phg foundation
making science work for health

The ethical and legal framework for a Genomics England and Sano Genetics participant engagement platform

Genomics england

Sano

Authors

Colin Mitchell, Tanya Brigden and Alison Hall

April 2021

A PHG Foundation report on the ethical and legal framework for a Genomics England and Sano Genetics participant engagement platform developed with funding from Innovate UK's Digital Health Technology Catalyst competition.

Acknowledgements

This report is funded through grant funding from Innovate UK *Industry Strategy Challenge Fund, Digital Health Technology Catalyst Round 4: Collaborative R&D,* grant number 105415.

We thank Fiona Maleady-Crowe, Shahla Salehi and Christine Patch (from Genomics England), Patrick Short (Sano Genetics) and other members of the project team for their input.

Declaration of interests

Alison Hall is a member of the Ethics Advisory Committee of Genomics England.

Disclaimer

The following report is intended to provide general information and understanding of the law. It should not be considered legal advice, nor used as a substitute for seeking qualified legal advice.

URLs in this report were correct as of March 2021

Written and produced by PHG Foundation

2 Worts Causeway, Cambridge, CB1 8RN, UK +44 (0)1223 761900

www.phgfoundation.org

Correspondence to:

intelligence@phgfoundation.org

The PHG Foundation is a health policy think-tank and linked exempt charity of the University of Cambridge. We work to achieve better health through the responsible and evidence-based application of biomedical science. We are a registered company, no. 5823194.

UNIVERSITY OF CAMBRIDGE

Executive summary

The development of digital technologies has greatly increased the potential for more active involvement by participants in large-scale, long-term health research. This includes the development of digital platforms, websites and apps which participants can join to connect with existing or new research in a dynamic way. The benefits may be considerable, both for our understanding of health and for the individuals and families who donate their time, data and samples to research.

For researchers, digital technologies could greatly enhance how they update and communicate their progress to participants, obtain the insights of those who have a lived experience of their condition, identify and invite more targeted groups to participate, and to enrich their study by collecting new and diverse data in an ongoing way. For families and individuals, digital tools could provide much more frequent and direct feedback from research, highlight new developments and discoveries, facilitate choices about new research opportunities and even in future, feedback medical or non-medical information from analysis of their data collected in research or via wearables and other devices. Digital technology provides the platform for all these and further potential functions.

With funding from Innovate UK, Genomics England, Sano Genetics and Zetta Genomics are working together to pilot a patient engagement platform that could facilitate improved communication between researchers and participants. The PHG Foundation has been commissioned to assess the legal and ethical implications of such a platform in light of Genomics England's 'internal' legal and ethical framework and within the wider ethical and legal landscape for genomic research. Because participant engagement platforms are a relatively new development and vary considerably around the world, there is not a significant literature discussing their ethical and legal implications. This is a novel review in its consideration of the ethical and legal implications of a wide range of potential features and applications of an engagement platform and in the specific consideration of Genomics England and the National Genomic Research Library (NGRL) context.

An understanding of the key ethical and legal considerations is informative for the partners in this specific project as well as those developing similar applications and contributing to research where comparable technologies are anticipated or even already implemented.

Key findings

There are some overarching ethical and legal issues that apply to the development of any digital participant engagement platform, regardless of the features or applications that they enable. Promoting transparency and managing participant expectations will be crucial in fostering participant and public trust in the existing research and in a digital platform. Transparency is required both about the development of a platform and in the surrounding processes and infrastructures that demonstrate that the endeavour is trustworthy.

Regard for participant expectations is important to ensure they are not surprised by how their data are being used and it is important to consider whether different cohorts may have different expectations of research or a platform. As with any novel technology in health research, there must also be careful consideration of the potential impacts on equity and equality. Digital platforms have the potential to remove some barriers to access to health research and to advance the understanding of disease in minority groups. However, there must be care that they do not lead to inequity for groups who are less digitally literate or less able to participate using digital means.

Consent is a core requirement for medical research. Ensuring informed consent for a digital platform will mean explaining (among other things) how data are to be used, who they are shared with and the risks and benefits involved. It will also require personalised choices about how data are used and shared. The digital setting makes determining the capacity of participants more complex, both when invited to join a platform and over time. Those lacking capacity or with fluctuating capacity may be supported to contribute by consulting relatives and carers and the development of material to facilitate their involvement as far as possible.

Genomics England has robust systems in place to safeguard privacy and protect personal data and it is important that these standards are maintained with a digital platform. In particular, this may mean that any integration of data collected from participants by the platform with data from existing datasets is only carried out by Genomics England within the secure Research Environment. Any differences in how data are used or safeguarded by the platform need to be made clear to participants and there may also need to be procedures developed to assess requests from participants or their family members to access data held about them by the platform.

As well as these overarching issues, the range of potential features and applications offered within the platform generates further specific ethical and legal considerations. In this review we consider seven potential applications:

My contributions and updates – a system to allow participants to visualise and track their contribution to research

Providing more information and updates to participants about research is potentially very beneficial. However, there are a few elements that may require some careful consideration to ensure that any risks of harm are kept to a minimum. These include taking care to ensure that participants are not surprised to learn of their risk of disease based on research that they have been included in, without being aware of their phenotypic or genotypic basis for the risk. As with several other prospective applications, this potentially engages a ‘right not to know’, and either providing a choice about such information or taking care not to disclose it are steps that could be taken to limit unwanted new information.

Participant voice — a system to allow participants to indicate what kind of research they might be interested in, contributing to research priority setting and early feedback for research proposals

This is a broad potential application that could apply both at an aggregated level, through representation of research participants in the research process as a whole, and at the level of individual research participants. Again, while there are significant potential benefits for researchers and participants, there is also the need for careful consideration of systems to avoid inadvertent disclosures of private information by participants and ongoing evaluation to ensure those participants who have been given a voice are as representative of the wider body as possible.

Research catalogue — enabling participants to browse studies and apply for those they are interested in

Some of the most promising applications for a participant platform centre on the ability to connect participants with new research studies. This could include studies that have been through the existing GEL governance and approvals processes, but it could also potentially include other research outside the GEL framework. If this is the case, it is important that any differences between external research and the GEL governance arrangements and safeguards in place for GEL research should be highlighted to participants.

Research matchmaker — a message, notification or signal to highlight research eligibility

As with 'My contributions', signalling eligibility to participants raises potential challenges of disclosing risks that they are unaware of and infringing a participant's right not to know. This is particularly the case if studies seek to 'recruit by risk' in a phenotype blind manner. Appropriate responses to this challenge could include providing a choice about matching for research topics which are not related to a 'primary' condition, or, to initially disclose only minimal information about new research opportunities if they relate to a risk of which participants may not be aware.

Data and reports — the reporting of results based on their genetic data to participants via the platform

In the future, returning extra findings to participants in the form of medical and non-medical reports may present an exciting opportunity to 'bridge the gap' between research and healthcare. It does, however, raise challenges around the variety of information being returned (e.g. results with differing levels of severity, actionability, and certainty) and the need to ensure proportionate consent and communication strategies in each instance. Delivering reports digitally raises particular challenges around communication of genetic related information - traditionally the domain of in-person genetic counselling. Ensuring that participants are sufficiently informed and supported may require different modes of communication, tailored to the individual's personal preferences and the nature/severity of the health information.

Longitudinal data collection — the ability to collect new data from individuals via the platform

A digital platform can provide a powerful new means of collecting further data from participants, either as a one-off addition or in an ongoing manner. There is a wide range of potential forms of data

collection from short and engaging online surveys, to more extensive patient reported outcome measures (PROMS) such as medical questionnaires, or in person data collection alongside further sample collection. One important consideration is whether such data collection falls within the category of 'research' and should be governed accordingly. We suggest that any intention to collect data for future research use should abide by research governance standards, including ethical approval where appropriate, consent to disclosure of confidential information and the development of policies in relation to any incidental findings that could be generated.

Wearables and symptom tracking — a specific form of longitudinal data collection using continuous monitoring technologies such as wearables devices and sensors

Collecting longitudinal data through wearable technologies, symptom trackers and other devices could help to generate real time estimates of disease risk and progression, providing a more comprehensive picture of individual health. However, the ability of these technologies to generate meaningful insights depends upon their ability to capture data reliably, and participants being willing to wear and engage with them over a sustained period of time. Barriers might include the inability to purchase and use wearables, or an aversion to 'collecting data for the sake of more data'. As such, strategies for minimising health inequalities and identifying where these technologies could generate the most value for different disease cohorts will be key.

Conclusions

Our ethical and legal analysis has led to a number of 'considerations' highlighted in the report which are applicable to the development of a participant engagement platform in general, and to the development of the more specific functions and applications that it enables. We also suggest a range of 'mitigations' that could be adopted to address the challenges we identify. These fall into three main themes. The first is the need for assessment of, and *clarity* about, the wider benefits, burdens and risks associated with the platform. The second is to ensure appropriate *personalisation* in the functionality and content of the platform according to the personal preferences of participants. The third is a commitment towards *engagement* and a meaningful and sustained consultation with participants about the development of the platform.

Together, the consideration of the key issues highlighted in this review and continuation of the thorough deliberation and consultation being carried out by the partners in this project should ensure that the promise of this novel technology is realised whilst minimising potential associated harms. If so, the development of this digital participant platform is likely to become an exemplar for practice in the rapidly evolving space connecting research, technology innovation and healthcare.

1. Introduction

This report concerns a novel patient engagement platform which is being developed by Sano Genetics (Sano) in collaboration with Genomics England (GEL) and Zetta Genomics (Zetta). The platform is the culmination of a project to develop the technology enabling a patient engagement platform for use by population-scale genomics programmes. The ultimate aim is to develop a novel system which will allow patients and their caregivers to supplement the information which is held about themselves in research databases. Funded by Innovate UK as part of its competition *Digital Health Technology Catalyst Round 4: Collaborative R&D,* the project started in October 2019 and aims to go to market in 2021.

The plans for the platform are ambitious, straddling six different applications. In combination, these applications serve two overriding objectives:

1. To embed patient experience in order to facilitate patients and research participants being genuine partners in the research process
2. To enrich the data held in the research database, through direct symptom tracking, observation or by patient-derived insights, in order to learn more about disease progression and management.

The PHG Foundation has been commissioned by GEL to provide a review of the proposed patient engagement platform and to assess its compatibility with existing legal, regulatory and ethical frameworks.

Scope and methodology

The scope of this report includes a review of the platform and its intended operation, taking account of:

- The internal governance and processes adopted by GEL
- The applicable ethical framework that Genomics England/wider UK research organisations must work within
- The applicable regulatory framework that Genomics England/wider UK research organisations must work within including relevant legislation (e.g. the Mental Capacity Act and General Data Protection Regulation, the UK GDPR and the Data Protection Act and the common law (consent and confidentiality))

The analysis has been primarily desk-based involving a detailed review of existing literature. Where relevant this has included legal and ethical sources and search engines such as PubMed to access medical and scientific material. Policy literature has been accessed through comprehensive scanning of our known contacts, networks and snowball searches of relevant literature. Search strategies have been limited to documents that are in the public domain and published in English.

In order to evaluate the potential nature, operation and processes involved in the novel platform we were provided with the following information:

- Information about the proposed participant platform including materials developed by Sano. This included the Sano Participant Portal Report, Sano-Zetta Data Flows Overview, Sano Information Sheet and Consent Form template, and Product Development Survey Plan
- Policies and documentation relating to GEL's internal governance structures and processes. This included materials in the public domain (e.g. the GEL protocol) and confidential information (policies on data handling, infrastructure and governance, airlock policies, data flows and public release); and patient facing material (patient/participant information sheets and consent forms)
- Materials developed by Sano and GEL relating to two workshops held in April 2020 with research participants including 100,000 Genome Project participants to inform the development of the platform and prioritise potential applications under consideration
- A participant survey conducted by Sano in September 2020
- A summary of the key features of the platform and operational processes likely to be taken forward by the collaborators into further testing and development (the Minimal Viable Product)

These materials were supplemented by six semi-structured interviews with representatives from GEL, Sano and other interviewees identified by the project team. The individuals were selected for interview on the basis of their knowledge and expertise of the platform, of the 100,000 Genomes Project, GEL, the National Genomic Research Library (NGRL), the additional pilots and cohorts now integrated within NGRL, or the internal or external ethical and legal/regulatory environment. Details of the interviewees and their affiliations are attached in the Appendix.

In order to ensure that our research was as informed as possible about the current plans for the platform, we had fortnightly meetings with Sano at the outset of the project, and four meetings with the project team over the project term.

For internal documentation, we adopted a system of data extraction whereby two members of the team read all documentation, and key ethical and legal issues were extracted in tabular form. These extractions were cross-checked and supplemented by the team. They were then organised into four sets of overarching issues (transparency, inequity and inequality, consent and capacity, and privacy data protection and confidentiality). The remaining issues were assigned to particular features of the patient portal in the order they first arose. In our analysis within each section, where appropriate, we have addressed external issues before addressing internal factors.

Co-development of the patient engagement platform

The workshops and survey described above refined and prioritised the potential product features in order to inform the production of a non-functioning product concept of the patient engagement platform, and ultimate production of a functioning prototype.

Through workshops and surveys, participants ordered these features into a rank-order based on the favourability to participants in the post-workshop survey. These features are: research matchmaker, participant voice, research catalogue, my contributions, wearable devices/symptom tracking, and data & reports. In our analysis, in order to minimise the duplication of issues within the report, we have re-ordered this list, so that the features are discussed iteratively. Our review makes the assumption that all features will be rolled out, although it is clear from existing empirical work that some features are more remote than others. We have also added an additional category, longitudinal data collection, which captures some missing elements. Our revised list of features on which this report is based are: my contributions and updates, participant voice, research catalogue, research matchmaker, data and reports, longitudinal data collection and wearables and symptom tracking (section 3).

This report utilises the features and definitions described in the case study report prepared by the project team.[1]

Table 1. Description of the Product Features to be included in the Patient Portal.

Product Feature	Description
My Contributions	A system to allow participants to visualise and track their contribution to research.
Participant Voice	A system to allow participants to indicate what kind of research they might be interested in, contributing to research priority setting and early feedback for research proposals.
Research Catalogue	A system that allows participants to 'browse' ongoing research projects they might join which require additional information such as a survey completion or submission of a sample to do additional testing.
Research Matchmaker	A system to allow researchers to specify criteria for joining a study (e.g. diagnosis of a condition, or a specific genetic variant) and participants

[1] Sano Genetics, Genomics England, Zetta Genomics. What research participants want: case study. 2020.

Product Feature	Description
	to specify types of research they are interested in. Participants would be securely notified of potential matches to opt-in/out to new research that is potentially relevant to them.
Data & Reports	A system to allow participants to receive genetic reports based on their whole genome sequence or other genetic data. The workshop participants considered two potential routes: (1) strictly non-medical reports (e.g. ancestry/ethnicity) (2) diagnostic/medical reports made available through an online platform and added by a medical professional, with the ability to selectively share and discuss these reports with peers / others affected by the same conditions.
Longitudinal data collection[2]	A system to allow the collection and retention of clinical and non-clinical data over time.
Wearable Devices / Symptom Tracking	A system to allow people using wearable devices (e.g. heart-rate monitor watches or other wearable sensors) to link this data with their medical / genetic data in research databases, bringing more information from the 'real world'.

This platform is still in development with a range of specific features that could be developed over time, including some more hypothetical applications that are contingent on future developments in technology and the research environments. The findings from these workshops suggested that some of these features were regarded as being more desirable and/or feasible than others. However, our report addresses all six features (plus an additional category of longitudinal data collection) although we recognise that some of these might be aspirational, and may not be reflected in the development process and functional prototype in the short term (within this funding cycle).

