



The National Genomic Research Library Governance Framework

Contents

1. Introduction	4
1.1 Genomics and whole genome sequencing	5
1.2 About Genomics England	6
1.3 Healthcare	7
1.4 Research	7
1.5 Governance Committees	8
2. Genomics England and the NHS Genomic Medicine Service	11
2.1 The NHS GMS	11
2.2 Responsibilities	13
3. The National Genomic Research Library (NGRL)	14
3.1 What is the NGRL?	15
3.2 Data and sample types	16
3.3 Aims of the NGRL	18
4. Data Governance and Security: The Five Safes	21
4.1 Safe Data	22
4.2 Safe Settings	23
4.3 Safe People	23
4.4 Safe Projects	23
4.5 Safe Outputs	24
5. Access to NGRL data	26
5.1 The Access Review Committee	26
5.2 Who can apply for access	26
5.3 The application process	27
5.4 Requirements on researchers	28
6. How NGRL data can be used	31
6.1 Prohibited uses of NGRL data	31
6.2 Acceptable uses of NGRL data	32
6.3 Exporting data: The 'Airlock'	33

7. Recruiting participants	35
7.1 Routes to participate in the NGRL	35
7.2 Where recruitment procedures happen	36
7.3 Finding potential participants to invite	37
7.4 Learning and support	37
8. Participant consent	39
8.1 Ways of giving consent	39
8.2 Who can give consent?	40
8.3 Consent materials	42
9. Recontacting participants	44
9.1 Returning clinical diagnoses	45
9.2 Offering research opportunities	45
10. Withdrawing from the NGRL	47
10.1 What happens if a participant chooses Full Withdrawal	48
10.2 What happens if a participant chooses Partial Withdrawal	49
10.3 What happens when a participant dies	49

Introduction



Introduction

1.1 Genomics and whole genome sequencing

Genomics is a branch of biology focused on studying the whole of a person's genetic material, known as a genome. It involves analysing and understanding the structure, function, mapping and interactions of all the genes present in a person's DNA.

Whole Genome Sequencing (WGS) is a technique within genomics. It 'reads' DNA by finding the sequence of someone's unique 3 billion letters of DNA. This sequence forms the basis of the body's 'instruction manual', governing how each of us develop and function. Through WGS, we can learn more about how DNA can help identify the cause of genetic conditions.

For example, some rare conditions are caused by as little as a single change or 'variant' in someone's DNA. Looking at the genome of a person affected by a rare condition, compared to others who do not have the condition, can help find which variant (or variants) might be causing the condition. In cancer, tumour cells develop a different genome to healthy cells. Comparing the normal and cancer genomes may give clues about ways to treat the cancer. In some cases, knowing more about a person's genome may allow clinicians to recommend a specific treatment.

Genomics and WGS are now playing an increasing role in healthcare. This is most prominent in diagnostics, where genomic sequencing is now part of routine care in England for rare conditions and some forms of cancer through the NHS Genomic Medicine Service. The role of genomics is also growing in personalised, preventative and predictive care, for example through the potential of pharmacogenomics: a way to understand how a person's genome affects the efficacy or risk of side effects of different drugs. New sequencing technologies, and with them a new era of gene discovery and the routine use of WGS in diagnostics, are helping people with rare conditions and some forms of cancer.

However, many conditions remain undiagnosed, and the genome is phenomenally complex: it may be possible to read all 3 billion letters in a genome, but we are a long way from knowing what they all mean and how they work together or interact with environmental and lifestyle factors to cause or protect against disease.

New areas of research are deepening our understanding of these complexities and how they may contribute to health and illness, for example through transcriptomics (exploring patterns of how genes are expressed in the body), and proteomics (exploring the structure and function of proteins).

1.2 About Genomics England

[Genomics England](#) is a company owned and funded by the Department of Health and Social Care. It was set up in 2013 to deliver the UK Government's flagship 100,000 Genomes Project. This Project sequenced 100,000 whole genomes from NHS patients with rare conditions or cancer and paved the way for incorporating genomic medicine into routine care in the NHS.

Though recruitment to the 100,000 Genomes Project has ended, its impacts are still being realised, enabling new research into genetic conditions and bringing advanced diagnoses and personalised treatments to those who need them.

Today, Genomics England builds on the foundations of the 100,000 Genomes Project to develop the potential of genomic research and medicine. It operates at the intersection of healthcare and research, providing a clinical service for the NHS while also enabling linked genomic and clinical data to be used for a range of research purposes.

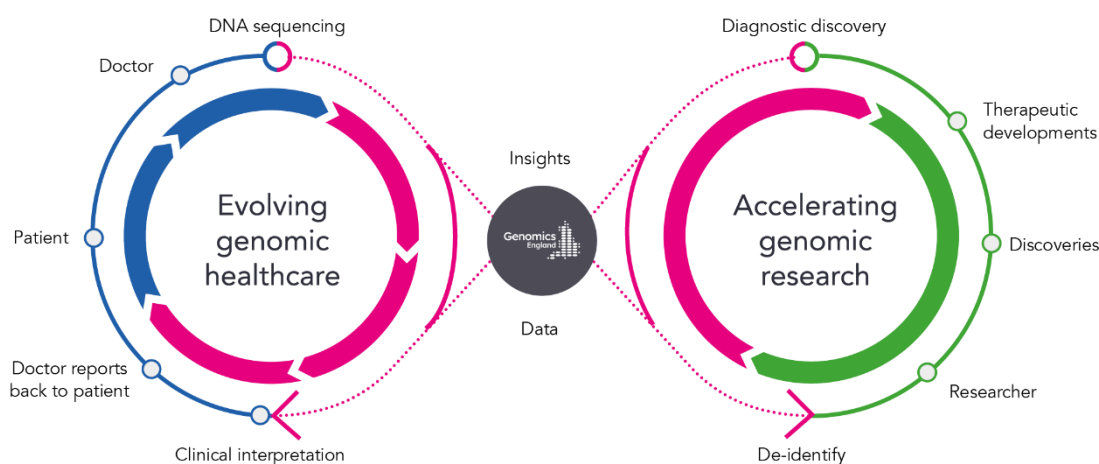


Figure 1: The Genomics England "infinity loop" showing the interaction between clinical care and research

Genomics England contracts with third party suppliers to undertake the sequencing and they provide the sequence data to us for storage and analysis. An exception to this is where we are testing new sequencing technologies such as Long Read Sequencing in-house, and comparing to established technologies for WGS. We aim to keep at the forefront of genomic research and medicine by expanding the types of genomic (for example gene panel data) and other 'omic' data we collect and analyse. Sitting across the domains of healthcare and research, Genomics England has a unique set of roles and responsibilities in both.

1.3 Healthcare

Clinical Support

We provide support for Whole Genome Sequencing (WGS) in the delivery of the NHS Genomic Medicine Service. This includes undertaking genomic sequencing and providing accredited bioinformatics pipelines for the interpretation and analysis of WGS results for NHS clinicians. Genomics England's pipelines are regulated as an In Vitro Diagnostic (IVD) device.

Clinical Research Interface

The Clinical Research Interface (CRI) team within Genomics England connect researchers with clinicians to help potential diagnoses discovered as part of research to be clinically validated, with the potential for them to be fed back to participants if the finding is sufficiently robust. Together with NHS England, the CRI team facilitate the return of clinically relevant findings to the NHS and recontact with participants for research purposes.

