcsbi/sites/default/files/PrivacyProgress508\\_1.pdf](https://bioethicsarchive.georgetown.edu/pcsbi/sites/default/files/PrivacyProgress508_1.pdf) (Accessed 7 May 2018); and The Presidential Commission for the Study of Bioethical Issues, 'Anticipate and Communicate: Ethical Management of Incidental and Secondary Findings in the Clinical, Research, and Direct-to-Consumer Contexts,' , 2013. Available at: [https://bioethicsarchive.georgetown.edu/pcsbi/sites/default/files/FINALAnticipateCommunicate\\_PCSBI\\_0.pdf](https://bioethicsarchive.georgetown.edu/pcsbi/sites/default/files/FINALAnticipateCommunicate_PCSBI_0.pdf) (Accessed 7 May 2018)

<sup>xxiv</sup> Steinsbekk K, Kare Myskja B, Solberg B. Broad consent versus dynamic consent in biobank research: Is passive participation an ethical problem? *Eur J Hum Genet* 2013; 21, 897–902

<sup>xxv</sup> These include the Human Rights Act (1998), Human Tissue Act (2004), Mental Capacity Act (2005), Declaration of Helsinki (2008, World Medical Association), General Medical Council 'Good Medical Practice' (2013), Joint Committee on Genomics in Medicine 'Consent and confidentiality in genetic practice', (2006, Joint Committee on Genomics in Medicine); 'Informed Consent Elements Tailored to Genomics Research', National Human Genome Research Institute, National Institutes for Health, 2012.

<sup>xxvi</sup> See, for example, the Diagnosing Developmental Disorders Project, at <https://www.ddduk.org>; and PHG Foundation, 'Realising Genomics in Clinical Practice,' 2014. Available at: [http://www.phgfoundation.org/documents/426\\_1419240712.pdf](http://www.phgfoundation.org/documents/426_1419240712.pdf) (Accessed 7 May 2018)

<sup>xxvii</sup> This includes: Health Research Authority, 'Consent and Participant Information Sheet Preparation Guidance,' 2014. Available at: <http://www.hra-decisiontools.org.uk/consent/> (Accessed 7 May 2018); and GMC, 'Good Practice in Research and Consent to Research,' 2013: *Giving information in a way that people can understand*. Available at: [https://www.gmc-uk.org/static/documents/content/Good\\_practice\\_in\\_research\\_and\\_consent\\_to\\_research.pdf](https://www.gmc-uk.org/static/documents/content/Good_practice_in_research_and_consent_to_research.pdf) (Accessed 7 May 2018)

<sup>xxviii</sup> Anonymised defined as per GMC 'Confidentiality:good practice in handling patient information,' Jan 2017. Available at: [https://www.gmc-uk.org/-/media/documents/confidentiality-good-practice-in-handling-patient-information---english-0417\\_pdf-70080105.pdf](https://www.gmc-uk.org/-/media/documents/confidentiality-good-practice-in-handling-patient-information---english-0417_pdf-70080105.pdf) (Accessed 7 May 2018)

<sup>xxix</sup> See GMC Good Medical Practice, 'Confidentiality:good practice in handling patient information,' Jan 2017; Paras 73-76: Disclosing genetic and other shared information.

---

and Paras 134-8: Disclosing information after a patient has died. Available at: [https://www.gmc-uk.org/-/media/documents/confidentiality-good-practice-in-handling-patient-information---english-0417\\_pdf-70080105.pdf](https://www.gmc-uk.org/-/media/documents/confidentiality-good-practice-in-handling-patient-information---english-0417_pdf-70080105.pdf) (Accessed 7 May 2018);

<sup>xxx</sup> We follow the recommendations of the Expert Advisory Group on Data Access (EAGDA) statement of 2013 here: <https://wellcome.ac.uk/sites/default/files/data-management-plans.pdf>

<sup>xxxi</sup> See GMC, 'Good Practice in Research and Consent to Research,' 2013. Available at: [https://www.gmc-uk.org/static/documents/content/Good\\_practice\\_in\\_research\\_and\\_consent\\_to\\_research.pdf](https://www.gmc-uk.org/static/documents/content/Good_practice_in_research_and_consent_to_research.pdf) (Accessed 7 May 2018); Health Research Authority, 'Consent and Participant Information Sheet Preparation Guidance,' 2014. Available at: <http://www.hra-decisiontools.org.uk/consent/> (Accessed 7 May 2018); and British Medical Association Consent tool kit. Available at: <https://www.bma.org.uk/advice/employment/ethics/consent/consent-tool-kit> (Accessed 7 May 2018)

<sup>xxxii</sup> As GMC 'Good Medical Practice,' 2013, also notes: See the Declaration of Helsinki and Medicines for Human Use (Clinical Trials) Regulations 2004, which requires parental consent to complement even competent under-16s' agreement to involvement in trials.