In this report we have reviewed these features in an iterative fashion, starting with the features that seem the least burdensome raising fewer ethical and regulatory issues, moving to those that seem more burdensome raising additional issues. In order to avoid duplication we have not replicated all the applicable ethical and regulatory issues that apply within the narrative relating to each feature.

[2] This has been added as an additional category, as it represents an intermediate step which is not reflected in the features list identified in the Sano case study.

Minimum Viable Product

Following the participant workshops that identified and prioritised desirable features for the platform and a follow-up participant survey, the project partners have developed a proposal for a Minimum Viable Product (MVP). This is a first step in testing the proposition for the participant portal, user experience hypotheses and the potential impact for Genomics England and their participants.[3] The aim is to prototype and test key features and functionality with a cohort of ~150 participants who share a condition, to gain insight into how objectives can be met, and to prioritise for future development. In this test phase, a cohort of at least 500 participants with a shared 'higher level' condition will be contacted with the goal of one hundred and fifty 100,000 Genomes Project participants registering interest in testing the minimal viable product.

Because this is a step in a dynamic process of developing and testing the portal, we have not limited our ethical and legal analysis according to the information provided in the MVP and we do not wish to make any assumptions about the goals or technical nature of the ultimate platform on the basis of this information. However, the detail in the MVP proposition has helped to inform aspects of our analysis. For example, we developed a separate heading of 'longitudinal data collection' on the basis of the MVP to analyse this function in greater depth. In this section we provide further detail on the goals and features of the MVP to inform our discussion of the potential form and functionality of the completed platform.

Goals

Three goals have been identified for the participant portal:

- Goal 1 Recontacting capability: Onboarding by invitation results in enriched participant profiles and additional recontact channels
- Goal 2 Research Matching: Relevant, actionable content and research contribution opportunities
- Goal 3 Longitudinal Data Collection: Reciprocal 'updates' foster a culture of collaboration between researchers and (anonymous) participants, resulting in higher response rates to surveys and data requests

Features

To test how these goals can be achieved, the MVP sets out some minimal features with corresponding objectives for onboarding and account creation, channels for recontact, research matching and longitudinal data collection. This is helpful to understand the potential nature of the platform and to identify emerging features.

[3] Sano participant portal progress update and MVP overview, shared with PHG Foundation in December 2020.

Onboarding

The test group users will be invited to register for the platform by Genomics England.[4] This invitation will be for participants to access a new service 'for people who want to make even more impact with 100,000 Genomes Project data, and get regular research updates'. Participants can follow a link or QR code to access a Sano and GEL co-branded landing page which provides information outlining key objectives for the platform, how their data will be used and how information has been vetted by GEL. It will also clarify how data inputted via the platform will be processed, held and linked with other data stored by GEL. Participants are invited to provide their explicit, affirmative consent for Sano to share their profile with Genomics England. Subject to clarifying their registration status, and consenting to terms of service and the privacy policy, they will then be able to join the platform.

For the purposes of this testing process, this is an adapted version of the Sano user interface and will involve the creation of a Sano account as though they are a new Sano user. The participants' information is not pre-populated from GEL profile information and participants are asked to identify whether they are registering on their, or another's behalf.

Participants are asked to input their own health histories by choosing from boxes based on ICD-10 and Orphanet disease classification standards and further detail, such as when a participant was diagnosed, or other details about symptoms may be gathered. As a custom step in onboarding Genomics England participants the platform will capture whether or not they have already received results from the 100,000 Genomes Project and, if so, if there were any findings.

An optional mobile number will also be collected if the participant chooses to provide it and they will be asked if they are interested in receiving push notifications, so they do not have to return to the app/website to check.

Research matching

Once registered, participants arrive at a customised co-branded user interface. This includes a welcome banner from Genomics England and will include a 'My studies' module highlighting research studies they already contribute to as well as an 'Updates' module which features new participation opportunities.

By navigating to a 'Research' page, the MVP states that participants will be shown a range of 'matched research'. Information is labelled with one of several tags.[5]

[4] We note that this invitation is currently for the testing and research phase of the platform development. Our analysis in sections 2.4 and 3.3 relating to the invitation to join the platform extends to the process of inviting participants to join a finalised platform, which influences aspects of the legal analysis in particular.

[5] This section quotes verbatim text from the Sano participant portal progress report and MVP overview, shared December 2020.

- New opportunities: These are studies participants may be eligible for, all approved by Genomics England. These may be online, remote or require a site visit. They are matched by condition, profile and location data. This is done in a non-scalable way for the test cohort and MVP. Each new study has a page for information and enrolment.
- Participating: Studies that the user has actively enrolled for. Each of these studies has a page with contribution updates, reports and related content.
- 'Powered by': These are studies 'powered by' the participant's data via the 100,000 Genomes Project (and studies completed online on the portal). This surfaces impact stories that Genomics England participants may not yet be aware of. Each 'powered by' study has a page with contribution updates and related content, if available.
- Research papers: Links out to a curated collection of free-to-access papers, relevant to condition.

Longitudinal data collection

The platform will also include a range of potential updates from researchers using messages, short observations, introductions to team members, relevant technology and plans, via video, podcast and webinar to test what formats and tone are likely to foster a relationship that supports longitudinal data collection. An additional element is the award of an 'impact award' or 'virtual thank you card' which participants will be prompted to share on social media.

To simulate the capture of longitudinal data, participants will receive two scheduled invitations to contribute 'meaningful health outcome' views, symptom information or participant research priorities relevant to their condition via short 'pulse surveys'. A further survey (or sequence of surveys) will simulate the capture of longitudinal data. For the MVP purposes, these 'pulse surveys' will take no more than 1 - 2 minutes to complete and there will be feedback at the end. Once an agreed volume of results per survey are available (to protect anonymity of respondents), users will be shown their results in the context of community responses.

Our approach

This report has been prepared using materials and information provided by the project team. However, we recognise that the platform is at an early stage of development and that current plans may change in future. For the purposes of this report, we have made the following assumptions:

1. The **key characteristics** of the platform:
    a. the primary aim of the platform is to facilitate research that is *'supplementary to or predicated on an existing research endeavour.'* However further use for secondary purposes (such as to support clinical care) are not directly considered in this report and will be subject to additional NHS healthcare governance

   b. the platform will be used to facilitate participant engagement and to enable additional data collection
   c. this may involve flows of data from participants to the platform, to the NGRL and to researchers, and vice-versa
2. The platform is being developed by Sano. For some activities, (such as the invitation sent to participants), documentation may be jointly badged by Sano and by GEL. Sano will be responsible for the day-to-day operation of the platform
3. The **users** of the platform will be:
   a. researchers
   b. participants/patients
   c. Sano and GEL staff or delegates
4. The **recipients** of the data collected using the platform are likely to be:
   a. researchers from both the public and private sectors who have met the governance requirements required by internal GEL practices/procedures [**researchers internal to GEL**]
   b. researchers from both the public and private sectors who are subject to additional external governance (such as approval from an additional research ethics committee) [**researchers external to GEL**]
   c. GEL staff and delegates
   d. Participants/patients
   e. Family members of participants who may be research participants in their own right
   f. Legal representatives/parents/carers as proxies for those lacking capacity
5. The early stage of development means that key aspects of the design and operation of the platform may change in future. As noted above, in this report we have not restricted our analysis to current features of the MVP, since these are more limited than is intended once the platform is fully operational
6. Our major focus is on the ethical, legal/regulatory position relating to England and Wales.

Exclusions

In view of the extensive scope of this project, the ethical and regulatory issues relating to children, such as parental responsibility, best interests, operational issues, lack of capacity and, where relevant, the challenges raised by their developing maturity, are excluded from this project. Neither does our analysis cover governance of matters explicitly excluded from the GEL protocol such as the application of relevant international standards, such as ISO 15189 regarding the future clinical bioinformatics service.

Outputs

The primary audience for this report is the project team, but in view of its general applicability, the report will be made publicly available to policy-makers, researchers, study participants and funders. In addition, an accessible summary comprising key findings from the main report will be developed, aimed at a more general stakeholder audience, including study participants, the general public, politicians and policy-makers.

Our terminology

In order to enhance the readability of this report, we have used the following conventions:

- we refer to individuals who join the platform as **participants** (although they may be participants/patients/family members of patients depending on context)
- we refer to the participant engagement platform as **the platform**
- we refer to Sano as the **host**

2. Key ethical and legal issues relating to participant engagement platforms

In this section, we discuss some overarching ethical and legal issues that are raised by the development and use of the proposed digital participant engagement platform. In the next section we develop a more detailed analysis of specific issues for the different features that a platform could incorporate.

Participant engagement platforms in general

As information and communication technology (ICT) has advanced, it has created opportunities for the improvement of multiple aspects of health care delivery and medical research. These include opportunities enabling the collection and analysis of increasing amounts of data and opportunities for improving the engagement and involvement of patients and participants in medical research.[6] As Hamakawa and colleagues have recently described, the global landscape of participant engagement and digital research platforms is diverse[7] with no clear definitions or boundaries between types of platforms. In their scoping review of digital participant-centric initiatives (PCI) which generate data for medical research, they identify a range of approaches:

- Various participants: Platforms facilitating direct communication between participants, or between participants and researchers
- Various objectives: Platforms implemented for research focusing on various diseases (e.g. PatientsLikeMe, 23andMe or Promise for Engaging Everyone Responsibly (PEER))[8] or, more commonly, research focused on specific disease areas (e.g. Rare and Undiagnosed Diseases Study (RUDY))[9]
- Various tools: Digital tools enabling a diverse range of data collection methods including uploading the results of DNA testing or electronic health records. The most common form of data gathered by these patient-centric initiatives are questionnaires.

[6] Hamakawa N, Nakano R, Kogetsu A, et al. Landscape of participant-centric initiatives for medical research in the United States, the United Kingdom, and Japan: Scoping review. *Journal of Medical Internet Research*. 2020; 22(8): 1–17.
[7] Ibid.
[8] Promise for Engaging Everyone Responsibly (PEER). Genetic Alliance, USA. URL: http://www.geneticalliance.org/programs/biotrust/peer ; cited by Hamakawa et al. (fn. 6).
[9] RudyStudy.org. National Institute for Health Research [updated 2021; cited 24 March 2021]. Available from: https://research.ndorms.ox.ac.uk/rudy/

Depending on the participants, objectives and tools involved, different consent models (broad, specific or dynamic) have been used for the further research that is enabled by these platforms.

Consideration of these platforms has tended to focus on one of a range of characteristics: their role in participant involvement or engagement, or, their facilitation of ongoing dynamic consent for example.[10] This report is novel in that it provides ethical and legal scrutiny of a wide range of potential functions of a digital research platform rather than just exploring one or two elements. We now turn to consider overarching ethical and legal issues for the digital participant platform.

2.1 Participant expectations and transparency

Promoting transparency and managing participant expectations are key issues relevant to the use of a participant platform, and also more broadly to the use of data in biomedical research and healthcare. These complementary obligations are promoted through Genomics England policies, as well as by external ethical and legal frameworks, with the aim of fostering public trust and accountability and protecting the interests of research participants. Crucially, these two imperatives are connected and overlap, as being transparent about the aims of the platform, uses of data and what can reasonably be achieved for individual participants and wider society can help facilitate realistic expectations for those stakeholders (provided that this is communicated in an accessible way).

Transparency

Transparency is a core principle underpinning the ethical governance of data, and is the first of three overarching principles that underpin the Government's Data Ethics Framework. This framework sets out the behaviours expected from those developing, deploying and using data driven technologies,[11] to ensure that they abide by the 'ethical principles for data initiatives' developed by the Nuffield Council on Bioethics: respect for persons, respect for human rights, participation and accounting for decisions.[12] Transparency is necessary to promote these principles, as individuals have a profound moral interest in controlling others' access to and the disclosure of information relating to them.

[10] Haas MA, Teare H, Prictor M, et al. 'CTRL': an online, Dynamic Consent and participant engagement platform working towards solving the complexities of consent in genomic research. *European Journal of Human Genetics*. 2021. https://doi.org/10.1038/s41431-020-00782-w

[11] Government Digital Service. *Data Ethics Framework*. June 2018 [updated September 2020]. Available from:https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/923108/Data_Ethics_Framework_2020.pdf

[12] Nuffield Council on Bioethics. *The collection, linking and use of data in biomedical research and health care: ethical issues*. 2015.

Transparency is widely recognised in UK healthcare policy and genomics initiatives. In 2016, the Chief Medical Officer's report 'Generation Genome'[13] called for a reframing of the implicit social contract for medical research and medical practice between the NHS and patients/publics. This social contract builds upon notions of solidarity and reciprocity, with publics participating and relying upon the trustworthiness of the health system, and the health system earning this trust by accepting responsibilities for improving information security and governance, including transparency and accountability. This responsibility is echoed in the 2020 government report 'Genome UK: The future of healthcare' which sets out the vision to extend the UK's leadership in genomic healthcare and research. Amongst its shared principles is the commitment that 'all genomics healthcare and research programmes will incorporate robust ethical frameworks to maintain best practice, transparency, and trust.'[14] The importance of transparency in research is the focus of a recent Health Research Authority strategy 'Make it Public' which emphasises transparency in regards to three key aspects of research: registering research studies, reporting results and feeding back to participants.[15]

Trust and trustworthiness

Transparency is often viewed as a means to securing trust. However, some commentators, most notably Onora O'Neill, caution against this, arguing that a myopic focus on transparency can actually increase deception and erode trust.[16] Transparency requires disclosure, but does not mandate that this disclosure be accessible to its audience. Although transparency can support trust, it is not sufficient. What is needed to increase trust is to ensure that demonstrable infrastructure and processes are in place (for example good governance, data stewardship and competent staff), for those 'trusting' to be able to understand, interpret and check that information. This requires that ethical values that may be important are also taken into consideration, as pointed out by Woolley and colleagues who note that 'in order to merit and garner trust, guardians of citizens' health data ought to ensure that they respect the values of the people who are expected to trust them with their data'.[17] Therefore, efforts should instead be directed towards supporting

[13] Department of Health and Social Care. Annual Report of the Chief Medical Officer 2016: Generation Genome. 2017 [cited 24 March 2021]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/631043/CMO_annual_report_generation_genome.pdf

[14] Department of Health and Social Care. Genome UK: The future of healthcare. 2020 [cited 24 March 2021]. Available from:https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/920378/Genome_UK_-_the_future_of_healthcare.pdf

[15] NHS Health Research Authority. Make it Public: transparency and openness in health and social care research. 2020 [cited 24 March 2021]. Available from: https://www.hra.nhs.uk/planning-and-improving-research/policies-standards-legislation/research-transparency/make-it-public-transparency-and-openness-health-and-social-care-research/

[16] O'Neill O. Lecture 4: Trust and Transparency [Lecture] Reith Lectures. 2002.

[17] Woolley JP, McGowan ML, Teare HJA, et al. Citizen science or scientific citizenship? Disentangling the uses of public engagement rhetoric in national research initiatives. *BMC Ethics*. 2016; 17(1).

trustworthiness. In other words, transparency should result in the disclosure of information that assists participants to intelligently place trust in the platform.

It should also be noted that transparency is relevant, not only for participants, but for the wider public too in order to show research to be a trustworthy endeavour.