1.4 Research

Genomics England operates as a longitudinal bioresource, providing technical and governance infrastructure to facilitate genomic research through access to data and samples. There are two core components of the resource itself:

- The **National Genomic Research Library** (The NGRL). This is the collection of the genomic, clinical and other health data and samples from participants who have consented to participate. The NGRL is a longitudinal research resource, regulated as a Research Tissue Bank by the Health Research Authority. See [section 3](#) below for details.
- The **Research Environment**. This is the secure environment that enables approved researchers to access and analyse the data held within the NGRL, through both on premises and cloud-based infrastructure. It provides analytic software and compute to academic and life science industry researchers for genomic research, medicine and technology development.

In addition, Genomics England supports specific programmes of research, expanding and enriching the data held in the NGRL. There are two avenues for this:

- **Research Partnerships:** working with other organisations and researchers to recruit specific groups of people or 'cohorts' to the NGRL. This enables targeted recruitment to help research that focuses on specific phenotypes and populations who may stand to benefit from genomic research.
- **Research Programmes:** Genomics England runs UK Government-funded specific research programmes designed to generate a robust evidence base for new innovations

in genomic medicine. For 2021-25 these are:

- **Cancer 2.0:** Exploring long-read sequencing technology and multimodal (particularly imaging) data to support earlier, faster diagnosis of cancer.
- **The Diverse Data Initiative:** Aiming to improve the representativeness of genomic data, reduce health inequalities and improve patient outcomes in genomic medicine for minoritised communities.
- **The Newborn Genomes Programme:** Delivering the NHS-embedded [Generation Study](#), which explores the benefits, challenges, and practicalities of sequencing and analysing newborns' genomes for the screening of treatable, childhood-onset genetic conditions.¹

1.5 Governance Committees

To guide Genomics England's corporate governance and policy, there are five independent committees:

- The **Participant Panel** is a voluntary advisory group, representing the interests of the thousands of people whose data is held by Genomics England in the National Genomics Research Library (NGRL). The Panel works with Genomics England to ensure that the diverse voices of participants, patients and their families are heard and understood at all levels of the organisation. Any individual who has participated in the NGRL is eligible to apply to join the Panel.
- The **Access Review Committee (ARC)** provides an independent review and approval process to requests for access to data consented for research purposes, including genomic and health data, held by Genomics England. Participant members sit on this committee. The ARC is a decision-making committee and is the appeals body for sanctions if a researcher is found to be in breach of data access and use rules.
- The **Ethics Advisory Committee (EAC)** provides independent ethical advice and oversight to Genomics England's programmes, processes and partnerships. It acts as an advisory committee to the Genomics England Board on ethics issues related to the

¹ The Generation Study is governed by a separate protocol and approval, IRAS 324562.

activities of the organisation and the wider genomics ecosystem. Participant members sit on this Committee.

- The **Science Advisory Committee (SAC)** advises the Genomics England Board on scientific aspects of the organisation's activities and programmes. This includes overseeing disease inclusion criteria and participant recruitment strategies. The Committee considers the interests of patients, the public, scientists and clinicians engaged in genomic research or genomic medicine.
- The **Research Network Committee** oversees the operation of the Research Network and includes a range of representatives from across the Research Network communities. Participant members sit on this Committee.

Genomics England and the NHS Genomic Medicine Service



Genomics England and the NHS Genomic Medicine Service

The [NHS Genomic Medicine Service](#) (NHS GMS) is the arm of the NHS working to enable the NHS to harness the power of genomic technology and science to improve the health of our population.

2.1 The NHS GMS

The aims of the NHS GMS are:

- Embedding genomics across the NHS, through a world leading innovative service model from primary and community care through to specialist and tertiary care.
- Delivering equitable genomic testing for improved outcomes in cancer, rare inherited and common conditions and in enabling precision medicine and reducing adverse drug reactions.
- Enabling genomics to be at the forefront of the data and digital revolution, ensuring genomic data can be interpreted and informed by other diagnostic and clinical data.
- Evolving the service through cutting-edge science, research and innovation to ensure that patients can benefit from rapid implementation of advances.

As part of the NHS GMS, whole genome sequencing is offered as standard of care for patients with rare conditions and some cancers. For cancers, this may involve sequencing of both the germline and tumour.

Genomics England contracts with third party sequencing services, and provides analytical and IT systems to sequence whole genomes from these patients (and their family members where needed), and to support interpretation by NHS clinicians.

When patients are offered WGS as part of their care in the NHS GMS, they are also asked whether they wish to consent to having their genome and health data and/or their sample (blood/saliva/tissue, etc.) being accessible for research, via participation in the NGRL. The consent model and process is set out in [section 8](#).

Potential results for GMS participants found as part of research (for example, a new diagnosis) are able to be fed back by Genomics England to the NHS GMS laboratories, where they can be validated and acted upon to inform clinical care. Clinical care, diagnosis and treatment, are provided by the NHS.

Whole genome sequencing, the patient journey and the NHS Genomic Medicine Service