<sup>xxxiii</sup> For further information about parental responsibility, see GMC, 'Good Medical Practice' 2013, 0-18 years guidance: Appendix 2, 1.1, Parents and parental responsibility. Available at: [http://www.gmc-uk.org/guidance/ethical\\_guidance/children\\_guidance\\_appendix\\_2.asp](http://www.gmc-uk.org/guidance/ethical_guidance/children_guidance_appendix_2.asp) (Accessed 7 May 2018)

<sup>xxxiv</sup> Including: Medical Research Council, 'Medical research involving children' 2004. Available at: <https://mrc.ukri.org/documents/pdf/medical-research-involving-children/> (Accessed 7 May 2018); and Royal College of Paediatrics and Child Health: Ethics Advisory Committee, 'Guidelines for the ethical conduct of medical research involving children,' *Arch Dis Childhood* 2000;82:177-182.

<sup>xxxv</sup> Health Research Authority 'Principles of Consent, Children and Young People' (England, Wales and Northern Ireland) Childrens/Young people's wishes and assent: Available at: <http://www.hra-decisiontools.org.uk/consent/principles-children.html> (Accessed 7 May 2018)

<sup>xxxvi</sup> Case law suggests that if a young person has sufficient understanding and intelligence to understand fully what is proposed, and can use and weigh this information in reaching a decision (i.e. often called being 'Gillick- or Fraser- competent'), he or she can give consent to treatment. See *Gillick -v- West Norfolk AHA*. 3 All Er 402, at 423-4.

<sup>xxxvii</sup> Guidance on this in England and Wales available at: <https://www.hra.nhs.uk/planning-and-improving-research/policies-standards-legislation/research-emergency-settings/> (Accessed 7 May 2018)

<sup>xxxviii</sup> HRA definition of consultee: Under the Mental Capacity Act (enforced in England and Wales) before an adult who lacks capacity to give consent can be included in research, the researcher must take reasonable steps to identify someone to consult (a consultee), to determine if participation in research is appropriate. The consultee must be involved in the person's care, interested in their welfare and must be willing to help. They must not be a professional or paid care worker. They will probably be a family member, but could be another person. A person is not prevented from being a consultee if they are an attorney authorised under a registered Lasting Power of Attorney or are a deputy appointed by the Court of Protection; but that person must not be acting in a professional or paid capacity (for example, person's solicitor). Available at <https://www.hra.nhs.uk/planning-and-improving-research/policies-standards-legislation/mental-capacity-act/> (Accessed 7 May 2018)

<sup>xxxix</sup> Mental Capacity Act Code of Practice, 2013 Available at: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/497253/Mental-capacity-act-code-of-practice.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/497253/Mental-capacity-act-code-of-practice.pdf) (Accessed 7 May 2018)

---

<sup>xi</sup> Health Research Authority, Principles of consent: Adults not able to consent for themselves (England and Wales), Participants regaining capacity during your study. Available at: <http://www.hra-decisiontools.org.uk/consent/principles-ALC-EnglandandWales.html> (Accessed 7 May 2018)

<sup>xii</sup> See Medical Research Council, 'MRC Ethics Guide: Medical research involving children,' 2004

<sup>xiii</sup> Health Research Authority, Principles of consent: Adults not able to consent for themselves (England and Wales), Participants regaining capacity during your study. Available at: <http://www.hra-decisiontools.org.uk/consent/principles-ALC-EnglandandWales.html> (Accessed 7 May 2018)

<sup>xiii</sup> HRA, 'Principles of Consent: Deceased People-consent and participant information sheet preparation guidance.' Available at: <http://www.hra-decisiontools.org.uk/consent/principles-deceased.html> (Accessed 7 May 2018)