The relationship between transparency and expectations

A function of transparency is that it contributes towards realistic participant expectations. Participant expectations in this report refers to the participant's beliefs about the nature and anticipated benefits and harms of using the platform. Setting out mutually acceptable expectations that participants should have, especially in relation to things that they value e.g. the use of their data, is important for the promotion of principles underpinning healthcare research such as respect for autonomy and beneficence/non-maleficence. The legal framework also seeks to protect and safeguard the interests of individuals where they have a reasonable expectation of privacy and confidentiality.[18] Within the health and social care system, this is now an explicit part of the Caldicott Principles for sharing personal confidential data. Principle 8 requires healthcare professionals and others to take steps to ensure there are no surprises regarding how the data is used or shared.[19]

Legal frameworks

Transparency is an essential component of the legal frameworks which govern the collection and use of personal data for research, most notably in regard to (a) the transparency of information that needs to be provided as part of the consent process, and (b) the transparency necessary for the processing of personal data. Consent and data protection shall be considered in more detail in following sections but are briefly outlined below to highlight their interactions with transparency.

Consent – Consent underpins the ethical and legal practice of healthcare research, and is essential for protecting individual choice and autonomy. An important facet of valid consent is that it be fully informed. This requires not only that relevant information is disclosed but also that it is communicated in a way that the patient understands. In this way, it goes one step further than transparency. The consent form and accompanying participant information also contribute to a larger body of information provided by GEL (together with their policies, frameworks and protocols) in relation to which individuals place their trust.

Data protection – Transparency is a fundamental principle for the processing of personal data. Article 5(1)(a) of the UK General Data Protection Regulation (UK GDPR) requires that 'personal data shall be

[18] *Campbell v MGN Ltd* [2004] UKHL 22, [2004] 2 AC 457, [2004] 2 All ER 995.
[19] National Data Guardian. The Caldicott Principles. 2020 [cited 24 March 2021]. Available from: https://www.gov.uk/government/publications/the-caldicott-principles

processed lawfully, fairly and in a transparent manner in relation to the data subject'. This means that data subjects should be provided with sufficient information to enable them to understand, and if appropriate challenge, how their personal data is being processed. In addition, transparency covers how data controllers communicate with data subjects in relation to their rights under the UK GDPR and how they facilitate the exercise by data subjects of such rights. The accountability principle contained in Article 5(2) requires that controllers must be able to demonstrate that personal data is processed in a transparent manner. One of the ways in which the requirement for transparency is upheld is through mandating provision of information to the data subject (see section 2.4).

Participant and public engagement

In the context of a participant engagement platform for participants involved in research conducted by GEL, transparency and participant expectations are reflected in the policies and the internal approach to legal and ethical issues set out in the NGRL Protocol.[20] This Protocol outlines the applicable principles and standards for inclusion of an individual's data and samples in the National Genomic Reference Library. As well as detailing patient and clinical benefits and scientific and transformational objectives, it also describes the ethical and governance frameworks required for research.

This public-facing document reflects GEL's commitment to provide transparency around the research activities undertaken in the NGRL. Transparency is one of GEL's four central aims and together they express a commitment to patient and public involvement (PPI) and stakeholder engagement. One of the ways in which this is reflected is through the GEL Participant Panel, made up of participants and carers of participants, who help to guide decisions such as how results are returned and how advice and support should be framed.[21] Members of this panel also serve on other consultative groups, including the Genomics England Ethics Advisory Committee and the Access Review Committee. Additionally, efforts are being made to engage all groups involved in research, including ethnic groups and minority populations and those disproportionately affected by rare disease.[22] [23] In doing so GEL are providing pathways for participant expectations to feed back into and shape the project and the wider research activities of the organisation.

[20] Genomics England. The National Genomic Research Library protocol v5.1. 2020. Available from https://www.genomicsengland.co.uk/national-genomic-research-library/

[21] Ibid, section 12.4 p.85

[22] For example, this could include people of Ashkenazi Jewish descent who may have increased risks of certain inherited diseases. See UK Rare Diseases Framework https://www.gov.uk/government/publications/uk-rare-diseases-framework for a general description of the need for representation from all populations, especially those who may go unheard.

[23] Genomics England. The National Genomic Research Library protocol v5.1. 2020. Section 12.3, p.85. Available from https://www.genomicsengland.co.uk/national-genomic-research-library/

More broadly, the Protocol emphasises the importance of transparency and wider confidence in research extending beyond research participants to the public. It states that 'gaining and retaining public trust and confidence is a key element of NGRL, both in terms of continued recruitment and wider societal confidence in the use of genomics in medicine'.[24] Issues identified that concern the public include those surrounding privacy and security in relation to data held by GEL, access to data services by commercial organisations, access by insurance companies and access by state agencies including police and border agencies.[25] GEL have led a programme of activities to facilitate a dialogue around some of these broader concerns. These include 'The Genomics Conversation', to understand public attitudes to genomic medicine, and 'The Public Dialogue'[26] a series of workshops exploring how genomic medicine might affect the social contract in healthcare (including the principles and red lines the public felt should be in place).[27]

Participant expectations

Participants using the platform may currently be recruited from three main cohorts: participants in the 100,000 Genomes Project, patients who are part of research conducted by the GenOMICC COVID-19 study and patients who have had a test through the NHS Genomic Medicine Service. However, in future, any participant, including participants from additional cohorts might be invited to join and use the platform. It is likely that each of these cohorts of participants will have different expectations around how their data will be used by GEL, who else will have access to it (researchers/ commercial access), how long for, the benefits that they anticipate and the risks that they are prepared to tolerate. These expectations around use may be bound up in their existing relationship with GEL and the aims of the research project that they originally signed up for, and may motivate their participation in the platform.

Participants may expect commercial partners to access their data for research purposes, but perhaps not that information about them will be collected by a commercial partner, and held on a separate platform operated independently from the NGRL: their expectations may be that the platform is solely a GEL research endeavour. The invitation letter inviting potential participants to join the MVP describes this as a collaborative effort between GEL and SANO as part of a grant award. It indicates that the portal will be co-badged, illustrating to those using it the involvement of both parties. However, the extent of this involvement in each instance will need to be transparent and communicated in an accessible way. One

[24] Ibid, section 12.1, p.83
[25] Ibid.
[26] Ipsos MORI, Genomics England, Sciencewise, UK Research and Innovation. A public dialogue on genomic medicine for a new social contract. https://www.ipsos.com/sites/default/files/ct/publication/documents/2019-04/public-dialogue-on-genomic-medicine-full-report.pdf
[27] Genomics England. The National Genomic Research Library protocol v5.1. 2020. Section 12.2, p.83-84. Available from https://www.genomicsengland.co.uk/national-genomic-research-library/

aspect of this will be to have clear demarcation of 'research' from other forms of activity and information gathering.

> Consideration: The presentation of the platform, branding and other associated aspects should reflect the involvement of the respective parties and their functions and responsibilities.

2.2 Inequality and inequity in health

The concepts of inequality and inequity are important in the context of health care and medical research. Understanding the determinants of health and societal responses to these differences including potential mitigations are key when considering the development and implementation of any novel technology.[28] There are multiple definitions and concepts of inequality and health inequity. For the purposes of this report, we describe health inequality as a description of differences in health status or health outcomes between individuals or groups.[29] By contrast, health inequity refers to distributions of health status or outcome that are unjust or arise due to injustice. This requires a normative judgment premised on theories of justice, together with an assessment of any background injustices that might lead to differences in health outcome or status. In the health context, commentators such as Julian Tudor Hart have noted the 'perverse relationship' in which those most in need of medical care were least likely to receive it, with those in least need using services more often and more effectively.[30]

Obligations 'to reduce inequalities' are enshrined within the Health and Social Care Act 2012,[31] and local clinical commissioning groups are obliged to reduce inequalities relating to health services access and the outcomes from health services provision. In the medical research context, the Health Research Authority is required to meet the requirements of the Equalities Act 2010 for its employees and for research ethics committee members, and to take account of protected characteristics. However, the UK policy framework

[28] Burton H, Ordish J, Cook S. My Healthy Future: Health technologies and social impacts report. PHG Foundation. 2019. Available from: https://www.phgfoundation.org/report/health-technologies-social-impacts

[29] Kawachi I, Sabramanium, SV. A glossary for health inequalities. *Journal of Epidemiology and Community Health*. 2002; 56(647).

[30] Hart JT. The inverse care law. *Lancet*. 1971; 297(7696): 405-412.

[31] Department of Health and Social Care. Reducing health inequalities – The Health and Social Care Act 2012. Fact sheet c2. 2012. Available from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/138267/C2.-Factsheet-Tackling-inequalities-in-healthcare-270412.pdf

for health and social care research does not specifically address the requirement for equality, but general obligations pursuant to the Equality Act 2010 are noted.[32]

Diffusion of new technologies

New innovations typically diffuse unequally through populations, with those of higher socioeconomic status benefiting sooner than lower socioeconomic groups.[33] Some commentators have suggested that this is because socioeconomic status embodies an array of resources, including money, knowledge, prestige, power and beneficial social connections that have a protective effect against repeated health insults. Certainly, empirical work suggests that those who are more educated, are quicker to take advantage of technological advances, including those with potential to improve health.[34] The advances conferred by education seem two-fold: those with more education tend to be more positive about innovation, but they are also more likely to have the skills, expertise and resources (such as digital expertise) needed to support the adoption and implementation of those technologies.

Potential to impact existing inequalities

Despite these generalised findings, technology can also have a transformative impact on existing social determinants of health by altering the mechanisms through which social determinants exert their influence. Depending on their purpose and application, complex technologies may increase inequalities, whilst technologies aimed at simplifying management may reduce health disparities.[35] As the efficacy of technology grows 'so too does the burden on society to provide access to technology equitably to all those in need'.[36]

These general observations about the motivation and ability to utilise novel technologies are highly relevant to the proposed platform and its development and implementation. The platform has been funded by Innovate UK on the basis that it offers unprecedented opportunities to enrich the NGRL dataset and to build participant engagement with research. By providing a secure and proportionate route for communication between participants and researchers, the platform has the potential to impact multiple points in the research pathway including enabling more targeted recruitment and the consequent

[32] NHS Health Research Authority, The UK Health Departments. UK policy framework for health and social care research v3.3. 2017. Available from: https://s3.eu-west-2.amazonaws.com/www.hra.nhs.uk/media/documents/Final_Accessibility_uk-policy-framework-health-social-care-research_.pdf

[33] Korda R, Clements M, Dixon J. Socioeconomic inequalities in the diffusion of health technology: Uptake of coronary procedures as an example. *Social Science & Medicine*. 2011; 72: 224e229.

[34] Glied S, Lleras-Muney A. Technological innovation and inequality in health. *Demography*. 2008; 45: 741-61.

[35] Weiss D, Rydland HT, Øversveen E, et al. Innovative technologies and social inequalities in health: A scoping review of the literature. *PLoS One*. 2018; 13(4):e0195447.

[36] Wise PH. Emerging technologies and their impact on disability. In: *The Future of Children.* 2012; 22(1): 169-191

development of safer more effective interventions and treatments. By empowering eligible research participants to become more activated and motivated, research participants may be more likely to adhere to research regimes. Together these developments could facilitate improvements in diagnostics, therapeutics and patient management.

By facilitating engagement and creating new routes to participate in research, the hope is that this novel platform could mitigate against existing disparities and help to enfranchise those who lack access to research. However, it is also possible that there will be a group of potential participants who are unable or unwilling to use the platform because they choose not to, although they have the expertise and the resources to do so. We explore the implications relating to this group in section 3.2.

Lack of access to digital technologies

In particular, some participants may be disenfranchised because they lack digital literacy, or computer/internet access. This may be more likely in older patients/participants or those who are disabled or economically deprived. Office of National Statistics data suggests that 80% of households with one adult aged 65 years and over are connected to the internet, representing the lowest proportion of internet connections.[37] Figures derived from the Labour Force Survey in 2019 suggest that of the 4 million adults who had never used the internet in 2019, more than half were aged 75 and over.[38] Indeed most adults aged 75 and over (53%) had not used the internet in the last three months preceding the 2019 Labour Force Survey.[39] In 2019, 7.5% of adults had never used the internet compared to 22% of disabled adults (i.e. those who self-assess that they have a disability in line with the Equality Act).[40] This suggests that there is potential for the platform to have a differential impact on users and non-users. If Wise's findings[41] are replicated here, it seems likely that as reliance on the platform increases, so efficacy may increase with potential to aggravate existing inequalities. Some claim that this lack of access to digital information and communication technologies, compounded by a lack of digital literacy, leads to a 'vicious cycle' of disadvantage.[42]

[37] Serafino P. Exploring the UK's digital divide. Office for National Statistics. 2019, figure 8, citing Office for National Statistics – Internet Users, Labour Force Survey.

[38] Prescott C. Internet users, UK: 2019. Office for National Statistics released 24 May 2019. These figures on internet usage are drawn from three months of internet usage from January – March 2019 and are derived from the Labour Force Survey. https://www.ons.gov.uk/businessindustryandtrade/itandinternetindustry/bulletins/internetusers/2019

[39] Ibid.

[40] Ibid.

[41] Wise PH. Emerging technologies and their impact on disability. In: *The Future of Children.* 2012; 22(1): 169-191

[42] Baum F, Newman L, Biedrzycki K. Vicious cycles: Digital technologies and determinants of health in Australia. *Health Promotion International*. 2014; 29(2): 349-360.

Expectations of consumption rather than research

Certain aspects of the platform, such as the potential for non-medical information to be returned to participants, may imply a consumer/provider model, regulated by contract, rather than an altruistic model predicated on a research paradigm, and could alter participants expectations of the project as a whole. This could fuel concerns voiced by some commentators such as Savard, that pursuit of personalised prevention through gaining predictive genetic information may 'simply privilege consumption' with the danger that it will undermine universal care, increase health inequities and 'lead to further injustice for those already without a voice.'[43] This tension may be problematic, given that the information gained about an individual within the 100,000 Genomes Project relies heavily on public funds and knowledge gained from population studies.

Disparities in genomic representation

Another ethical challenge concerns the potential disparity between the amount of genomic information about populations of European descent (Caucasian populations) and those of other ancestries. This disparity is a legacy from early genetic studies, and reflects the fact that genetic discovery efforts to date heavily under represent non-European populations globally. Although considerable efforts are being made to make genetic and genomic databases more representative, the practical effect of this disparity is that analysis and interpretation of variants in non-European individuals are less accurate than in European individuals. Substantial global efforts are being made to supplement and augment genomic data about non-European populations to mitigate against this historic inequity.

Operational issues

Moving forward, it is important that the platform is developed with a view to signalling a commitment to minimising existing and potential inequities and inequalities wherever possible. For example, careful decisions should be made about the language used on the platform, and whether translations into other languages are available. There should also be sufficient provision for those with disabilities (such as specific assistance for those with impaired sight).

While eligibility for preliminary testing of the platform will be targeted at participants with a shared 'higher level' condition, once the platform becomes operational, it is important that participants who have extremely rare conditions are not disqualified from using the platform because they do not share a phenotype or genotype with other participants.

[43] Savard, J. Personalised Medicine: A Critique on the Future of Health Care. *Bioethical Inquiry*. 2013; 10: 197–203.

Consideration: Active efforts should be taken to optimise the accessibility of the platforms, so that the users can be as representative as possible.

Consideration: Careful assessment should be made of the potential for causing or aggravating inequalities and inequities at each point that a decision about implementation is made during the development of the platform. Wherever reasonably possible, potential inequalities and inequities should be minimised by providing alternatives for those who cannot or chose not to use the platform.

2.3 Consent and capacity

The general ethical and legal framework relating to consent and capacity has a profound impact on healthcare and medical research. For this reason, in this section we address the framework relating to consent and capacity before applying it to the specific context of the participant platform. Consent is a key ethical principle underpinning much of healthcare and medical research. Consent also serves as the legal basis for many activities that form part of healthcare and research, either as a core activity (such as patients consenting to being touched as part of care, and to interventions such as testing or diagnosis) or to support care and research (such as agreeing to data sharing). In legal terms, there are many synergies between the processes supporting consent, such as information provision, and mechanisms that are used to meet expectations of participants through high levels of transparency (as described in section 2.1).