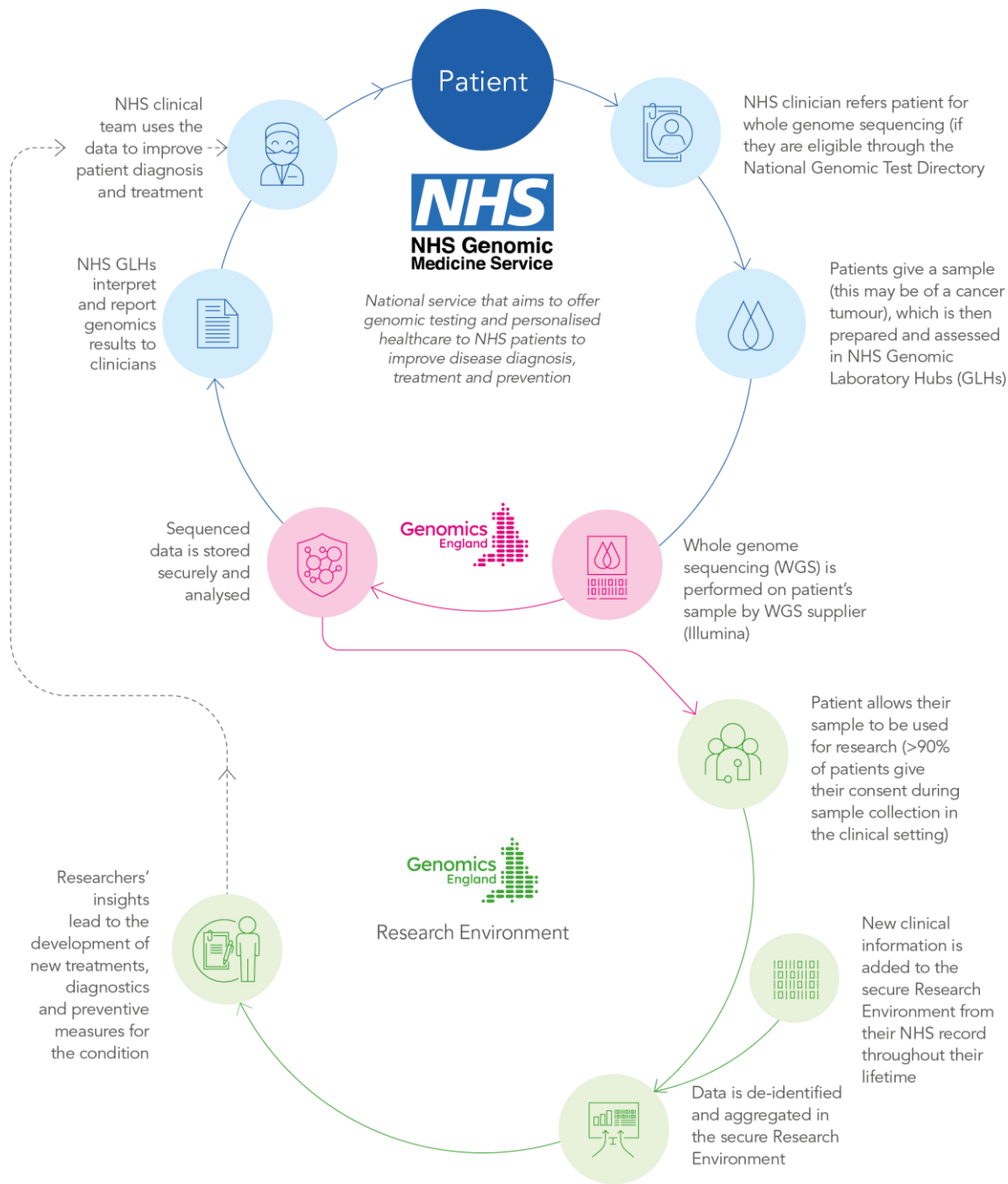


Figure 2: A diagram showing Genomics England's role in supporting the GMS

2.2 Responsibilities

Within this partnership between Genomics England and NHS GMS there are delineated responsibilities:

Table 1: The distributed responsibilities between NHS GMS and Genomics England

NHS GMS	Genomics England
Providing clinical infrastructure: Genomic Laboratory Hubs, Genomic Medicine Service Alliances	Contracting with sequencing partners to provide WGS pathways
Obtaining consent and samples from patients	Providing technical infrastructure of the Research Environment and the NGRL
Providing clinical scientists and clinical geneticists to interpret and act on genomic test results	Providing bioinformatics infrastructure to receive, process and analyse genomic data
Actioning clinically relevant findings	Returning clinically relevant findings arising from research on NGRL data to NHS clinicians
Acting as data controller for clinical care	Acting as data controller for research

The National Genomic Research Library (NGRL)



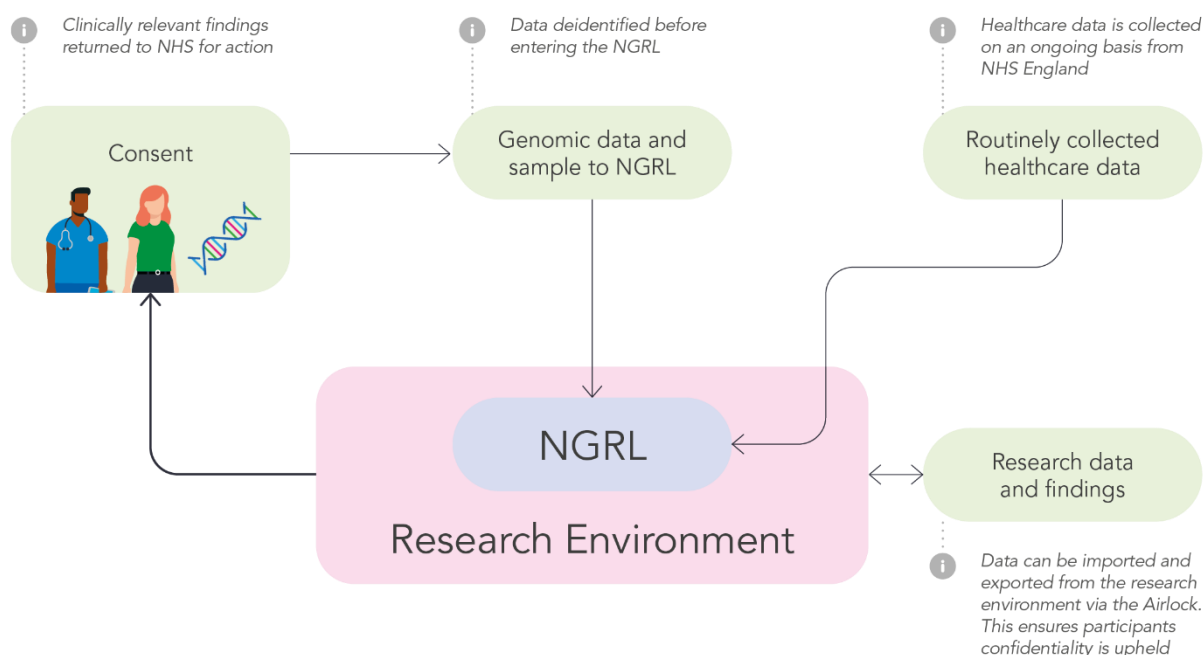
The National Genomic Research Library (NGRL)

3.1 What is the NGRL?

The NGRL is a bioresource, consisting of healthcare data and biological samples, established to facilitate future research rather than being a specific research project itself. For the purposes of Research Ethics Committee (REC) approval it is a Research Tissue Bank (RTB), which permits Genomics England to approve researchers access to the bioresource without the need for further REC approval for each individual study. Genomics England is the custodian of the data and samples in the NGRL with responsibility to act within the law, the conditions of REC approval and in line with participant expectations. It is the data controller for the use of the data in research.

The NGRL grew out of the 100,000 Genomes Project, and is an important national asset as a genomic research resource. By comparing the genome sequences and health data of multiple people in one place, researchers are able to see patterns that may not be visible with the data from only a few participants. Being able to explore all this data in one place means researchers can better understand diseases, develop new treatments and work toward new scientific discoveries.

The NGRL is held within a secure Research Environment and access to it is governed by an independent [Access Review Committee](#) (see [section 5](#)). It is expected that approved researchers wishing to access NGRL data conduct their analyses within this environment. No identifying information is permitted to leave the Research Environment. Any researcher wishing to extract data, for example, in order to publish, must apply for (see [section 4](#)).



The NGRL is made up of data from participants who have taken part in:

- the NHS Genomic Medicine Service (NHS GMS)
- the 100,000 Genomes Project
- the Generation Study (recruitment starting 2024)
- Covid-19 participants who were recruited during the pandemic as part of the GenOMICC (IRAS 269326) and REACT (IRAS 291775) studies.
- Other partnerships that have used an NGRL consent as part of their recruitment.

The NGRL stores all the data collected from these participants, including data collected under earlier versions of this governance framework.

3.2 Data and sample types

The following types of data and samples can be held in the NGRL:

- Biological samples taken from participants
- Digital genomic data generated from participants' biological samples
- Data from participants' healthcare records, from multiple sources including NHS England, received under Data Sharing Agreements
- Imaging data such as histopathology slides or scans
- Data from other sources, such as research studies or clinical trials, Patient Reported Outcome Measures (PROMS) and mortality data.

3.2.1 Biological samples

The NGRL includes a high-quality biological sample resource stored at [UK Biocentre](#), which is responsible for the ongoing processing and storage of samples. Genomics England's protocols govern the tracking, audit of and access to the samples. The sample resources support a range of potential analytic options across different 'omics' platforms, for example transcriptomics, proteomics and metabolomics.

The biological samples available in the NGRL include:

- Blood, serum and plasma
- Tumour biopsies
- Saliva
- Cord blood (from the Generation Study)

Other types of biological sample may be collected in future; these will be handled in the same way and subject to the same governance processes. As depletable resources, special consideration is given to the proposed scientific and research value of any proposed use of NGRL samples.

3.2.2 Data from participants' healthcare records

Evaluating genomic data in the context of rich and extended phenotypes from health records — such as symptoms, imaging from scans and histopathology tests, and medication histories — adds significant research value. To enable this, the NGRL combines participants' genomic data with data from their healthcare records. Genomics England has data sharing agreements in place with NHS England and other appropriate organisations that hold relevant health data about NGRL participants.