<sup>xiv</sup> See Joint Committee on Genomics in Medicine, section 4.2.6 'Who can give qualifying consent for analysis of DNA in cellular tissue?' in 'Consent and confidentiality in genetic practice', 2006. Available at: <https://www.rcpath.org/resourceLibrary/consent-and-confidentiality-in-genetic-practice.html> (Accessed 7 May 2018)

<sup>xiv</sup> See Human Tissue Authority Code of Practice (1) Consent, 'Consent requirements - Part 2: Tissue from the deceased'. Available at: [https://www.hta.gov.uk/sites/default/files/HTA%20Code%20A\\_1.pdf](https://www.hta.gov.uk/sites/default/files/HTA%20Code%20A_1.pdf) (Accessed 7 May 2018)

<sup>xvi</sup> *Ibid.* This states at para 92 that 'If the deceased person has not indicated their consent (or refusal) to post-mortem removal, storage or use of their body or tissue for scheduled purposes, or appointed a nominated representative, then the appropriate consent may be given by someone who was in a 'qualifying relationship' with the deceased person immediately before their death. Those in a qualifying relationship are found in the HT Act in the following order (highest first).

1. spouse or partner (including civil or same sex partner) The HT Act states that, for these purposes, a person is another person's partner if the two of them (whether of different sexes or the same sex) live as partners in an enduring family relationship.
2. parent or child (in this context a child may be of any age and means a biological or adopted child)
3. brother or sister
4. grandparent or grandchild
5. niece or nephew
6. stepfather or stepmother
7. half-brother or half-sister
8. friend of long standing.

It should be noted that the qualifying relatives for adults in Scotland is different and is set out in the Human Tissue (Scotland) Act.

<sup>xvii</sup> Under the Human Tissue Act 2004 in England, Wales and Northern Ireland, a 'child' is defined as being under 18 years old. Under the Human Tissue (Scotland) Act 2006, a 'child' is defined as being under 16 years old.

<sup>xviii</sup> This is because in strictly legal terms, the products of a pregnancy of less than 24 weeks gestation are considered to be the mother's tissue. (The law does not distinguish between fetal tissue and other tissue from the living; fetal tissue is regarded as the mother's tissue). For further details, please see the Human Tissue Authority Code of Practice (1) Consent, 'Consent requirements - Part 2: Tissue from the deceased'. Available at: [https://www.hta.gov.uk/sites/default/files/HTA%20Code%20A\\_1.pdf](https://www.hta.gov.uk/sites/default/files/HTA%20Code%20A_1.pdf) (Accessed 7 May 2018)

---

<sup>xlix</sup> See Human Tissue Authority Code of Practice (1) Consent, 'Consent requirements - Part 2: Tissue from the deceased', 'Qualifying relationships' Para 92, Available at: [https://www.hta.gov.uk/sites/default/files/HTA%20Code%20A\\_1.pdf](https://www.hta.gov.uk/sites/default/files/HTA%20Code%20A_1.pdf) (Accessed 7 May 2018)

<sup>l</sup> See Joint Committee on Genomics in Medicine, section 2.4 'Use of Medical records to confirm diagnosis', Consent and confidentiality in genetic practice' in 'Consent and confidentiality in genetic practice', 2006. Available at: <https://www.rcpath.org/resourceLibrary/consent-and-confidentiality-in-genetic-practice.html> (Accessed 7 May 2018)

<sup>li</sup> British Medical Association 'Parental responsibility,' 2008: Guidance from the .

<sup>lii</sup> See *Gillick -v- West Norfolk AHA*. 3 All Er 402, at 423–4.

<sup>liii</sup> General Medical Council, Good Medical Practice (2013), 0-18 years Guidance: Research: point 38

<sup>liv</sup> Modi N *et al.* Guidance on clinical research involving infants, children and young people: an update for researchers and research ethics committees, *Arch Dis Child*, 2014; 99(10):887-91

<sup>lv</sup> General Medical Council, Good Medical Practice (2013), 0-18 years Guidance: If a young person refuses treatment. Available at <https://www.gmc-uk.org/ethical-guidance/ethical-guidance-for-doctors/0-18-years/making-decisions> (Accessed 7 May 2018)

<sup>lvi</sup> General Medical Council, Good Medical Practice (2013), Consent Guidance, making decisions when a patient lacks capacity. Paras 75 and 76. Available at: <https://www.gmc-uk.org/ethical-guidance/ethical-guidance-for-doctors/consent/part-3-capacity-issues#paragraph-75> (Accessed 7 May 2018)