The Health Research Authority advises that -

> 'Seeking informed consent is central to the conduct of ethical research. Seeking informed consent properly respects a person's right to determine what happens to them'.[44]

Thus consent can advance individual autonomy by ensuring that a potential research participant has a genuine understanding of what is proposed and is how it is relevant. However, in the research context, routinisation, a desire for comprehensiveness, and increasing regulatory complexity have resulted in longer, complex patient facing materials (information sheets and consent forms) arguably at the expense of genuine understanding.[45] Onora O'Neill has argued that consent should properly be viewed as a 'propositional attitude to a proposition', and that a genuine consent is contingent on patients being neither

[44] NHS Health Research Authority. HRA Guidance on applying a proportionate approach to the process of seeking consent 2019 v1.02 Final. [Cited on 13 Jan 2021.] Available from: https://s3.eu-west-2.amazonaws.com/www.hra.nhs.uk/media/documents/Proportionate_approach_to_seeking_consent_HRA_Guidance.pdf

[45] O'Neill O. Some limits of informed consent. *Journal of Medical Ethics.* 2003; 29: 4-7.

coerced nor deceived, and also having control over the amount and form of information they receive. Providing 'easy' ways of rescinding that consent demonstrates that patients have not been coerced. Patients should be free to change their minds about participation and be able to revoke or withdraw their consent at any time. Supporting the autonomous choices of patients in this way can be logistically challenging, especially if individual patient data has been anonymised or shared.

A proportionate approach

Concerns about increasing length and complexity of materials and processes undermining the spirit and legal validity of consent has fostered a more proportionate approach. Thus the Health Research Authority supports methods and procedures which are proportionate to the nature and complexity of the research; the risks, burdens and potential benefits (to the participants and to society) and the ethical issues at stake.[46]

Health Research Authority guidance supports seeking, confirming and documenting consent by electronic methods, provided that these facilitate a two-way communication in real time, that the participant's identity can be assured, and a decision recorded 'in writing'. This guidance advises that using advanced methods of validation involving tracing a participant's signature with a stylus may be needed for research involving more than minimal risk.[47] These methods have been invaluable in facilitating research to proceed despite the restrictions imposed by COVID, but the onus remains with the researcher to ensure the participant (or their legal representative) has understood the information provided, and that their confidentiality has been maintained. Specific patient populations may have special information needs and might need additional resources to support their participation. When communication is limited to electronic methods, being sure that the participant has fully understood what is being proposed might be challenging.[48]

Key elements of consent

The common law (namely judge-led decisions through courts and tribunals) requires that participants be informed, in broad terms, of the **nature and purpose**[49] of the research and the **material risks, benefits and**

[46] NHS Health Research Authority. HRA Guidance on applying a proportionate approach to the process of seeking consent 2019 v1.02 Final. [Cited on 13 Jan 2021.] Available from: https://s3.eu-west-2.amazonaws.com/www.hra.nhs.uk/media/documents/Proportionate_approach_to_seeking_consent_HRA_Guidance.pdf

[47] Ibid.

[48] MHRA, Health and Care Research Wales, NHS Scotland, Health and Social Care, NHS Health Research Authority. Joint Statement on seeking consent by electronic methods. 2018 [cited 24 March 2021.] Available from: https://s3.eu-west-2.amazonaws.com/www.hra.nhs.uk/media/documents/hra-mhra-econsent-statement-sept-18.pdf

[49] *Chatterton v Gerson* [1981] QB 432, [1981] 1 All ER 257.

reasonable alternatives.[50] A valid consent must be freely given (with no undue influence); by a person with necessary mental capacity; who has been adequately informed. Those asked to give their consent to a research study should, in addition, neither be coerced nor deceived; be able to control the amount of information they receive, and have the opportunity to withdraw a previously given consent. There is increasing emphasis on conveying information in a way that is understandable and accessible, as well as layering information so that it can be provided in ways that meet individual patient needs and desires.

The ubiquitous nature of consent in informing, bounding and justifying care and research provides a framework for exploring different types of consent. Just as an affirmative action by a patient or participant can authorise a proposed action or intervention (in Onora O'Neill's terms), specific and informed consent can, under the General Data Protection Regulation (GDPR)[51] or version of that Regulation adopted by the United Kingdom (the UK GDPR), provide a legal basis[52] for processing personal data[53] (i.e. identifiable data relating to an individual). In practice however, the conditions placed on utilising consent as a legal basis under the GDPR/UK GDPR mean that other legal bases such as contract and legitimate interests provide more flexible and practical alternatives. Indeed, the Health Research Authority has advised that the legal basis of consent should be avoided whenever possible.[54]

In addition, a narrower form of consent[55] (explicit and informed consent for specified purposes) can legitimate processing of special categories of data including genetic, biometric and health data. Here, consent provides a more attractive option where the alternatives of legitimate interests and contract are not available. A further form of consent may serve to authorise the disclosure of confidential health information in some circumstances, even where less information might be available that would typically be associated with an 'informed consent'.[56]

[50] *Montgomery v Lanarkshire Health Board* [2015] UKSC 11, [2015] AC 1430.

[51] Council Regulation (EC) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [2016] OJ L119/1.

[52] UK GDPR Article 6 (1).

[53] UK GDPR Article 4(1) defines personal data as 'any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person'.

[54] NHS Health Research Authority. *Consent in research*. 2018 [cited on 24 March 2021.] Available from: https://www.hra.nhs.uk/planning-and-improving-research/policies-standards-legislation/data-protection-and-information-governance/gdpr-guidance/what-law-says/consent-research/

[55] UK GDPR Article 9(2)(a).

[56] Chico V, Taylor M. Using and Disclosing Confidential Patient Information and The English Common Law: what are the Information Requirements of a Valid Consent? *Med Law Rev*. 2018; 26(1): 51-72.

Mental capacity

The elements of a legally valid consent point to a second consideration which is extremely important for this project, namely that the person giving the consent has the necessary mental capacity. Determining 'necessary' mental capacity is not straightforward even when done in a face-to-face interaction. The Mental Capacity Act[57] defines a lack of capacity in relation to a matter

> 'if at the material time he is unable to make a decision for himself in relation to the matter because of an impairment of, or a disturbance in the functioning of, the mind or brain.'[58]

A person is unable to make a decision for himself if he is unable to understand, retain, use or weigh the relevant information or to communicate his decision.[59] Sections 3 and 4 elaborate on this to enforce a person-centred approach[60] which takes account of all relevant circumstances[61] and 'as far as reasonably practicable' supports the person to participate 'as fully as possible in any act done for him and any decision affecting him'[62] ensuring a heavily context driven approach. Section 30 of the Mental Capacity Act provides additional safeguards for research, stipulating the circumstances in which intrusive research can be carried out, for example as a clinical trial, connected with an impairing condition, or the treatment of a person who lacks capacity, and the approvals or safeguards that must be put in place.[63]

Consent and the platform

There is an emerging literature on participant-centric initiatives, and the approach that they have taken regarding consent. The design choices that are adopted are heavily dependent on the target user group and on the context of ongoing patient engagement and data collection. In their review of participant-centric initiatives, Hamakawa et al[64] noted that a specific consent model was adopted in half the participant-centric initiatives studied, whereby a participant's consent is requested each time they participate in a new study. Only one initiative adopted a broad consent model (PatientsLikeMe) and a minority (14%) implemented a dynamic consent model allowing their participants to make granular consent decisions over time.

[57] Relevant legislation in the devolved nations are the Adults with Incapacity (Scotland) Act 2000, and the Mental Health (Northern Ireland) Order 1986 (outgoing) and the Mental Capacity Act (Northern Ireland) 2016 (incoming).
[58] Mental Capacity Act 2005, s 2(1).
[59] Mental Capacity Act 2005, s 3(1).
[60] 'appropriate to his circumstances'.
[61] Mental Capacity Act 2005, s 4(2).
[62] Mental Capacity Act 2005, s 4(4).
[63] Mental Capacity Act 2005, s 30-33.
[64] Hamakawa N, Nakano R, Kogetsu A, et al. Landscape of participant-centric initiatives for medical research in the United States, the United Kingdom, and Japan: Scoping review. *Journal of Medical Internet Research*. 2020; 22(8): 1–17.

In the development phase of this platform, selected participants will be invited to access the new platform and register to use the service. By responding to a registration request, potential participants will access a landing page outlining key objectives for the platform, and information about how data relating to them will be processed, held and linked. Participants are asked for explicit, affirmative consent for Sano to share their profile with Genomics England 'to connect you with the most relevant research and keep records up to date'. After clarifying the registration status, (on whose behalf the account is set up), and consenting to terms of service and privacy policy by clicking a 'register' tab, an activation link will be sent via email, to verify the identity of the user. Once this is activated, the participant will be able to access the platform.

Meeting the consent criteria in the UK GDPR

Establishing a valid consent, and ensuring that users have capacity to join and continue to use the patient portal, are key elements of the proposed platform. However, as with many remote platforms, these objectives are logistically challenging, especially given its long-term nature. The model platform envisages that, by pressing the 'register' tab, participants will affirm their consent to proceed. Article 4(11) of the UK GDPR defines consent as 'any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement of by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her'. The prototype includes the terms of service and the privacy policy as optional additional information which is accessible by clicking on the link. There is no guarantee that a potential participant has read these terms before registering for the service.

> Consideration: As part of the onboarding process, participants should have access to a plain language summary of key aspects of the terms of service and privacy policy so that they are sufficiently informed about the platform and its operation for the consent to be informed.

Determining capacity

Determining a person's capacity to make decisions for themselves, for example to agree to a suggested intervention in health care, or take part in medical research forms an integral part of health care and medical research. When done face to face, health professionals use verbal and non-verbal cues to make these assessments. In the context of the platform, determining that a participant is eligible through demonstrating capacity to consent raises additional challenges, in part due to virtual technologies being used and the lack of a mechanism for corroborating a person's competence. As with all longitudinal research, although the question of capacity is raised most acutely at the time a person is invited to join the platform, it may also be an ongoing issue for those participants who have registered to use the portal but who lose capacity over time or who subsequently experience fluctuating capacity, so that they have capacity to consent to participate intermittently.

Where a single email address is used for multiple participants within a family (e.g. a child and his/her parents), reliably confirming the identity of the person registering for and using the platform may be a logistical challenge rather than a question of medical assessment.

As described above, the Mental Capacity Act enforces a patient centred interpretation of capacity, promoting an approach where the individual is supported to participate as fully as reasonably practical, taking account of all relevant circumstances. Instead of using stated age as a proxy for capacity (as the GDPR does for use of information society services),[65] the minimal viable product platform provides for invitation letters to be sent to eligible participants.

It is not clear on what basis 'eligibility' will be determined. Monitoring the ongoing capacity of family members of patients, who are themselves participants of the 100,000 Genome Project may be feasible through ongoing linkage of electronic health records.

> Consideration: In order to reduce the potential for those lacking capacity to be inadvertently recruited, checks on the eligibility of potential participants should be carried out before invitation letters are issued. These could include an assessment of the potential participant's age, record of impaired or reduced capacity in electronic patient records, legal power of attorney, or history of reducing and fluctuating capacity relating to a diagnosis.

In the absence of notification, the GEL programme makes an assumption that participants continue to have capacity for a limited period of five years. The protocol provides for NHS Genomic Medicine Service to implement a specific check of an adult patient's capacity around 5 years after their consent was first given to join the library.[66] It is not clear whether a similar assumption has been made in relation to demonstrating consent for the participant portal, and whether the tests done by clinical staff to demonstrate continuing consent to remain within the Project (which have been incorporated within the commissioning process) could or should be adapted to cover continuing consent to use the portal (including data provision for research purposes). It also is not clear whether the expectation that clinical staff monitor the ongoing capacity of participants extends to family members of participants, who are themselves participants of the 100,000 Genomes Project, unless such monitoring is triggered by loss of capacity being reported to GEL by NHS GMS. For some subsets of cohort members (e.g. cancer patients), there will not necessarily be an ongoing clinical relationship with the NHS GMS, and it might be more appropriate for general practitioners to be involved in this process.

[65] GDPR Article 8: Conditions applicable to child's consent in relation to information society services.

[66] Genomics England. The National Genomic Research Library protocol v5.1. 2020. Section 11.1.7, p.71. Available from https://www.genomicsengland.co.uk/national-genomic-research-library/

Notwithstanding these challenges, this regulatory requirement reflects the ethical principle of inclusivity which emphasises the importance of supporting as many people to participate as possible/practicable. Depending on the application, it may be appropriate for those lacking capacity to be supported in participating, especially if carers can support collection of patient related outcomes (PROMS) through their detailed knowledge of the person, e.g. when they are in pain, distressed or tired.

> Consideration: In line with the Mental Capacity Act, those lacking capacity, or having fluctuating capacity should be supported to contribute to the platform. Depending on the intended use of the platform for this group, this may require a person with the appropriate legal authority to act as consultee for the person lacking capacity. It may also require appropriate material to be developed to facilitate the involvement of the person lacking capacity to the extent that they are able.

However, eliciting ongoing contact with non-responsive participants might be undesirable for them, or even intrusive and upsetting, particularly if the patient suffers from a rare disease or cancer which is life-limiting. This may also be highly relevant to the GenOMICC COVID-19 study, where many patients who have been recruited with severe COVID infections may have died after being recruited to the study. If the condition of the patient has deteriorated or they have died, contact from platform administrators could be insensitive or harmful. More efficient checks may be facilitated by NHS Digital moving to a new patient identifier system.[67]

There should be clarity about the extent to which the two activities (inclusion in the research database and participation in the participant portal) should be fully aligned. For example, this includes whether withdrawal from the research database or the NGRL necessarily results in withdrawal from the participant portal and vice versa, or whether the two activities can operate independently.

> Consideration: Additional work is needed to determine the thresholds for demonstrating continuing capacity or the onset of incapacity for those participants who register for the participant portal, and the process by which this could be achieved.

> Consideration: Reasonable checks should be put in place to minimise the potential for causing distress through contacting participants who have died or who are severely ill.

[67] Imminent changes to the Master Person Service person identifier may facilitate more efficient linkage of episodes relating to individual patients by GEL: https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/hospital-episode-statistics/hospital-episode-statistics-data-changes-in-2021

2.4 Privacy, data protection and confidentiality

A key set of ethical and legal issues for any participant platform relates to privacy, confidentiality and the framework for the protection and governance of personal or identifiable data. These apply both to the uses and storage of participant information via the platform and to any sharing or linkage of data between the platform, the genomic initiative it serves and other third parties.

Privacy and confidentiality

Privacy and confidentiality are core components of the ethical and legal framework for medical research (e.g. the World Medical Association's Declaration of Helsinki, as amended October 2013).[68] The right to privacy is contained in the European Convention on Human Rights (1950)[69] which provides for a right to private life (Article 8) including (as interpreted by the European Court) a right to self-determination and a right to privacy of personal information. This is incorporated in UK law through the Human Rights Act[70] in 1998 which also requires UK courts to take account of decisions of the European Court of Human Rights. Under this framework, a platform must ensure that the privacy of individual participants and their private information are protected from unauthorised invasion. This will require a range of technical and organizational measures to safeguard the data collected by and visible on the platform. For example, it may require specific measures, for example a passcode or biometric login, to ensure that particularly sensitive information is not accessed by family members or others who could come across a device.