The biological and genomic data in the NGRL is linked to each participant's NHS electronic health record. This record is lifelong and regularly updated. It includes data from hospital visits, tests and treatments, and other care records. Data is sent to Genomics England from multiple sources where select individuals process it to ensure that it is de-identified before it can be accessed by researchers. A link is held by Genomics England to ensure data can be matched with the correct record, and so that relevant clinical findings can be fed back to NHS clinicians for specific patients. This process of removing identifying information from the dataset, but retaining a 'key' to be able to link to participant identities, is called pseudonymisation.

Consent to participate in the NGRL also enables Genomics England to bring in data from national and local patient-level registries for several relevant areas, including cancer and rare diseases. For instance, the National Cancer Registration Analysis Service collects data on every patient diagnosed with a cancer-registerable condition across England, unless they opt out.

The data is collected from sources covering the whole pathway from referral and screening, to palliative care and mortality.

3.2.3 Data from other sources

Genomics England may link to other data sources that enrich the NGRL dataset and aid scientific discovery, where the appropriate participant consent is in place. These sources typically include information from specific health research studies and clinical trials, and may include more detailed data relevant to a condition or treatment under investigation, such as lifestyle questionnaires, patient reported outcome measures, scans and drug responses.

This enriched data could enable deeper phenotyping of specific participants with genetic conditions.

3.3 Aims of the NGRL

Participant data contained within the NGRL may have been collected in the context of a clinical test or research study into a particular condition, but the purpose of the NGRL itself is to enable a wide range of research to be undertaken beyond the participants' conditions of interest. The aims of the NGRL fall into three broad categories:

3.3.1 Benefits to patients

It is made clear in the consent that joining the NGRL may not yield any direct benefits to participants' own care and treatment. However, being able to compare participants' data together at scale allows Genomics England and researchers to better understand conditions, which can lead to:

- Diagnosis for participants with rare conditions, where research can yield a potential new finding that is not yet in established clinical literature
- Discovery of new causes of conditions, providing evidence to offer a tailored choice of therapies
- Development of more effective treatments for patients
- Building an evidence base in support of potential commissioning decisions or service evaluation

3.3.2 Enabling research

Having data together in one secure place enriches the possibilities for discovering new insights about the connections between genomics, health and illness. We support this by:

- Centralising data from many different cohorts of people linked to their ongoing healthcare data, to enable researchers to find novel patterns between genotypes and

phenotypes, or molecular changes giving rise to illness

- Increasing the types of clinical data and sample analysis possible, for example through new -omics technologies
- Driving cutting edge tooling, for example through providing high performance compute resource and analysis software
- Fostering international collaboration with academics and industry

3.3.3 Making the UK a good place to do research.

The UK has a strong international reputation in genomic research and medicine, both for the quality of science and for navigating ethical and regulatory issues for the benefit of patients and the health system. The NGRL seeks to contribute to the development of the life sciences sector in the UK through:

- Working with the life sciences industry and academia to add to the knowledge of the genetic basis of disease, increasing opportunities for clinical trials, and building the evidence base to accelerate new technologies into healthcare.
- Stimulating and enhancing the UK life sciences industry and investment by providing appropriate industry access to the unique research dataset in the NGRL, helping to develop knowledge, methods of analysis, medicines, diagnostics and device

Data Governance and Security: The Five Safes



Data Governance and Security: The Five Safes

The Five Safes framework is a set of principles that enable data services to provide safe research access to data. Developed by the Office of National Statistics, it is a framework that is widely adopted among custodians of large sensitive datasets in the public sector. Genomics England follows this best practice model.

4.1 Safe Data

De-identification and Data Sharing Agreements (DSAs)

Keeping participants' identities confidential is a critical part of the NGRL. To do this, we de-identify participant data before researchers can access it. This involves:

- Removing anything that might identify a participant from this data, such as their full name, date of birth, NHS number, address.
- Only allowing de-identified data to be accessible in the Research Environment.

De-identification means that researchers only have access to information that does not identify participants. Genomic data is special category personal data which means that it is very sensitive and requires more protection than most types of data. However, genomic data is very granular and many of the participants in the NGRL may have unusual characteristics, such as a very rare genetic variant. This means that we cannot completely guarantee that data is anonymous and a researcher could never re-identify a participant in the dataset. For example, de-identification cannot technically prevent a researcher from using their own knowledge (for example, clinical knowledge of a patient with a rare condition) to identify an individual. This activity is prohibited by our data access rules (see [section 6](#)) and any breaches can result in sanctions by the Access Review Committee (see [section 5](#)).

An important caveat to the principle of de-identification in the NGRL is that if a researcher does discover a new research insight that could have important implications for a (unnamed to them) participant's health and care, they can flag this to the Clinical Research Interface Team within Genomics England. This internal team does have access to the 'keys' to re-identify the participant as the purpose of data use would then be to inform that participant's care. If needed they can get in touch with the participant's clinical team to take the finding forward and potentially feed it back to the participant.

Genomics England receives data from other sources, such as NHS England, under legally binding Data Sharing Agreements. These include stipulations about the security of the data, imposing conditions on use and requiring us to provide guarantees about the technical standards we adhere to. As an organisation receiving NHS data, we conform to the requirements of the Data Protection and Cyber Security Toolkit.

4.2 Safe Settings

The Research Environment

Genomics England operates a secure and controlled [Research Environment](#) (RE) providing access to NGRL data, and to tools used to analyse it. The RE provides user interface tools, shared storage drives and software through a virtual desktop. It uses a blend of on-premises and cloud-based compute hosted on UK-based AWS servers. A High-Performance Compute (HPC) cluster on AWS is also available for computationally complex tasks. Two-factor authentication is required to log in to the RE, and there is no internet access from within the RE.

4.3 Safe People

To be eligible to apply for access to the Research Environment, all researchers and their institutions must be approved.

There are different vetting and approval routes for academic and industry researchers, but they follow the same principles to ensure only competent, bona fide researchers are eligible to apply. Any researcher seeking to access data must be affiliated with an organisation that has signed an Access Agreement with Genomics England, and those institutions are responsible for ensuring their researchers comply with the terms of the agreements.

Researchers can be based in the UK or internationally. Access from some countries is prohibited; this list is updated based on UK Government advice.

4.4 Safe Projects

The independent Access Review Committee (ARC) is responsible for decision-making and oversight over applications to access NGRL data. It sets the approvals criteria for all projects based on the Acceptable Uses (see [section 6](#)), commitment to transparency and level of public or participant involvement in the research proposal. The ARC reviews commercial applications and delegates responsibility for reviewing and approving academic applications to the Genomics England Research Management team.

4.5 Safe Outputs

Data can only be extracted from the RE via what is termed the “Airlock” process (see [section 6.3](#)). Any researcher wishing to extract data, for example, the results of analyses they have undertaken in order to publish in a scientific journal, must submit a request to the Genomics England Airlock team.

Screenshotting or manual transcription of data from the virtual desktop is prohibited by the rules researchers agree to when signing the agreement with us.

Any breach of these rules can be met with sanctions imposed by the Chief Scientific Officer (CSO). Sanctions include withdrawal of access to the RE for a specified time period, requests

to journal editors to retract published papers that include prohibited data, and in the case of academic researchers, notification to their institution of the rules breach. This is also detailed in the Genomics England [Publication Policy](#) that researchers agree to abide by. The Access Review Committee is the appeals body for any researcher wishing to appeal against sanctions imposed.