The complementary ethical and legal duty of confidentiality protects against unauthorised disclosures of private or confidential information (e.g., without consent or a lawful basis). Under the UK common or judge made law, privacy and confidentiality will apply to information in circumstances that give rise to a reasonable expectation that privacy will be protected.[71] This has been found to clearly apply to medical information and is likely to apply to most information imparted in medical research. For the platform this means that disclosures of private and confidential information, for example sharing with researchers, will only be authorised with a valid explicit consent or other lawful means.[72] Although there are three main

[68] WMA Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Subjects, Adopted by the 18th WMA General Assembly, Helsinki, Finland, June 1964 and amended by the: 64th WMA General Assembly, Fortaleza, Brazil, October 2013.

[69] Convention for the Protection of Human Rights and Fundamental Freedoms as Amended by Protocols No.11 and No.14 (opened for signature 4 November 1950, entered into force 3 September 1953) CETS No. 005.

[70] Human Rights Act 1998.

[71] *Campbell v MGN Ltd* [2004] UKHL 22 at [14], [2004] 2 AC 457 at [14], [2004] 2 All ER 995 at [14] per Lord Nicholls of Birkenhead ('this cause of action has now firmly shaken off the limiting constraint of the need for an initial confidential relationship'); *Mosley v News Group Newspapers Ltd* [2008] EWHC 1777 (QB) at [7], [2008] EMLR 679 at [7], [2008] All ER (D) 322 (Jul) per Eady J ('the law now affords protection to information in respect of which there is a reasonable expectation of privacy, even in circumstances where there is no pre-existing relationship giving rise of itself to an enforceable duty of confidence').

[72] Other alternatives that may be relevant are a decision by the Confidentiality Advisory Group of the Health Research Authority that it is in the public interest for the duty of confidentiality to be set aside (s.251 approval) or a

groups of participants who may submit their genomic and phenotypic data into the NGRL – participants in the 100,000 Genomes Project, patients who are part of research conducted by the GenOMICC consortium and patients who have had a test through the NHS Genomic Medicine Service – the disclosure of their confidential information for research purposes will generally have been authorised by their explicit consent given to healthcare or research professionals when they agree to take part in research.[73]

For most uses of confidential information on the platform, consent is likely to provide the most suitable form of authorisation. It is possible that some functions could ultimately interact with individual healthcare, where the sharing of data between healthcare professionals may also be authorised by implied consent.[74] However, as we discuss in the context of longitudinal data collection in section 3.6 below, it may be harder to imply that consent has been granted in the digital platform context.

Finally, during the COVID-19 pandemic the Secretary of State for Health has made use of Regulation 3(4) of the Health Service (Control of Patient Information) Regulations 2002 (COPI) to direct the disclosure of some confidential medical information without consent. This could become relevant for the sharing of some information with the platform if it is addressing COVID-19 or communicable disease and the Secretary of State directs other organisations to share information with it, as has been the case with UK Biobank.[75]

Data protection

Closely allied with privacy and confidentiality is the ethical and legal imperative for the protection and control of personal data. Although not universally presented as an independent fundamental right, the Council of Europe has long recognised a right to protection of personal data in its 1981 Convention 108[76] and recently updated Convention 108+. This has been echoed by the European Union which also recognises an independent right to data protection in the EU Charter of Fundamental Rights and which has been heavily influenced by Convention 108 in its recent General Data Protection Regulation. With significant emphasis on transparency and individual rights to control and scrutinise data, the European approach to data protection is intimately linked to autonomy, human dignity and respect for persons as well as privacy.

disclosure in the public interests to protect others from significant risk of serious harm (e.g. *ABC v. St George's Healthcare NHS Trust & Ors* [2020] EWHC 455 (QB), [2020] 2 WLUK 400).

[73] It is possible that some patient information is disclosed to Genomics England from other studies where consent has already been obtained or even perhaps where there has been a s.251 decision from the Confidentiality Advisory Group of the Health Research Authority that the common law duty of confidentiality may be temporarily set aside for the purposes of specific research.

[74] General Medical Council. Confidentiality: Good practice in handling patient information. 2017, 13.

[75] https://www.gov.uk/government/publications/coronavirus-covid-19-notification-of-data-controllers-to-share-information/coronavirus-covid-19-notice-under-regulation-34-of-the-health-service-control-of-patient-information-regulations-2002-biobank

[76] Convention for the Protection of Individuals with regards to Automatic Processing of Personal Data [1981] ETS No.108.

The UK has left the EU but its data protection law remains almost exactly the same in the form of a 'UK GDPR' which is defined and amended by the Data Protection Act 2018. This applies to all forms of 'processing' of 'personal data' (which includes storage and almost anything that can be done with such data). Unlike privacy and confidentiality, the test for 'personal data' solely relates to information relating to an identified or identifiable natural person (Article 4(1))—regardless of whether it has a private or public nature. What matters most is if there is a risk that an individual could be identified from that data either on its own or in combination with other information. A broad contextual risk assessment (described in Recital 26) should be made, and 'account should be taken of all the means reasonably likely to be used' by 'another person' taking into consideration 'the available technology at the time of the processing and technological developments'.

When processing 'personal data', data controllers and processors are obliged to abide by a range of principles, rights and obligations. Among others, the UK GDPR sets principles for processing fairly, lawfully, and transparently. The law requires that processing is limited and that as minimal data as possible are used. Data must be accurate and not stored for indefinite periods and measures should be used to ensure the security of data and to demonstrate accountability for compliance with the law. These principles are supplemented by many more detailed obligations and data subject rights, including a right of access to data and rights to information about how personal data are being processed.

Safeguarding personal data – Because genomic data are multi-dimensional and capable of being linked to many other forms of information they are considered highly identifying,[77] (although not, in our view, *inherently* identifying even in full).[78] It has long been recognised that it is possible to identify an individual in otherwise 'anonymised' health data using only a small number of single nucleotide polymorphisms (SNPs),[79] so a cautious approach to safeguarding privacy is justified. This means that even the removal of obvious identifiers and their replacement with a code that is held separately and securely (pseudonymisation) will frequently be insufficient to reduce the risk of identification of an individual.[80]

The broad scope of personal data means that platforms need to be cautious about the potential for even innocuous participant data to be connected to an individual in combination with other information. For example, if connecting one participant to a research project could be used by someone else to cross

[77] Finnegan T and Hall A, Identification and genomic data. PHG Foundation. 2017. Available from: https://www.phgfoundation.org/documents/PHGF-Identification-and-genomic-data.pdf

[78] Mitchell C, Ordish J, Johnson E, Brigden T, Hall A. The GDPR and Genomic Data: PHG Foundation report on the impact of the General Data Protection Regulation and Data Protection Act 2018 on Regulating Genomic Technologies in Healthcare for the Information Commissioner's Office. PHG Foundation. 2020. Available from: https://www.phgfoundation.org/report/the-gdpr-and-genomic-data

[79] Shabani M, Borry P. Rules for processing genetic data for research purposes in view of the new EU General Data Protection Regulation. *European Journal of Human Genetics*. 2018; 26(2): 149-156; Lin Z, Owen A.B. Altman R.B. Genomic Research and Human Subject Privacy. *Science*. 2004; 305(5681): 183-183.

[80] Mourby M, Mackey E, Elliot M, et al. Are 'pseudonymised' data always personal data? Implications of the GDPR for administrative data research in the UK. *Computer Law & Security Review*. 2018; 34(2): 222-233.

reference data from other research and single out their unique data, there could be a reasonably likely risk of identification. There are strict controls against this possibility within the Research Environment. It is important that the collection and processing of personal data outside the Research Environment and via the platform also safeguards against these risks. For example, by strictly monitoring and controlling access to data and keeping abreast of new threats to otherwise anonymous datasets.

> Consideration: Where data are collected and processed outside the Research Environment, there should be careful monitoring and safeguards in place to ensure that risks of identification are minimal. This requires consideration of new technologies or developing threats to otherwise anonymous data.

Lawful processing – For participant platforms and the organisations who feed into them or receive data from them, compliance with data protection requires that they establish a lawful basis (Article 6) for each form of processing (for example, collecting profile data directly from participants, sharing data with researchers or potentially providing results and information back to participants). Consent is one option (Article 6(1)(a)) but the standard for and interpretation of consent in data protection law means that it is not necessarily the most straightforward legal basis in the research context. Indeed, the Health Research Authority (HRA) actively advises researchers not to rely on consent[81] for multiple reasons, including that consent is inappropriate where there is a 'clear imbalance' between the data subject and controller, which may particularly be the case where the controller is a public authority (although this will not apply to GEL as it does not consider itself a public authority for data protection purposes).[82] The HRA also flags the implications of the need to respect withdrawal of consent and the data rights that follow from this legal basis as potentially detrimental to the validity of research. Compounding these difficulties is a debate about how specific consent must be to data processing. Despite some recognition in the GDPR (Recital 33) that it is often not possible to fully specify the purposes of data processing at the outset of scientific research, this has generally been interpreted very restrictively by the European data protection authorities (the European Data Protection Board, or 'EDPB' and its predecessor the Article 29 Working Party, or 'WP29'). The guidance is that broader consent should be used as a limited exception which should be remedied with additional specific consents rather than allowing a one-off broad consent.[83]

There are alternatives such as the public interest (Article 6(1)(c)), legitimate interests (Article 6(1)(f)), or that processing is necessary for the performance of a contract (Article 6 (1)(b)) which are potentially better

[81] NHS Health Research Authority. Consent in research. Available from: https://www.hra.nhs.uk/planning-and-improving-research/policies-standards-legislation/data-protection-and-information-governance/gdpr-guidance/what-law-says/consent-research/ [Accessed 11th February 2021].
[82] Recital 43 GDPR (also incorporated as an interpretative aid to the UK GDPR).
[83] Hallinan D. Broad consent under the GDPR: An optimistic perspective on a bright future. *Life Sciences, Society and Policy*, 2020; 16(1), 1–18. https://doi.org/10.1186/s40504-019-0096-3

suited to this situation. Genomics England's privacy policy makes clear that legitimate interests provide the legal basis for almost all data processing for research across the different cohorts[84] whereas the current Sano privacy statement provides a contractual basis for most of its current personal data processing activities.[85] Whether this is appropriate for some of the potential activities enabled by the co-developed platform will depend on the precise nature of the processing. However, it could potentially be surprising for participants to learn that processing is based on a contract with the platform in this context. Although the provision of a service does not require financial payment and therefore the service provided by the platform, and the relevant terms and conditions which apply, may—as a matter of law—form a contractual relationship, the health research context is generally not characterised by participants regarding themselves as party to a commercial or contractual relationship. In this case, it could be preferable to consider alternative legal bases such as legitimate interests or even consent, where suitable. This does not apply to processing where a further service (such as direct to consumer genetic testing) is provided and of direct benefit to the participant.

> Consideration: In the health research context it may be surprising to participants that their data are being processed for the performance of a contract, even if the platform provides some services of benefit to the participant.

Special category data – Where data include health or genetic data ('special category data'), Article 9 of the GDPR/UK GDPR requires that a further condition, such as explicit consent[86] (Article 9(2)(a)), or research purposes (Article 9(2)(j)), is met. Genomics England's privacy policy makes clear the Article 9 condition for processing of health and genetic data is Article 9(2)(j), which allows processing for scientific research purposes in accordance with safeguards contained in the UK GDPR and DPA 2018. This should also be appropriate for most of GEL's processing in relation to the platform. However, it may not apply so easily to the processing involved in the invitation to join the platform (outside the specific context of current research and testing phase), to the extent that it involves processing participants' special category data.

This is because the UK law implementing the Article 9(2)(j) scientific research purposes option requires that, if research includes 'measures or decisions with respect to a particular data subject', it must have been approved by a research ethics committee.[87] An invitation to join the platform could fall within this as it may influence a range of further measures and decisions, such as invitations to take part in further

[84] The situation may be slightly different for COVID 19 research which is a 'ring-fenced' part of the Genomics Research Environment and has a potential initial stage for some participants who have registered an interest to join the COVID-19 research, thereby providing consent under art 6(1)(a), and explicit consent under art 9 (2)(a) of the UK GDPR for the follow up contact by GEL.

[85] https://sanogenetics.com/privacy

[86] (the additional explicit element simply requiring a positive statement such as a signed consent form).

[87] Data Protection Act 2018, section 19(3) & (4).

research or even return of some results. In general, this is not a problem as GEL has obtained ethical approval for its Protocol and this also includes the potential for recontacting participants about further research and the development of methods to facilitate clinical trials referral and patient recruitment.[88] However, this platform may have multiple potential functions and it may not necessarily be entirely straightforward describing an invitation to join the platform as recontact for further research opportunities. This will require further consideration but if it is felt that the invitation to join the platform cannot be characterised as part of GEL's approved research purposes it may be better to seek an alternative Article 9 condition, such as explicit consent (Article 9(2)(a)) instead. Despite the challenges of consent already mentioned, for the specific purpose of processing data to invite participants to join the platform, consent could still provide a suitable basis. This is because it would relate to a discrete form of processing for a specific and clear purpose. The challenge here is determining whether participants can be said to have consented to recontact for the purposes of an invitation to join a platform (as opposed to recontact for a specific research project).

> Consideration: if the invitation to join the platform relies on the processing of a participant's health, genetic or ethnicity data, there may need to be consideration of the appropriate legal basis for this 'recontact' by GEL, and further analysis of how this is contemplated in the approved Protocol and participant information.

The platform and further data processors or controllers could also rely on the scientific research basis for their processing (Article 9(2)(j)) but because much of the processing by the platform will involve 'measures or decisions with respect to a particular data subject' this will only extend to activities that fall within the scope of the existing approved Protocol. This means research that deviates from the extensive GEL policies for example, only allowing access to identifiable data within the Research Environment, would fall outside the scope of the approval. Nor would further forms of processing, such as the provision of ancestry results fit comfortably with this.[89]

As an alternative the platform could seek to rely on the Article 9(2)(a) explicit consent, since it is highly likely that participants will be given clear information and free choice about the multiple functions and attendant processing of data by the platform. Where this may become more challenging, is if the platform 'carries out its own' research, in which case this must be clearly specified and opted into by participants or there may need to be further ethical approval for this research if it cannot be carried out in line with the existing NGRL Protocol. This is discussed further in relation to longitudinal data collection in section 3.6 below.

[88] NGRL Protocol 5.3

[89] However, the return of medical results could be conducted in reliance on art 9(2)(h) medical purposes, provided that this is supervised by a healthcare professional subject to a duty of confidentiality (DPA 2018, s11(1)).

Overall, it is likely that different legal bases and Article 9 conditions will need to be chosen for different forms of processing by the platform but if this consent-based approach is adopted, it is important that there is a clear distinction between that processing and any subsequent research data processing carried out by researchers or third parties, perhaps with the assistance of the platform to collect this data. If this is based on the scientific research data protection basis, the distinctions between this processing (and the differences between subsequent rights and obligations that apply) must be made clear to participants.

> Consideration: Where data are processed for different purposes and rely on different legal bases or Article 9 conditions, this should be made clear to participants. In particular, it should be clear when their data are being used in further research and who is responsible for it.

Data protection rights and obligations – Data protection law sets out a range of data subject rights and further obligations that may apply. Where multiple parties are jointly determining the purposes and means of the processing (i.e. the type(s) of data to be processed, the period for which they would be retained, from which data subjects the data would be collected, who will have access to data and the recipients of data[90]) they will need to work together to determine their respective responsibilities for compliance with rights and obligations and how to make this clear to participants (Article 26).