Access to NGRL data



Access to NGRL data

5.1 The Access Review Committee

Genomics England is committed to maximising the use of NGRL data for research, whilst being a responsible, trustworthy guardian of that data. To do this, decisions on who can access NGRL data are guided by the independent Access Review Committee (ARC).

The ARC is the only independent decision-making group in the organisation's governance. It is made up of a mixture of people with different expertise and experience, including Participant Panel members whose data is held within the NGRL, data governance experts, geneticists and clinicians.

ARC reviews applications for access to the NGRL from life sciences industry partners, as well as informing Genomics England's data governance policies. It also monitors how data is being used within the NGRL and provides an appeals route for individuals who have received sanctions for breaching Genomics England's rules on data use.

Full details on the application process and how the ARC operates can be found on our [website](#).

5.2 Who can apply for access

Only bona fide researchers whose research is in line with Genomics England's mission can apply to access NGRL data. They can be based in the UK or internationally. These researchers come from two broad groups:

The [Research Network](#): Any non-commercial or non-profit organisation that predominantly engages in or supports academic research and activities to further public health. Examples include:

- Academic institutions (e.g. universities)
- Clinical institutions (e.g. hospitals/trusts)
- Governmental research organisations (e.g. research council institutes)
- Independent non-profit and not-for-profit research organisations (e.g. research accelerators, charities who themselves employ researchers and conduct research, etc.)

The [Discovery Forum](#): Industry researchers who work for for-profit organisations such as drug or technology companies who operate in the life science sector.

All organisations and researchers are vetted by Genomics England before being permitted to submit an application. This process will look to ensure:

- They are reputable and qualified or competent to undertake genomic research.
- They are based in countries that are not excluded from accessing the NGRL (for example, if there are Government sanctions in place).
- Their work is in line with Genomics England's mission to bring the benefits of genomic medicine to everyone.
- The appropriate and correct institutional affiliation is in place and that the individual or organisation is eligible to enter into an Agreement

Organisations and their researchers can only apply to access NGRL data after they have passed vetting.

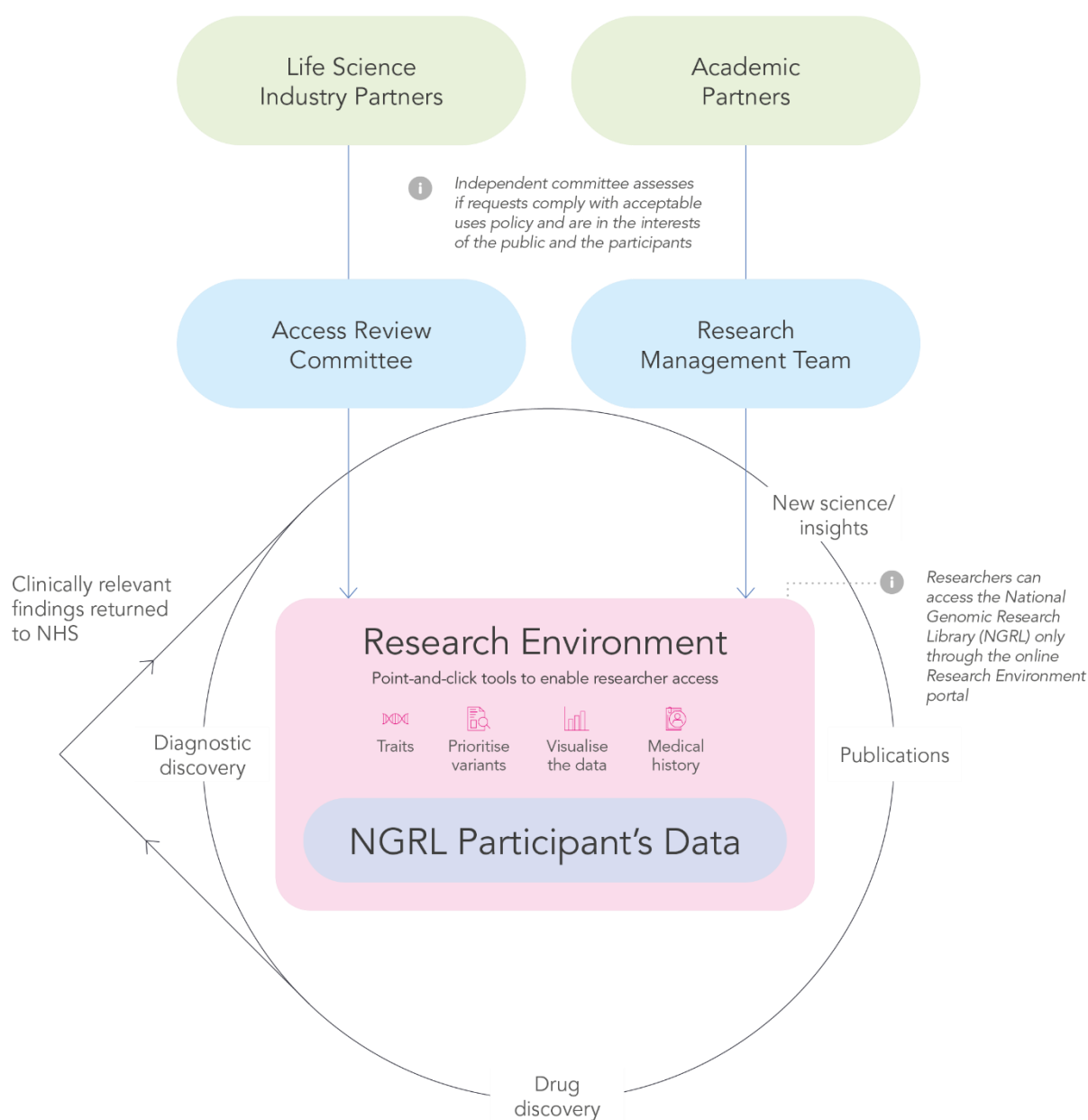
5.3 The application process

For commercial researchers

Once a commercial organisation has passed vetting, they must submit an application to the ARC. The application comprises two sections: one about the company, its research aims and objectives, and one about the specific initial research project they wish to undertake with NGRL data.

For clinical and academic researchers

The ARC has delegated the responsibility for reviewing and approving academic research applications to the Genomics England's Research Management team. Academic researchers wishing to join the Research Network must be affiliated with an organisation that has signed an Access Agreement with Genomics England. Academic researchers must then submit applications directly to this team. The Research Management team have clear routes of escalation to the ARC for any queries.



5.4 Requirements on researchers

5.4.1 Criteria for approval

All researchers must show that their projects fulfil the ARC criteria. ARC will assess whether proposals:

- Are in line with the acceptable uses for NGRL data ([section 6](#))
- Are in line with participant expectations

-
- Have involved patients, service users and the public where appropriate
 - Are sufficiently transparent about how they plan to use data

Many research projects using NGRL data will be far 'upstream' of potential benefits back to participants or the health system, for example they will be seeking to do exploratory research looking for potential molecular drug targets. Therefore, there is not a specific "public benefit" requirement for the use of NGRL data. The internal vetting of prospective research applicants ensures that only applicants whose research efforts align with Genomics England's role and mission would reach the submission stage.

5.4.2 Training and compliance

If projects are approved, all researchers involved must complete information governance and data protection training to use the Research Environment. They will need to complete this before permission is granted to access NGRL data. Researchers must also agree to the terms of use for data access. Their institutions are responsible for ensuring compliance with Genomics England's terms of use. Genomics England provides training materials, including live training sessions, so researchers can better navigate the Research Environment and get the most research value from using the NGRL data within this environment.