Determining which rights or obligations apply will depend on the precise nature of the processing and the legal basis and Article 9 conditions that have been chosen. However, some rights and obligations are relevant to the development of the platform in general. As already mentioned, one core obligation contained in the UK GDPR is transparency which requires that information is provided to participants, either where data are directly collected from them (Article 13) or as a secondary use of data (Article 14). This means that information should be available in the form of a privacy notice or other clear and accessible explanation, about (among other things) how 'personal data' are being processed, who (at least in terms of categories of recipients) it is being shared with, the legal basis for processing and how participants or 'data subjects' can enforce their rights. This should be provided when participants register with the platform as well as being available on the relevant parties' websites.

A further data subject right that may be particularly relevant to the platform is the 'right of access' under Article 15 UK GDPR. This provides a right for all platform participants to request a copy of their 'personal data' (as described above, this includes all the data which relate to them as an identifiable individual). There are logistical complications of this in the genomic context as a 'complete' genome may be between 80-200Gb and in a raw format (such as a BAM file) which could be a surprise to recipients. GEL already emphasise that they can provide this to data subjects but that they will not help them to interpret the data

[90] EDPS Guidelines on the concepts of controller, processor and joint controllership under Regulation (EU) 2018/1725, 3.1.3.

and that this information can only come from their healthcare team.[91] Because the genomic data held by GEL will not initially be incorporated within the platform, this possibility may not arise outside the normal GEL framework. However, a data subject may request their data held by the platform, which could include responses to surveys, data collected in other ways and the results of analysis of this data if these results are in identifiable form. Managing such requests may not be straightforward where the data also relate to other individuals (as is frequently the case in this context) and where they constitute data concerning health.

A request may be refused to the extent that doing so would involve disclosing 'information relating to another individual who can be identified from the information', unless the other individual has consented or 'it is reasonable to disclose the information to the data subject without the consent of the individual'.[92] Where data held by the platform relate to multiple family members there may need to be a mechanism for seeking their consent to disclosure and, in extreme circumstances if consent is refused, for determining whether it may be reasonable to disclose if consent is refused.[93] No health data can be disclosed by the platform or any controller who is not a health professional unless they have obtained the opinion of a health professional that disclosure 'would not be likely to cause serious harm to the physical or mental health of the data subject or another individual'.[94] This does not apply to the extent that the platform host or data controller is satisfied that the data concerning health has already been seen by, or is within the knowledge of the data subject.[95] Finally, where the data are being processed for scientific research purposes, the right of access will not apply 'to the extent that the application of these provisions would prevent or seriously impair the achievement of the purposes in question' so long as no research results or statistics are published in an identifiable form.[96] This could exempt some research data and results from a right of access but many other forms of data (and research which would not be adversely affected) collected and processed via the platform will be within the scope of a data access request.

[91] https://www.genomicsengland.co.uk/understanding-genomics/data/participant-data-requests/
[92] Data Protection Act 2018, sch 2 para 16
[93] A controller must have regard to all relevant circumstances and a range of additional factors set out in sch 2 para 16 (3).
[94] This is called the 'serious harm test' and the restriction for non-healthcare professionals is contained in DPA 2018, sch 3 para 6.
[95] Data Protection Act 2018, sch 3 para 5.
[96] Data Protection Act 2018, sch 2 para 27 (1), (3).

Consideration: The platform may require a data access review procedure with input from an appropriate healthcare professional to ensure that the disclosure of health information will not seriously harm the data subject or another person (e.g. their family members). There may also need to be a procedure to obtain consent from relatives for disclosure of data that could identify them or determine if disclosure is reasonable if they refuse consent.

A final set of data protection considerations that will be significant for the platform are an array of obligations to ensure the privacy and security of the potentially sensitive data they will be processing. These include the requirement for a data protection impact assessment (Article 35) where processing is 'likely to result in a high risk to the rights and freedoms' of the participants, which is likely to be the case where data are highly sensitive (as genomic data are), derive from 'vulnerable data subjects' and are processed for evaluation or health prediction purposes.[97] Articles 24 and 25 require the platform and other data controllers to implement 'appropriate technical and organisational measures' and to ensure 'data protection by design and default'. This means there will need to be a context-sensitive and proportionate implementation of safeguards, such as pseudonymisation, to fulfil data protection principles and protect the rights of data subjects. GEL has developed robust safeguards across its own databases already and this will shape the nature of the initial interaction between those data and any further data collected and processed via the platform.

Interaction between the platform and GEL databases

GEL's information governance is centred on protection of participant data within a secure and restricted Research Environment containing de-identified data. Only in rare circumstances are non-generalised or truly anonymised data allowed to leave the Research Environment. This makes it unlikely that participant data will be shared directly with the platform. The alternative is for any necessary linkage, processing or analysis of data collected by the platform in combination with participant records held by GEL to be carried out within the Research Environment, either by bringing data into the Research Environment, or by provision of individual level data (such as NHS number) to facilitate cross-referencing of records by GEL. It should not be possible for the platform and researchers using the platform to be able to identify an individual participant from other GEL data they may access.

This framework should give rise to a one-way system, whereby researchers are only able to analyse genomic data in combination with newly obtained platform data within the GEL controlled Research Environment. Any analysis conducted outside the GEL framework will therefore be limited to the data

[97] Data Protection Working Party. Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679. 2017.

collected via the platform and other external data that have been obtained by researchers, including new samples or sequencing of participants directly.

3. Specific applications and specific issues

In this section we review the seven platform features which have been prioritised by participant engagement carried out by Sano and GEL. This iterative analysis moves from evaluating features which note the generic contribution of participants (My contributions), facilitate participants having an increased role in the research process (Patient voice) and receiving generalised updates from research (Research catalogue), through to being matched for suitable trials (Research matchmaker), receiving individualised results (Data and reports), and finally enabling detailed data collection from individual participants (Longitudinal data collection, Wearables).

3.1 My contributions and updates

A system *to allow participants to visualise and track their contribution to research* at first glance, is closely related to the requirement for transparency. Providing details of how data and samples from an individual have been used, acknowledges the reciprocal relationship between patient and researcher which is based on altruistic donation for public benefit. However, many research participants also have an interest in how samples and data about them is used, because they hope that the research facilitated through this donation, will directly or indirectly have a positive impact on their care, treatment or management. Allowing participants to visualise and track their contribution to research could therefore be viewed as fulfilling part of a social contract between participant and researcher.[98] From an ethical perspective, there could therefore be clear benefits associated with this: the provision of information about the research which has utilised participant data and samples goes some way to satisfy the requirements for transparency described in section 2.1 above which is often missing in long-term large scale research. Participants who volunteer data, samples and time to research expect that research to proceed, and to generate findings. They may also hope that those findings will influence their future care, whether directly or indirectly. Indeed, the Health Research Authority 'Make It Public' strategy puts transparency at the heart of health and social care research, embedding the requirements for research registration, reporting and informing participants of the outcomes of research.[99] Despite the expectations that 'information about the findings of the research [should be] available in a suitable format and timely manner, to those who took part in it, unless otherwise justified' [100] 90% of clinical trials have not told participants about findings. The

[98] Generation Genome: Annual Report of the Chief Medical Officer 2016, chapter 16. Available from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/631043/CMO_annual_report_generation_genome.pdf

[99] MakeItPublic: Transparency and openness in health and social care research. NHS Health Research Authority. July 2020. Available from https://www.hra.nhs.uk/planning-and-improving-research/policies-standards-legislation/research-transparency/make-it-public-transparency-and-openness-health-and-social-care-research/

[100] NHS Health Research Authority, The UK Health Departments. UK policy framework for health and social care research v3.3. 2017. Available from: https://s3.eu-west-

Health Research Authority have committed to developing new guidance on how to inform participants about study findings, especially for vulnerable groups. Further engagement events are planned for summer 2021, and future plans include developing sanctions for non-compliance.

Embedding the obligation to return research results to participants is certainly becoming a strategic priority within the UK. However, questions arise about the extent, granularity, and timeliness of the information that is communicated. These communications could cause distress or anxiety if they fail to meet patient expectations in some way. This might be due to the data not being used for the volume or type of research anticipated by the participant, or conversely being used for types of research that were unexpected or unanticipated. The likelihood of this occurring may depend partly on the transparency governing communication between researchers and participants and also on the granularity of the information provided. As previously mentioned in section 2.4, appropriate measures, for example a passcode or biometric login, might be needed to ensure that particularly sensitive information is not accessed by family members or others via the platform.

The right not to know

One of the most challenging aspects might be if the participant becomes aware of their future risks of disease through inclusion in a research project. This might occur if they were recruited on the basis of their genotype. In such a case, communicating the details of a research project and its findings might breach the ethical principle of the 'right not to know' a genetic diagnosis.

International biomedical law supports a right not to know one's genetic status.[101] Article 10.2 of the European Convention states:

> 'Everyone is entitled to know any information collected about his or her health. However, the wishes of individuals not to be so informed shall be observed'

The theoretical foundation of the right not to know is widely regarded as being respect for individual autonomy, legitimised on the basis of the individual's interest in not being psychologically harmed.[102] The limits to this presumed right have been widely debated. Nevertheless, they are, to some extent, enshrined

2.amazonaws.com/www.hra.nhs.uk/media/documents/Final_Accessibility_uk-policy-framework-health-social-care-research_.pdf

[101] Convention for the protection of Human Rights and Dignity of the Human Being with regard to the Application of Biology and Medicine: Convention on Human Rights and Biomedicine ETS No.164; UNESCO Universal Declaration on the Human Genome and Human Rights, adopted unanimously and by acclamation at UNESCO's 29th General Conference on 11 November 1997.

[102] Andorno R. The right not to know: an autonomy based approach. *Journal of Medical Ethics*. 2004; 30: 435-440. doi: 10.1136/jme.2002.001578

in UK legislation. For example the UK Human Tissue Act 2004 preserves a form of the right not to know which allows genetic testing to be done for the benefit of another person (usually the proband seeking care from health professionals).[103] This can be utilised if reasonable efforts have been made to get the donor of the tissue (usually a family member) to decide whether to consent to the use of the material for that purpose and there is no reason to believe that that person has died, refused to consent or lacks capacity to consent. This allows health professionals to approach family members for their consent to testing their genetic material in ways that do not directly disclose genetic risk.[104] In the situation envisaged by the Human Tissue Act, the right not to know is implied by the donor's lack of action in responding to letters addressed to them requesting that they make contact with clinical services.

Some have suggested that the presumption of a right not to know is misguided, and that instead this right should be activated by an individual's explicit choice.[105] Graeme Laurie goes further, arguing that the right not to know is based on 'psychological spatial privacy' justifying a "prima facie" respect for the interest in not knowing, even in the absence of an explicit choice. This places the burden on the individual disclosing information, who prior to disclosure, should ensure that special conditions are fulfilled.[106] This prudent 'privacy enhancing approach' seems most justified where there is no prior indication of an individual's wishes. Indeed, Laurie advocates that because we cannot assume anything about what people want in the absence of actual knowledge about their wishes, then a measure of caution should be exercised in taking disclosure decisions.[107]

Regardless of what the communication contains, it may cause distress for the recipient, particularly if the person that it principally relates to (the family member with a rare disease or cancer) is gravely ill or has died. Care should be taken to guard against this potential distress by instituting routine checks of Hospital Episode Statistics and Death Registration Data prior to making contact.[108]

The participant portal and communication of research contributions

Before joining the 100,000 Genomes Project, potential participants will have been provided with a patient information sheet and consent form, and will have had an opportunity to ask questions about the research project. Consent will have been sought from participants in the pilot projects on a similar basis. Patients

[103] Human Tissue Act 2004, Schedule 4 section 9(3).
[104] Human Tissue Act, Schedule 4, section 9.
[105] Andorno R. The right not to know: an autonomy based approach. *Journal of Medical Ethics*. 2004; 30: 435-440. doi: 10.1136/jme.2002.001578
[106] Laurie G. *Genetic privacy: a challenge to medico-legal norms.* Cambridge. Cambridge University Press, 2002.
[107] Laurie G. A response to Andorno. *Journal of Medical Ethics*. 2004. 30; 439-440.
[108] NHS Digital. Hospital Episode Statistics: https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/hospital-episode-statistics

recruited into the research from the Genomic Medicine Service will have been provided with appropriate patient choice materials.

The 100,000 Genomes Project Rare Diseases Participant Information Sheet[109] states that

- the clinical team will communicate 'any results from further research', and that these might be communicated in an appointment or by letter [page 6/9]
- 'we won't tell you personally every time your data is studied' [page 7/9]
- 'you could get further reports about different conditions in the future' [page 7/9] [adding to list of additional findings]
- 'apart from these additional findings and carrier testing no other information will be looked for or reported'
- 'In future, … the Genomics England project team may contact you. This could be to ask you for more information. Or to invite you to take part in future research. Or to ask you for your views on the project. It is up to you whether you agree to take part in these studies. We may also send you information about the progress of the project'.

Thus the current participant information sheet clearly identifies current limits in the information that will be reported back to the participant (i.e. primary findings, additional findings and carrier testing). Currently there is no mechanism for extra findings from individual research studies to be fed back under the auspices of the 100,000 Genomes Project. However it does envisage information about the 'progress of the project' to be returned.

The consent form reiterates this content allowing GEL or the participant's clinical team to contact the participant to 'ask me to provide more information for the project; ask me to donate further samples if needed in the future; invite me to join other research and send me general updates about the project.' Again, there is no provision for a participant to agree to some of these provisions and not others.[110]

The consent form provides that reporting of results to participants takes place via the clinical team and that no other information will be looked for or reported.[111] There is provision for other results to be provided in

[109] The 100,000 Genomes Project Rare Diseases Participant Information Sheet. Version 2.3. 01-01-2017.
[110] 100,000 Genomes Project consent form version 2.3 for adults with a rare genetic condition and for their adult family members dated 01/01/2017. Page 1/7. Available from https://www.genomicsengland.co.uk/information-for-participants/participant-forms/
[111] Ibid, page 3/7.

the future if further conditions are added to the list of additional findings. However, it is implied that this change in scope will apply to all 100,000 Genome Project participants who have consented to have additional findings returned to them. There is also an option for a participant to consent to carrier testing if of reproductive age and their partner is part of the project and consents to testing.[112]

For patients recruited to NGRL via the Genomic Medicine Service, a record of discussion form applies which incorporates elements of consent to testing for clinical care, and the consent to participation in The National Genomic Research Library. A section on recontact allows NHS staff or Genomics England together with the NHS, to contact participants if the data or samples 'reveals clinical trials or other research that I might benefit from' allowing relevant information to be shared with the NHS clinical team (and by implication shared with the patient).

Currently the degree of contact between GEL and research participants is quite narrowly defined in terms of seeking updates about the project, namely all research facilitated by the 100,000 Genomes Project via the NGRL. If participants are recruited to supplementary research via the platform, it is important that participants should understand the likely benefit from that research. Additional patient facing materials including invitations, information sheets and consent forms should manage these expectations by not guaranteeing certain benefit from participation.

Indirect disclosure of non-validated results to research participants

Research participants may also have information disclosed to them through data access requests by which they can be informed about how identifiable data about them has been used by researchers (see section 2.4).[113] Participants may have rights to access these data under the UK GDPR even if they have withdrawn from the research project and put their samples and data beyond use. There is therefore a potential for indirect disclosure of information which is unexpected, anxiety provoking or burdensome through this route.

> Consideration: When considering the contents of notifications about research contributions, care should be taken not to inadvertently disclose information of which the recipient might be unaware.

[112] Ibid, page 5/7.

[113] Researchers may access identifiable data if of direct clinical benefit and agreed by the Access Review Committee provided that there are appropriate legal agreements and safeguards in place [NGRL Protocol pages 35 and 43/113]. Standard operating procedures are in place to manage data access requests [NGRL Protocol page 53/113].

Consideration: The platform user could be asked explicitly whether they wish to opt-out of receiving communications about certain topics.

Delivery of research updates via a participant platform

Evidence of the benefits and burdens associated with the use of a participant platform for the delivery of research updates is limited. Participant platforms typically are developed with a range of applications in mind, ranging from improved patient engagement (including providing new mechanisms for research updates), to piloting return of non-clinically significant results.[114]

In order to comply with the terms of existing GEL participant and patient information sheets and consent forms, it is important that there is clarity about the provenance of the information that is fed back, the extent to which it is a trustworthy source and subject to appropriate governance and safeguards. The choice of how the platform and materials are badged, is a critical decision.

Participants should also be given information about the sensitivity of the information that is likely to be sent to them via the platform. This might avoid inadvertent disclosure of personal details to other family members or carers, and it might also act as a caution against participants voluntarily sharing sensitive medical data with peers or wider publics, without careful consideration.

3.2 Participant voice

A second proposed application is to develop *a system to allow participants to indicate what kind of research they might be interested in, contributing to research priority setting and early feedback for research proposals,* i.e. to provide a mechanism for patients and participants to have a 'voice' in research. Facilitating participants to have an active and meaningful role in research can be interpreted in two ways: at an aggregate level, through increasing the representation of research participants in the research process as a whole, and at the level of individual research participants.

Participant and patient involvement

Boote et al. note that there are three key arguments supporting public involvement in healthcare.[115] There is an epistemological argument relating to the knowledge and experiential insights that patients and

[114] Biesecker B et al. Web Platform vs In-Person Genetic Counselor for Return of Carrier Results from Exome Sequencing A Randomized Clinical Trial. JAMA Intern Med. 2018;178(3):338-346. doi:10.1001/jamainternmed.2017.8049

[115] Boote J et al Public involvement at the design stage of primary health research: A narrative review of case examples. Health Policy 95 (2010) 10-23.

members of the public can offer, which claims that those closer to direct experience and interpretation are best placed to offer accurate accounts of knowledge. This argument is encapsulated by the phrase 'nothing about us without us'.[116] A second argument is rooted in the publicly funded nature of the NHS, claiming that the public should be actively involved by virtue of its publicly funded status. A third, consequentialist argument states that public involvement allows potential improvement of the quality, relevance and impact of health research.

Patients with rare diseases often have a unique role to play in research. For some very rare conditions, especially those which are undiagnosed, patients contribute knowledge of the lived experience. Through their subjective experience of dealing with the challenges posed by illness, they may be able to identify key research questions; ensure that patient facing materials are relevant and accessible; contribute feedback on the practicality of what is being proposed in a research study, and be in a key position to assess the clinical utility of a new drug, treatment or management. In some jurisdictions (such as the US), they may also be able to actively influence whether a drug is prescribed to them via their health care professional. Indeed, groups of patients through health information sharing websites such as PatientsLikeMe have been mobilised into separate communities, able to share aggregate data and de-identified data[117] to collaborate with research partners in research but also to facilitate clinical trial involvement. This 'for profit' company now has the objective of redefining healthcare around the patient experience but has also facilitated its own research projects and research publications.

The importance of facilitating an active patient voice in research is gaining increased momentum in the UK. The National Institute for Health Research (NIHR) actively mandates public involvement (namely research carried out 'with' or 'by' members of the public rather than 'to', 'about' or 'for' them) in all research that they fund, providing a training and a 'matching' service[118] to connect life sciences with patients. It has also published a set of integrated, practical UK Standards for Public Involvement which are designed to improve the quality and consistency of public involvement in research through a practical toolkit comprising six elements: inclusive opportunities, working together, support and learning, governance, communication and impact.[119]

[116] Werner D. Nothing about us without us: developing innovative technologies for, by and with disabled persons. HealthWrights; 1998.

[117] https://support.patientslikeme.com/hc/en-us/articles/201245770-Does-PatientsLikeMe-sell-my-information-

[118] NIHR Patient Engagement in Clinical Development Service Patient Information Sheet for Life Sciences Organisations. Version: 1.0 - June 2019. Available from https://www.nihr.ac.uk/documents/patient-engagement-in-clinical-development-service-information-sheet-for-life-sciences-organisations/11659

[119] The National Institute for Health Research (NIHR), Chief Scientist Office (CSO) Scotland, Health and Care Research Wales, and the Public Health Agency Northern Ireland. UK Standards for Public Involvement in Research. NIHR November 2019. Available from https://sites.google.com/nihr.ac.uk/pi-standards/standards?authuser=0

This commitment to engage patients in research is also reflected in the research approval process. The Health Research Authority has entrenched this element within the research review process by having a question on the integrated research application system (IRAS) that allows patients, carers, service users and members of the public to offer insights on their health condition or experiences to help researchers with the design and set-up of their studies patient voice in research. INVOLVE is funded by the National Institute for Health Research to support active public involvement in the NHS, public health and social care research, and has a mandate for advancing it as 'an essential part of the process by which research is identified, prioritised, designed, conducted and disseminated'.[120]

However, despite public involvement being mandated at many levels of the research infrastructure landscape, numerous barriers and obstacles prevent these aspirations from being fully realised. Boote et al note in the primary care context that potential challenges include a clear understanding of health research methods, provision of sufficient funding at the outset of the research process and using accessible terminology to facilitate meaningful engagement. Active patient/public involvement is also sometimes rejected by commercial research sponsors on the grounds of commercial sensitivity of the drug/device being developed.

Decisions by individuals to share data

An alternative route for facilitating patient participation in research is through individuals making choices about their active participation in research by selecting how they wish to share their data. From the perspective of individual participants, facilitating individuals to express their preferences for research, and providing a mechanism for priority setting and research design, could be seen as a way of facilitating individual autonomy interests. One route for doing this is through providing increasingly granular forms of consent, whereby potential research participants can set out their preferences for research involvement. Providing a dynamic and granular form of consent enables participants to make flexible and responsive choices about their involvement in research which can take account of their changing circumstances and priorities.[121]

There has been a proliferation of platforms allowing individuals to share their data. Sometimes this is done informally via patient support groups or social media. Here, challenges include a lack of privacy awareness with individuals oversharing personal data that could be then utilised for secondary use without their knowledge or consent. Indiscriminate sharing of data, without understanding its potential sensitivity or longer term significance for the individual concerned, such as potential stigmatisation or discrimination,

[120] https://www.invo.org.uk/

[121] Kaye J, Curren L, Anderson N, Edwards K, et al. From patients to partners: Participant-centric initiatives in biomedical research. *Nature Reviews Genetics*. 2012;13(5):371-376.

might be a potential cause for concern, particularly for parents sharing their child's data to peers, even if that sharing is well-intentioned.

The MVP mentions an 'impact award' to incentivise participation on the platform and encourage sharing on social media. Providing incentives for sharing potentially sensitive data on social media even within a peer group setting could be inappropriate, for example, if they provide information about the future health of a child.[122] In order to mitigate against potential harms, participants could be reminded of the potential implications of posting on social media.

Operationalising participant choice

By permitting participants to express their preferences for the use of their data, certain types of research might be excluded, or made more onerous. Personalised data sharing has now become commercialised with commercial companies such as Private Access[123] offering consumers direct control over their entire medical record, with the ability to make this data available to researchers of their choice. In research carried out by Ipsos MORI, 25% of survey respondents stated that they would rather research did not happen if commercial organisations had access to the data and 17% of people said they would not want commercial organisations to have access to health data for research under any circumstances. In focus group deliberations of case studies, sharing of genetic sequencing data was considered to be the most risky example: 'genetic data both most private, and most potentially valuable'.[124]

Some of these observations concerning public attitudes to research involving commercial entities and research seem highly relevant. The platform potentially provides a mechanism for patients and participants to indicate their research interests and provide feedback at different points in the research pipeline. Research participants could potentially use the platform to contact researchers to identify research questions that deserve investigation, co-develop research proposals and materials, and for peer review by participants. There is evidence to suggest that patients are willing to improve research design in order to overcome challenges in recruitment and retention, especially if it can facilitate the development of new medicines.[125] PPI is already well integrated into the management and governance of GEL. The Participant Panel acts as an advisory body to the Genomics England Board, 'working to ensure that the health data held

[122] This guidance recommends that predictive genetic tests should normally be delayed until a young person can decide for themselves when or whether to be tested. British Society for Human Genetics (BSHG). Genetic Testing of Children. 2010.

[123] http://www.privateaccess.com/

[124] Ipsos MORI. The One-Way Mirror: Public attitudes to commercial access to health data. Wellcome Trust. Journal contribution published online 20.11.2017. Available from https://doi.org/10.6084/m9.figshare.5616448.v1

[125] DasMahaptra P. et al. Clinical trials from the patient perspective: survey in an online patient community. BMC Health Service Research 2017; 17; 166.

by Genomics England is looked after with respect and used in the best interests of the participants'.[126] The panel includes representatives from the 100,000 Genomes Project, GenOMICC COVID-19 study and NHS patients, and its members sit on the key governance committees of Genomics England (Access Review Committee, the Ethics Advisory Committee and the GeCIP board). Determining which research can be accessed via the platform will initially be determined via existing GEL governance requiring that research projects have prior approval from the Access Review Board,[127] incorporating a degree of PPI oversight.
At an individual level, the platform provides an infrastructure for patients and participants to assert their own views. The Ipsos MORI research was informative in that it also identified five different mindsets which shape the way the general public thinks about commercial access to health data. Researchers found that two mindsets were more prevalent among patient groups: a cohort actively supportive of health research and data sharing [the 'Monitor Me' cohort], for whom health data sharing offered life-saving impact, notwithstanding risks of identification or discrimination; and a deeply suspicious attitude towards commercial organisations in another group of rare disease patients and many healthcare professionals [the 'Fed Up' mindset]. These mindsets had a strong influence on individual approaches to weighing up the potential risks and benefits associated with research.[128] If the findings from this research are borne out, this suggests that platform users may not be representative of the entire patient/participant cohort, as there may be a minority who make a conscious choice not to engage with it.

> Consideration: There should be ongoing evaluation of the demographic characteristics of those who register for, and use the platform. Where information collected via the platform is used to inform wider research or policy, reasonable efforts should be made to identify the views of those people who do not use the platform.

3.3 Research catalogue

Some of the most promising applications for a participant platform centre on the ability to connect participants with new research. This could complement the aims of the NGRL to create new opportunities for research—including clinical trials referral[129] — and in turn enrich the existing dataset. For example, the

[126] https://www.genomicsengland.co.uk/participant-panel/
[127] Genomics England. The National Genomic Research Library protocol v5.1. 2020. Page 98/113. Available from https://www.genomicsengland.co.uk/national-genomic-research-library/
[128] Ipsos MORI. The One-Way Mirror: Public attitudes to commercial access to health data. Wellcome Trust. Journal contribution published online 20.11.2017. Page 83/161. Available from https://doi.org/10.6084/m9.figshare.5616448.v1
[129] Genomics England. The National Genomic Research Library protocol v5.1. 2020. Section 5.3.2. Available from https://www.genomicsengland.co.uk/national-genomic-research-library/

Research Environment currently enables (commercial) Discovery Forum members to generate hypotheses and select cohorts, and, providing research is approved, recruit subjects and perform a precision clinical trial[130] but the Protocol also recognises the potential further need for strategic partnerships to facilitate patient recruitment and trial facilitation.[131] If this is combined with the potential to develop a more detailed understanding of each participant's phenotype through the onboarding process or via further longitudinal data collection (as discussed in section 3.6 below), this function could also facilitate the identification of highly tailored cohorts.

Connecting participants with research can be enabled in two ways. First, and most straightforwardly, a research catalogue can be provided, enabling participants to browse studies and apply for those they are interested in. At minimum this is proposed to incorporate a way of highlighting potential relevance to participants. Second, and as a logical extension of determining potential relevance, an active 'research matchmaker' function may enable researchers to request that potentially suitable participants are actively notified of their research and asked to consider taking part.

At present, identifying suitable participants for further research in the GEL dataset requires relevant approvals and access to GEL data, combined with the ability to query that data for eligibility and then an application to recontact participants in line with the relevant policies for their cohort. Providing a new and dynamic way of asking existing research participants directly if they would like to contribute to further health research via the platform would mark a step change over the existing process that researchers must currently follow, but may require careful consideration to ensure appropriate legal and ethical safeguards are in place.

Invitation to join the platform

Currently, the participants who join this developing platform may come from three main cohorts: the 100,000 Genomes Project cohort, the GenOMICC study of people with severe infections and injuries (notably COVID-19) and a new route for patients who have been tested by the NHS Genomic Medicine Service and agree to contribute to further research. As discussed in section 2.1, the existing information provided to participants and the consent process they have been through have created expectations about the research they are contributing to. For all these cohorts, this includes information about, and a consent choice to, potential recontact for new research opportunities.

[130] Ibid. 5.3.
[131] Ibid. 5.3.3.

One complication for these groups joining the platform is that this is not necessarily the same as asking them directly if they would like to take part in further research but instead, if they would like to join a platform which, among other functions, enables further research recontact. This possibility has not been directly anticipated in the information provided to participants although some (such as the GMS Patient Information on research[132]) set out that participants may be contacted for a range of purposes, including 'if there is general news about the National Genomic Research Library'. Arguably, the research catalogue or matchmaker functions will bring contact to join the platform within the general expectations of participants, namely that they may be contacted about participation in research. However, given the range of additional potential platform functions, it may be advisable to expand the information provided to participants.

Consideration: It may be advisable to expand the information provided to participants to explicitly include potential future contact for new platforms or services supported by GEL.

There are additional differences between the cohorts in the information and expectations that have been provided about *how* participants may be recontacted for further research opportunities. For the 100,000 Genomes Project cohort and the NHS GMS groups, participants were informed that the clinical or GEL teams may contact them in future to ask them to take part in further research (or to ask their views on the 100,000 Genomes Project).[133] Participants in the GenOMICC study are similarly informed that they may be contacted again about 'other research opportunities'[134] but this information is less clear about who may contact them. The co-sponsors of the study are the University of Edinburgh and NHS Lothian and the participant information suggests that the study recruiters will be the ones to contact participants in the future.

In terms of legal requirements for the use of personal or private data, recontact by the GEL team to ask participants if they would like to join research should be well within the scope of the existing legitimate interests legal basis relied on by GEL as a research organisation and within the broad reasonable expectations of participants about the use of their private and confidential information as all have been informed about sharing of data with GEL.

The challenge is simply whether this specific recontact is expected and whether, for reasons of non-maleficence and trustworthiness in particular, recontact about the platform should be led by the

[132] GMS Patient Information Research v1.1.
[133] e.g. The 100,000 Genomes Project Rare Diseases Participant Information Sheet. Version 2.3. 01-01-2017 and the Patient Information Research v1.1 for GMS.
[134] GenOMICC healthy volunteer participant information sheet v2.4, 27th July 2020.

GenOMICC team for those study participants. Because it is likely that additional cohorts will be recruited into the NGRL in future and participants from those cohorts might be invited to join and use the platform, this issue is potentially of broader relevance. Potential participants should be made aware that they might be contacted directly by GEL team members about future research opportunities, including use of platform for this purpose.

> Consideration: Although all relevant participants have been informed that they will be recontacted about future research opportunities, the GenOMICC study cohort may not be expecting direct contact from GEL team members. This may require further exploration to ensure no surprises for this and other future cohorts if contacted about the platform.