5.4.3 Sanctions for misuse

Researchers who are found to have broken the rules of data access will have sanctions applied. These are proportionate to the level of rule breaking. Sanctions are issued by the Chief Scientific Officer of Genomics England. If a researcher appeals their sanctions, the ARC will review them and provide a final decision.

5.4.4 Auditing and reporting requirements

Researchers must update Genomics England on the progress of their work as part of an annual Project Audit managed by the Research Management team. Failure to respond to audit requests can result in sanctions including suspension of projects and removal of access to the Research Environment.

How NGRL data can be used



How NGRL data can be used

The NGRL is a healthcare research resource. It has a broad set of acceptable uses. Equally, there are some uses that will be prohibited to ensure the purposes of use align with participant expectations in contributing to a genomic health and research resource, and do not undermine their trust.

Some requests for access to data will not be permitted. The following will be refused outright.

6.1 Prohibited uses of NGRL data

6.1.1 Requests for insurance or marketing uses

Data or samples in the NGRL will not be allowed to be used by insurance companies or marketing companies, or for insurance and marketing purposes.

6.1.2 Paternity testing or level of relatedness testing

Family relationships are logged in rare conditions where families are recruited together as a group. Where a biological relationship doesn't match the expected relationship (for example a half-sibling who was expected to be a full sibling), data from the individual with the unexpected relationship will not go into the NGRL.

If a participant were to request results of this testing, this request will be rejected. However, participants can request their own data by submitting a Subject Access Request as per their rights under data protection legislation. This means that if several, biologically related participants requested their own data from the NGRL, and then took that data to a third party to analyse, it is conceivable that data could be used in this way.

6.1.3 Parents requesting their child's data once the child reaches 16

This applies irrespective of whether the child agrees or disagrees with the request. If the young person at 16 wishes to receive a copy of their data, they are able to do so via the Subject Access Request process.

6.1.4 Requests to link data with official records

This includes employment records, tax records, benefits records and records for non-scientific or non-healthcare related purposes.

6.1.5 Requests that may lead to discrimination

That could be discrimination against a specific person, or a group of people, based on their genetic or genomic characteristics.

6.1.6 Requests for forensic or criminal justice purposes

These could be from the police, a coroner or Government department. Any such requests would be promptly referred to Genomics England's General Counsel. If a court order is presented Genomics England would be obliged to comply, but the General Counsel would make representations to the court to ensure only the minimal data required to comply with the request is provided.

6.2 Acceptable uses of NGRL data

The following list sets out the range of acceptable uses of NGRL data, consistent with the terms of participant consent and expectations and aligned with Genomics England's mission:

Supporting Healthcare

- **Clinical care:** to support diagnoses, treatment and management of participants (where clinical Genomics England staff link NGRL data back to a participant identity specifically for the purpose of informing their care – see [section 1.3](#))
- **Clinical audit:** to assess whether clinical standards are met
- **Clinical commissioning/policy:** to inform decision making by a relevant commissioning body
- **Service evaluation/development:** to validate, improve and deliver new analysis and interpretation tools and tests, including bioinformatics pipelines
- **Public health:** to gain insights at a population level of the prevalence of health issues and protections
- **Learning for health professionals:** to inform education and training about genomic research and medicine, for example on the use, analysis and interpretation of genomic data.

Supporting Health and Care Research

- **Hypothesis driven research:** to test specific research aims and questions. Any use of the data to inform interventional research would require its own approval.
- **Non-hypothesis driven research:** to explore the data, or identify novel patterns and insights by analysing data in order to generate a hypothesis. For example, looking for variants of interest that could be a target for new drug molecules.

- **Clinical trials:** to support feasibility and eligibility testing for sponsors running clinical trials, for example to identify numbers of potential participants with a specific variant.

Supporting Participants

- **Subject Access Requests:** participants can request a copy of all of the data held on them by Genomics England, by submitting a request in line with their rights under data protection legislation. This could include a copy of the genomic data held in the NGRL.

6.3 Exporting data: The 'Airlock'

Keeping participant data safe and secure is critical to the trustworthiness of the NGRL. To help ensure this, data is not 'released' that could potentially identify individuals. Instead, data analysis can only happen within a secure, monitored environment.

However, it is important that the results of analyses are permitted to be extracted from the Research Environment, so that the new insights gained through research can be shared and built on to advance further understanding. This could be via journal papers, conference presentations, or new analytic models and algorithms. The process for extracting summary and results data is called the Genomics England 'Airlock'. The concept of the Airlock is that any extraction of data happens in a staged, supervised and controlled way.

All approved projects that want to move findings out of the Research Environment must first apply for Airlock approval. Where these requests are complex they are reviewed by the Airlock Committee. The committee comprises the Caldicott Guardian of Genomics England along with clinical, bioinformatics, data security, data governance and ethics expertise. They assess whether the request aligns with an approved and registered project, and whether the data requested is minimised to a sufficient degree so that nothing that could potentially identify an individual once it leaves the secure environment can be extracted. For example granular phenotypes such as HPO terms are grouped so that they are not highly specific and potentially identifying.

There is a balance to strike here between ensuring the data is sufficiently granular to be of benefit to a publication and the wider research community, while protecting participant identities from inadvertent disclosure. This is especially difficult for research into rare conditions, where only a few individuals may exist with the condition of interest. The Airlock process considers each request on a case-by-case basis.

The only exception to this rule is if a clinical researcher has a separate consent from a participant to publish information that may lead to them being identifiable. This may be the case if, for example, they have a rare condition or phenotype, and it would not be possible to make a contribution to clinical and scientific literatures without including a level of detail that could potentially identify the participant.

Full details of the Genomics England Airlock and associated processes are available publicly in the [Airlock Policy](#).

Recruiting participants



Recruiting participants

7.1 Routes to participate in the NGRL

Genomics England works with a range of external partners to identify and recruit participants into the NGRL. The two main routes to recruit participants are via the NHS GMS and via research ecosystem partnerships.

7.1.1 NHS Genomics Medicine Service

Patients who are eligible for WGS as part of their care will be asked if they would like to donate their sample and health and care data to the NGRL. For all patients who agree, their sample will be sequenced and then analysed by Genomics England. Their genomic and health and care data will then be ingested into the NGRL and will be made accessible to researchers in the secure Research Environment.

In future, Genomics England may also work with the health systems in the devolved nations to establish pathways for recruiting participants in Wales, Scotland and Northern Ireland into the NGRL.

7.1.2 Other partners

Genomics England works with other partners to progress genomic research on specific conditions or populations that may benefit from WGS or research into genomic technologies. These research initiatives include, but are not restricted to, research studies, clinical trials and other bioresources. Examples are:

Condition- or population-specific research cohorts

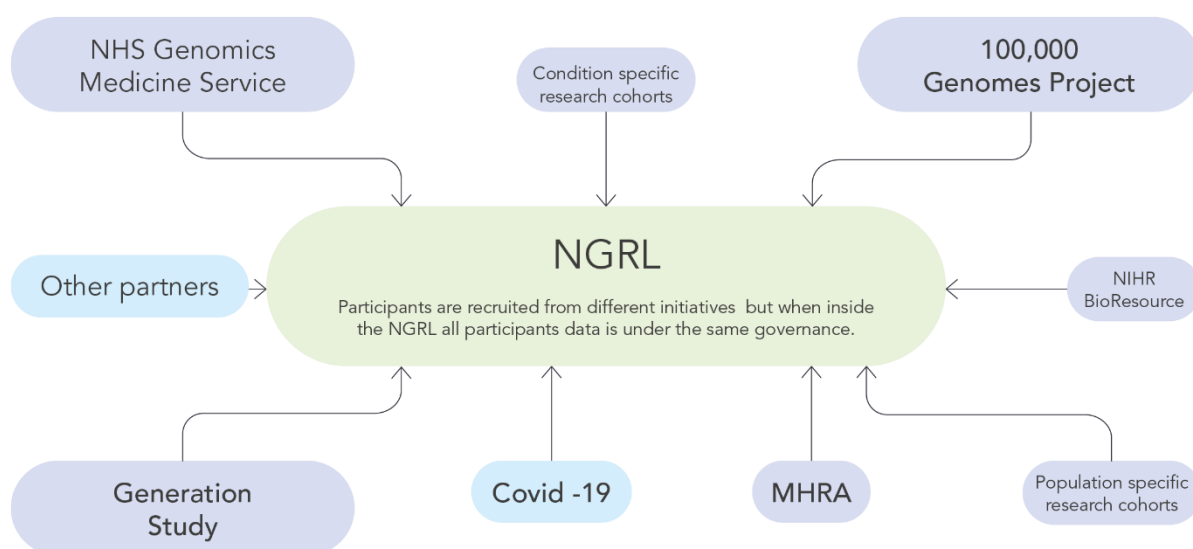
Researchers linked to universities, hospitals or charities who have approved genomic research projects may include wording about the NGRL in their consent and information materials, to give their research participants an opportunity to join the NGRL.

Research infrastructures

Collaborations with other research infrastructures such as NIHR (National Institute for Health and care Research) BioResource provide potential opportunities to enrich research across different cohorts, or de-duplicate between them where the same participant may be included in more than one resource.

Regulators

Collaborations with regulatory bodies enable us to explore new research possibilities while contributing evidence to inform regulatory standards, such as on drug safety. For example, we work with the MHRA (Medicine and Healthcare Products Regulatory Agency) on their 'Yellow Card Biobank' to help identify and approach individuals who have had an adverse reaction to a medicine, to ascertain if anything in their genomic profile could explain the adverse reaction.



7.2 Where recruitment procedures happen

To join the NGRL, participants must always give consent. In some cases, they will also give a blood, saliva, or other biological sample. In the NHS Genomics Medicine Service, routine clinical samples will be used. For other partnerships or studies this is not always possible or appropriate and new samples will be required.

7.2.1 NHS care setting

Participants may be offered the opportunity to join the NGRL at a routine clinical or specific research appointment. This could happen in either primary (GP (General Practitioner)) or secondary (hospital) care. Here, fully competent members of staff will discuss the NGRL with potential participants and share information and consent materials with them. All approved materials will be available online on the Genomics England website.

7.2.2 Outside of an NHS care setting

Consent and sample collection might be offered outside of a clinical setting, in a place that is convenient for the participant. This often happens in their home, but it is not restricted to this setting.

The organisation that has a relationship with the potential participant will make the initial contact. Potential participants must then consent to someone getting in touch to arrange an appointment at a time and place that suits them.

At this appointment, potential participants can review information and consent materials again and raise any questions before giving their consent to join the NGRL. Once they have given consent, a trained member of staff will be able to take samples.

7.3 Finding potential participants to invite

All potential participants will be identified by an organisation or individual who has an existing relationship with the person and is allowed to access the individual's identifiable data, for example their treating clinician.

Where this is face to face in the healthcare setting this will usually be undertaken by a member of the care team or research staff.

Where this is remote, initial contact will be made by an organisation with an existing relationship. This ensures there is no 'out of the blue' contact from Genomics England that might surprise or cause anxiety.

7.4 Learning and support

Learning and [materials](#) are available to support individuals who are having consent conversations with potential participants.

In partnership with NHSE and Health Education England, Genomics England have developed [specific learning](#) to assist with the recruitment process in the NHS.

Participant consent



Participant consent

There are different ways for informed consent to be obtained when a person is considering being part of the NGRL. No matter the method of communication, the following principles always apply:

- Participants have an opportunity to ask questions from a competent individual.
- Participants are given enough time and opportunity to agree or disagree.
- Information is available in different formats.
- All participant materials will be informed by principles of accessibility and user centred design.
- All core materials used for recruitment, for example information sheets and consent forms will be reviewed by Genomics England and a Research Ethics Committee.
- Where REC review is not required for supplementary materials such as infographics and easy read guides these will be reviewed and approved by Genomics England. See [section 8.3.5](#) for further details.

8.1 Ways of giving consent

8.1.1 Face-to-face

Participants will be offered the opportunity to discuss the NGRL with an appropriate individual. Information materials will be shared with potential individuals who will be given adequate time to consider and ask questions about the NGRL. If in agreement the participant will agree by recording their agreement either in writing or electronically on the relevant consent form. All individuals undertaking the consent process will always be competent to do so.

8.1.2 Telephone Consent

Where appropriate, recruitment into the NGRL may be undertaken by telephone conversation. All conversations will be scripted and staff trained and supported by Genomics England. All scripts will be developed with participant input, user tested and reviewed and agreed by Genomics England. All individuals undertaking the consent process will always be competent

to do so. Consent will be documented by the individual undertaking the telephone conversation.

Participants will be signposted to the relevant participant materials and sent copies in a suitable format.

8.1.3 Remote electronic consent

Some researchers and organisations working with Genomics England, and Genomics England, might use electronic consent platforms to allow remote consenting.

Any use of external systems used in consenting for inclusion in the NGRL will be strictly monitored by Genomics England: as a minimum checks will be undertaken to ensure the system is working correctly and that there are sufficient security measures in place to ensure the safety of that data. Genomics England will require assurance of clear identity verification of individuals to ensure accurate data and sample processing. An example of this is 2-step verification where an individual is sent a code to their email or phone which is then entered into the consent platform, also known as two-factor authentication.

A competent person will be available to answer any questions that a participant may have. Contact details will be provided as part of the information in the system.

All electronic systems and processes are regularly reviewed by Genomics England to ensure compliance with this framework.

8.1.4 Postal Consent

Materials can be posted to participants and consent forms signed and returned to Genomics England via the post. Materials will contain contact information to allow people to seek further information or speak to a competent individual about joining the NGRL.

8.2 Who can give consent?

8.2.1 Adults with capacity

Some adults will be approached because they have a condition or a particular phenotype, others will be approached as they are parents of child with a condition. In the instance of parents they may be consenting for themselves (as part of a family group) and/or on behalf of their child.

Ongoing capacity of a participant is assumed; where Genomics England is made aware of a fluctuation in or loss of capacity (as assessed by a clinician) 8.2.2 will be followed.

8.2.2 Adults who lack capacity

Some adults will lack capacity to consent to participation in the NGRL at the time of potential recruitment. Where this is the case (as assessed by a clinician), a suitable consultee will be identified and approached for their advice. Specific consultee materials will be provided and advice recorded.

Where a participant regains capacity they will be approached for their consent to join the NGRL. Where consent isn't given by the individual their data that is in the NGRL will be put beyond use, as per the withdrawal policy (see [section 10](#)).

8.2.3 Young people

16 years old and above

Those over 16 years old will be treated as adults for the purposes of joining the NGRL. If they lack capacity, the process for adults who lack capacity will be followed.

Less than 16 years old

Those under 16 who are deemed to have sufficient understanding and intelligence to understand fully what is proposed, and can use and weigh this information in reaching a decision (Gillick competent) can provide consent to participate in the NGRL.

Those under 16 who are not Gillick competent will not be able to consent for themselves. An adult with parental responsibility will need to give consent on behalf of the young person.

The consent process in this situation should ideally be a familial conversation. Children who are old enough should be engaged in the conversation by the person obtaining consent. Genomics England has developed specific [young person's information materials](#) for this purpose.

As a lifelong resource many young people will turn 16 whilst in the NGRL. Genomics England or its partners endeavour to contact young people around the time of their 16th birthday to see if they wish for their samples and data to stay in the NGRL as an adult. Specific information materials have been developed for this purpose.

Where the young person lacks capacity they will be treated as an adult who lacks capacity and a consultee will be sought for advice.

Where Genomics England does not receive confirmation of a young person's consent after their 16th birthday, their genomic data and healthcare data collected up until they are 16 will continue to be used in the NGRL but no ongoing healthcare data will be collected.

8.2.4 The deceased

In rare cases an individual may die suddenly before an opportunity to obtain consent for research has arisen. Where a sample was taken for genetic diagnosis, but the patient died before the diagnosis was received, if relatives would benefit from receiving a genetic diagnosis then Genomics England would consider inclusion of samples and data to the NGRL on the advice of their clinical team.

In order to use the samples and data in research, consent would be sought from somebody in a [‘qualifying relationship’](#) to the individual.

8.3 Consent materials

Genomics England produces consent materials in [different formats](#). These include:

- Physical printed copies
- Online digital versions of physical copies
- Videos and animations

8.3.1 For the NHS Genomic Medicine Service

Consent materials are co-produced and branded, with responsibility shared between Genomics England and NHS England Genomics Unit.

8.3.2 For the Generation Study

Recruitment and consent materials for the Generation Study have specific HRA (Health Research Authority) Approval IRAS 324562.

8.3.3 MHRA Yellow Card Biobank

Recruitment processes and participant materials are co-produced and branded by Genomics England and MHRA. These processes and materials are governed under this framework.

8.3.4 For research cohort partnerships

Genomics England reviews all consent materials. This ensures that data protection requirements and participant expectations are sufficiently addressed, while enabling information materials to be as widely accessible and understandable as possible in the interests of enabling equity of access to genomic research.

Some cohorts will be research initiatives that have their own Health Research Authority (HRA) approval. If a cohort wants to update their consent materials to enable participants to join the NGRL, Genomics England works with them to make sure that any agreed wording and amendments are completed before recruitment starts.

Other cohorts may use Genomics England's NGRL consent and information materials in addition to their own materials.

8.3.5 Developing materials

In the development of its own consent and participant-facing information materials Genomics England will gather input from:

- Users including participants, patients and the public
- Clinical Team
- Data Protection Team
- Ethics Team and, where needed, Ethics Advisory Committee
- Content Designers

All materials are reviewed and approved by Genomics England's Ethics and Data Protection Group. This group ensures that materials are in line with data protection requirements and clearly set participant expectations. The group includes Genomics England's Caldicott Guardian, Data Protection Officer and Senior Information Risk Owner as well as representatives from Ethics and Clinical teams.

Recontacting participants



Recontacting participants

The NGRL is a de-identified dataset. However, there are circumstances under which participants may be re-identified and recontact sought with them.

9.1 Returning clinical diagnoses

Participants who are part of the 100,000 Genomes Project, NHS GMS or Generation Study may have results and diagnoses returned.

In these cases, results are sent back to a clinical team in the NHS. To learn more about how clinically relevant findings are returned to clinicians and participants, visit [Genomics England's results page](#). For information specific to the Generation Study, visit the [Generation Study protocol](#).

The types of results returned for participants in the 100,000 Genomes Project or NHS GMS include:

- **Potential diagnoses:** Genomics England has established pathways to return relevant findings to the NHS for clinical action. The bar for evidence of genomic diagnosis is set intentionally high, to reduce the risk of causing harm or distress by returning an incorrect diagnosis. Both the NHS GMS and the Generation study have designated Standard Operating Procedures for the return of clinical findings.
- **Unexpected findings:** these are clinically relevant findings not related to the reason why a participant has had WGS. They are not actively sought, but may, very rarely, be discovered through the course of research. Unexpected findings may be relevant to the patient or to their family members. They can only be returned in exceptional circumstances.

To qualify as exceptional, a finding must fulfil all the following:

- have clear evidence of impact of the molecular change identified (e.g. pathogenic or likely pathogenic variant);
- constitute a high risk of serious harm such as sudden death or significant detriment to quality of life, based on current knowledge of the condition and likely penetrance;
- have clear potential for therapeutic benefit (e.g. medication, surgical option, preventative screening or action); and
- where an intervention is critical to improving health. This may include a clinical trial or therapy not yet approved for general use in the NHS.

9.2 Offering research opportunities

All participants in the NGRL agree to be recontacted in the future about relevant research opportunities. These opportunities will always be related to an individual's health, but may not

be related to the reason (condition) a participant joined the NGRL. For each recontact request

Genomics England will take into account whether a participant is likely to be aware of the reason for approach (area of study) before progressing the recontact

Participants maybe recontacted for research that is in relation to their health.

This may be because of:

- A confirmed diagnosis (confirmed by the NHS whether by WGS, or other genomic testing, or other standard diagnostic test). Genomics England undertake checks to ensure evidence of confirmed diagnosis is in place before participants are recontacted.
- The presence of a variant of interest that may show a likelihood to respond to treatments, or increased risk of side effects from certain treatments.
- The presence of a variant of interest that may make them eligible for a clinical trial.

Withdrawing from the NGRL



Withdrawing from the NGRL

Participants can withdraw from the NGRL at any time and without providing a reason, by contacting Genomics England directly via the Genomics England Service Desk. A record of withdrawal will be provided to the participant.

Withdrawal is available on two levels:

- Full ('No further use'), or
- Partial ('No further contact').

10.1 What happens if a participant chooses Full Withdrawal

10.1.1 Withdrawal prior to sequencing

Samples that are in-transit or in storage that have not yet been sequenced will be destroyed. No data will enter the NGRL.

10.1.2 Withdrawal prior to data entering the NGRL

If sequencing and analysis is complete and results have been made available, genomic data will be retained internally for the purposes of analytical validity and improvement. This data will not be included in the NGRL and will never be made available to researchers. Any remaining samples will be destroyed. All identifiable data, less that required for a record of withdrawal, will be deleted. Data that is not deleted will be ringfenced and made inaccessible.

10.1.3 Withdrawal after data enters the NGRL

If a withdrawal is received after data has been included in the NGRL, it will continue to be used in research that is already underway, as it is not possible to selectively delete it from analyses that have begun already or been completed. If a diagnosis is discovered as part of this research it would still be returned to the NHS. It will be excluded from further versions of data release into the Research Environment, so it will not be used in research in the future.

An audit record of withdrawal will be maintained, which will include some identifiable information being retained by Genomics England. This is separate from the NGRL and not available to researchers.

10.2 What happens if a participant chooses Partial Withdrawal

Samples and data will be processed, and data from clinical records will continue to be linked to the participant's genomic and demographic information.

However, Genomics England will no longer contact the participant. They will not receive exceptional findings, nor will they receive updates or invitations to participate in further research.

For patients of the NHS GMS, information that is relevant to their, or their family's, health will continue to be returned to the NHS.

In the case of participants who joined the NGRL when they were under 16 and were partially withdrawn — when they turn 16, all identifiable data, less that required for record keeping, will be put beyond use.

10.3 What happens when a participant dies

When a participant dies their data and sample will continue to be used in the NGRL. Relatives of the deceased participant may request for withdrawal of the participant by either contacting Genomics England or the participant's treating clinician. Requests for withdrawal made by relatives will be honoured.