Follow up information and consent

A research catalogue enables participants to signal their interest in a specific research project and, where relevant, this allows for a research team to follow-up with them and seek their consent. There is the potential for some of this follow-up communication to be conducted via the platform but equally, this could signpost participants to a further study website and initiate separate communication outside the platform between researchers and potential participants. Because these processes are yet to be determined we do not propose or recommend that certain approaches should be followed or not followed, other than to emphasise that these should be tailored to the type of research, level of invasiveness and risk, and the complexity of the information that will need to be communicated to participants. This should be a proportionate approach, commensurate with the balance of risks and benefits, as emphasised by the HRA.[135]

For example, it is possible to anticipate simple questionnaire or survey-based research with very minimal risk of harm to participants being included in the catalogue. In such cases it may be appropriate to provide information and seek consent online via the platform, following relevant best practice.[136] However, for more invasive research with greater potential risks- including informational risks to privacy or data protection rights- personalised conversations and information may be required. This is obviously the case

[135] HRA Guidance on applying a proportionate approach to the process of seeking consent v1.02 FINAL https://s3.eu-west2.amazonaws.com/www.hra.nhs.uk/media/documents/Proportionate_approach_to_seeking_consent_HRA_Guidance.pdf

[136] HRA Consent and Participant Information Sheet Preparation Guidance http://www.hra-decisiontools.org.uk/consent/ HRA Guidance: Applying a proportionate approach to the process of seeking consent (2017) https://www.hra.nhs.uk/planning-and-improving-research/best-practice/informing-participantsand-seeking-consent/MHRA etc, Joint Statement on Seeking Consent by Electronic Methods v1.2 September 2018

where further samples may be desired and this will require compliance with the legal and ethical processes for clinical trials or clinical trials of investigational medicinal products.

If consent is relied on to provide researchers with their legal basis for processing data (and explicit consent to processing of special category data) under the UK GDPR, the information provision and processes should follow relevant guidance, including from the Information Commissioner's Office.[137] If, as the HRA recommends,[138] consent is not the chosen lawful basis, this should also be explained to participants.

Finally, it will be important to ensure that a participant has capacity to consent to further research. This will be more challenging if this takes place entirely online, therefore, individual telephone calls or follow up may be advisable even in low-risk research. It is also important to note that digital and online processes for research participation and engagement may not work well for all groups. However, our analysis starts from the position that a platform is being put in place and that those who have engaged with it are therefore not those who will find digital or online communication or processes more challenging, or who are even unable to engage with them completely.

> Consideration: There should be a proportionate approach to the follow-up consent process and information provided in accordance with the balance of risks and benefits of each research project.

What research is included in the catalogue and how is it governed?

A second set of considerations arise when determining what research should be included on the platform for participants to browse. It is important to distinguish between research that is an extension of existing GEL facilitated research—i.e. research that is conducted in the Research Environment and subject to all the requirements and policies set out by Genomics England— and 'new' research that may be largely unrelated to current GEL research. The platform may enable both forms of research, working as an additional digital tool for GEL approved researchers, or, providing a connection between platform participants and entirely new research projects.

'Extended' GEL research

[137] https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/consent/what-is-valid-consent/

[138] https://www.hra.nhs.uk/planning-and-improving-research/policies-standards-legislation/data-protection-and-information-governance/gdpr-guidance/what-law-says/consent-research/

If the platform is used to facilitate supplementary research by already approved researchers relying on existing consents and authorisation by the relevant ethics and scientific review committees, the primary obligation is to ensure that the research is in accordance with the information given to participants and relevant internal GEL policies and the overarching NGRL protocol. We consider how these policies and protocols may shape the nature of research that can be ethically and legally carried out using the platform further below, under the heading of 'longitudinal data collection'. If new forms of data collection are planned which would not sit comfortably within the existing GEL/NGRL framework, this may be better thought of as 'new' research.

'New' research projects

In the current stage of development, it is not yet clear if the platform will only serve to connect participants with research activity that is entirely in accordance with the NGRL framework, or, if it could allow participants to choose to apply to research that has not been subject to the same policies and approvals set out in the NGRL protocol. For example, the Protocol anticipates that clinical research organisations may engage with Genomics England to identify eligible patients (for example, see section 11.1 of the Protocol v 5.1) but it is also clear that no results or summary level data can be exported without review and approval in accordance with the established framework.[139] However, the platform could facilitate the connection of individual participants with other researchers on the basis of information the participants have provided when joining the platform (for example their phenotypic or genotypic self-reported information). This would potentially mean that external researchers would not need to apply to Genomics England for approval because they would not need to access or analyse any data held in the NGRL to recruit participants.

Notwithstanding the potential enhancements for scientific discovery that this could enable, there may need to be careful consideration of whether allowing 'new' and unrelated research on the platform would circumvent important aspects of the carefully considered policies and governance that GEL has put in place. For example, when commercial researchers and industry apply for access to data within the Research Environment, they will only be granted access subject to the considerations according to 9.2 of NGRL protocol:

- 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

[139] NGRL Protocol section 5.3.2 Genomics England. The National Genomic Research Library protocol v5.1. 2020. Section 5.3.2. Available from https://www.genomicsengland.co.uk/national-genomic-research-library/

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.
- Prioritisation of access according to resource availability
- Facilitation of high quality research

There are also a significant further range of information governance requirements that are required of research to ensure that patient confidentiality and protection remain a 'key cornerstone' of the Library.[140] These include legally binding data sharing agreements setting out acceptable uses and measures to safeguard patient confidentiality and data privacy. As the Protocol sets out, '[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.'[141] It is possible that the platform can be integrated with the Research Environment in a way that allows this secure approach for 'internal' GEL research (see further below in section 3.6 on longitudinal data collection) but if the platform connects to 'external' research it is not certain that an Research Environment equivalent will be in place to safeguard the data in that project.

It is also important to acknowledge that GEL research participants have had expectations set about the nature of the endeavour they are joining, including that they are part of an ethical and transparent programme based on informed consent. Individuals and families who are taking part in GEL research are partly acting altruistically to enable new scientific discovery and medical insights; even those who hope for individual results are aware that they may not be conclusive. They are also informed about the strict safeguards (in particular the restricted Research Environment) in place around who has access to their data and how their privacy is protected. Although it may be feasible for researchers to seek consent to quite different research (e.g. open access data or even non-health research) hosted on the platform there are potential dangers that participants assume that the same governance arrangements and strict review procedures are in place for such research as are applied to the research that takes place within the Research Environment.

Careful consideration may be required to ensure that the trust placed in GEL by participants is maintained and that careful developed policies are not circumvented through signposting to further research via the platform. If there are differences between the research allowed within the GEL framework and research hosted on the platform (e.g. because it allows the transmission of personal data outside the GEL controlled

[140] Ibid, section 8.1.
[141] Ibid.

Research Environment) transparency about this is crucial so that individuals and families can make an informed choice to take part.

> Consideration: Where research hosted on the platform differs from the governance arrangements and safeguards in place for GEL research this should be highlighted to participants.

As a starting point in considering the appropriate governance of new or external research on the platform are the well-established fundamental ethical obligations for biomedical research including—amongst others—to ensure that proposals are scientifically valid, adequately resourced and that they pursue an objective which outweighs any risks and burdens to the research subjects.[142] It is not appropriate to rehearse the full range of legal and ethical obligations here but it is worth noting that Research Ethics Committees will consider these aspects in their review so ensuring REC approval is a crucial part of the governance framework for research hosted on the platform. We discuss the boundary between activities that may be covered by the existing GEL ethical approval and those which may fall outside that (and may require independent ethical approval) further below in section 3.6.

Determining suitability for participation

A research catalogue could simply provide information about different research projects and a mechanism for those browsing this information to apply to take part. This has potential dangers in terms of participants hoping that they or their family are eligible for potentially therapeutic research when in fact they are not, or alternatively those who are eligible may not realise they are. To maximise the potential of a research catalogue and to assist users in determining eligibility, the platform could provide an indication of their suitability for the project when browsing the catalogue. This is a likely aspect of the proposed platform,[143] but it also raises additional issues which we consider in relation to a function of 'research matchmaker', in particular in relation to the accuracy of matching and the potential for matching to lead to unwanted information for participants.

3.4 Research matchmaker

A research matchmaker builds on the research catalogue with a more active function enabling researchers to actively seek new recruits by setting parameters for eligibility and trigger a message, notification or

[142] WMA, Declaration of Helsinki. WMA Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Subjects, Adopted by the 18th WMA General Assembly, Helsinki, Finland, June 1964 and amended by the: 64th WMA General Assembly, Fortaleza, Brazil, October 2013

[143] See section 1.1. on the Minimum Viable Product.

other signal to participants who have chosen to be open to new research opportunities on the platform. Building on the considerations outlined in relation to a research catalogue, the matchmaker function raises a number of additional issues.

How are potential participants matched?

It is important that participants are accurately matched to potential research studies. If matching is not sufficiently accurate then it risks imposing a significant burden on participants and a consequent loss of trust or confidence in the platform and its research matching. This is particularly important for research that may require volunteers to travel a long distance or provide a sample. It could be important for this reason not to rely on self-reported data (for example, health information provided when joining the platform) as part of matching unless there is a way of verifying its accuracy.

If participants are provided with probabilistic indicators about how well they match to a research project these will also need to be as accurate as possible and may need explanation[144] so that individuals can determine if they would like to put their effort into applying to take part. How participants perceive and respond to these probabilistic indicators should also be kept under review to ensure that they are neither failing to capture the attention of relevant participants nor leading to disappointment for participants who are not in fact eligible for research.

> Consideration: A proportionate approach should be taken to ensure the accuracy of matching so that burdensome research or recruitment processes are not embarked on unnecessarily. This will depend on the nature of the research. Surveys and online research are less of a risk than invasive research which requires considerable effort from the participant.

Matching by risk and unwanted information

Where participants are matched to research according to their phenotype/genotype, in particular in disease specific research, there is a potential danger that they will be unaware of the 'risk' which makes them suited to a study. This is particularly the case where the study is unrelated to the condition or disease which has led them to have their genome sequenced and it gives rise to the challenge of respecting a potential 'right not to know' as discussed in section 3.1 above. A range of measures could be taken to safeguard a right not to know or minimise potential invasion of informational privacy as far as possible in the context of research matching. One could be to limit the research opportunities participants are matched with to the disease/conditions they are already aware of (i.e. the primary condition and reason

[144] e.g. 'close match' or 'highly likely to be a match'.

they have undergone genome sequencing in the first place). Alternatively, participants could be informed about the range of potential new information that participants could discover about them or their family, and be asked to choose or reject matching for research topics which are not related to their 'primary' condition. It could also be possible to try and provide only minimal information about new research opportunities if they relate to a risk that participants may not be aware of, and allow participants to decide whether they would like to learn more about a study (and their potential risks).

> Consideration: If participants are recruited according to a 'risk' that they may be unaware of, careful consideration should be given to policies or approaches that could be adopted to minimise a potential breach of a right not to know.

One further aspect which could benefit from participants setting their preferences is the number of active matches that are put to them for consideration to avoid overburdening participants or risking a drop in engagement or even withdrawal. If this is not the case, it could be appropriate to set a policy limit on the number of opportunities that are presented to participants in a given period (e.g. no more than four a year).

3.5 Data and reports

A function of the digital platform could be to enable the reporting of results based on their genetic data to participants via the platform. These could be classed as (a) non-medical reports (e.g. ancestry or ethnicity related information) or, (b) diagnostic/medical reports, added by a medical professional, with the ability to selectively share and discuss these reports with peers affected by the same conditions.[145]

These categories (medical and non-medical) are frequently treated differently in ethical and legal guidance; medical reports are likely to be subject to stricter controls such as additional validation or verification, or not returned within the research setting at all, or certainly not without access to clinical expertise. Whilst in some instances the boundaries may be clear cut, there may be others instances where they are not, and even if reports are 'medical' in nature it may not always be clear whether disclosing them will be beneficial or conversely cause harm. Additionally, seemingly innocuous results may in combination lead to health inferences.[146]

[145] In discussion with the project partners, the ability to share and discuss reports with peers was seen as being more remote due to the additional issues it raises around confidentiality and curation, and will not be discussed in detail in this report.

[146] The potential for incidental findings to arise when information from multiple datasets is combined is discussed further in section 3.6.

At present, these applications are not part of the participant platform, but they raise a number of ethical, legal and practical considerations which should be considered to assess whether, and how, they could be integrated in the future, including the type of information included, whether appropriate consent had been sought from participants, and how the results of the report are delivered and communicated.

3.5.1 Non-medical reports

The feedback of non-medical results such as ancestry information or eye colour can be engaging and interesting to some participants. In general, there is limited risk of harm to participants from this information but there are some considerations, relating to the context of the platform and the nature of Genomics England's research endeavour, which could influence whether it is appropriate to incorporate this form of feedback on the platform. The feedback of non-medical results is often associated with commercial direct-to-consumer genetic test services. In contrast, given that GEL is wholly owned by the UK government, and the 100,000 Genomes Project has been largely publicly funded, generating and communicating these reports could be regarded as an inappropriate use of scarce resources, or may undermine the spirit and purpose of the research endeavour, giving the impression of a commercial relationship rather than one of a public institution working to protect current and future generations.[147]

Given the aims and nature of the project, feedback of non-medical reports may require careful consideration including consultation with participants and empirical research to explore how this function is received by these cohorts. A precautionary approach may also be required to the return of non-medical results that could later generate health implications (perhaps as a result of an evolving understanding of variants, or combined with other sources of data).

> Consideration: There should be further exploration of the potential implications of reporting non-medical information on the platform with stakeholders, including participants and wider publics.

3.5.2 Medical reports

The feedback of medical reports containing extra findings of potential medical significance to participants is another promising application that may be introduced in the future. This could include findings from current GEL research or from research and other analysis enabled by the platform. This presents an exciting opportunity to further 'bridge the gap' between research and healthcare by making research reports more readily available. Establishing appropriate boundaries for the severity and actionability of the information

[147] Interestingly, so long as the return of medical results was carried out ethically, in accordance with appropriate consent mechanisms, and supported, where appropriate by relevant clinical expertise, the return of non-medical results was perceived by some interviewees to be the greater concern.

fed back and ensuring that these are communicated effectively and in a manner that supports participants will be essential for the legal and ethical incorporation of medical reports into the platform.

Disclosure of results

The return of a defined group of medical results to participants is already facilitated by GEL and guidance and policies enforcing strict parameters have been developed. Participants of the 100,000 Genomes Project and the NHS GMS cohorts are fed back results that may be relevant to the explanation, diagnosis or treatment of their disease and are clearly actionable.[148] In addition to pertinent (or primary) findings, participants can also *choose* a small number of additional or secondary results to be looked for and fed back. Management of secondary findings is challenging, as they are likely to be unexpected and the significance for each individual is often uncertain. These have therefore been limited, with 'clinical actionability' in mind, to include:

- 'Additional findings' that are unrelated to the cancer or rare disease that led them to take part. These are limited to a select panel of known pathogenic mutations with high clinical relevance[149]. This panel is more restrictive than the panel of 59 genes used by the American College of Medical Genetics.[150]
- Incidental findings, which are only fed back when there is an exceptional reason for doing so. This is in line with European Society of Human Genetics guidance.[151]
- Carrier status in the CFTR gene, as long as both members of a couple are taking part in the project and both request that it is looked for. This may be expanded to include carrier status for other recessive conditions in the future.

Currently, no results are returned by GEL beyond these strict parameters and reports feeding back a broader range of medical results via the platform would mark a departure from the GEL policies and NGRL protocol. Recommendations in the literature vary regarding the return of results to research participants. However, most commentators agree that results that *may* be offered to participants should be analytically valid; reveal an established and substantial risk of likely health importance, reproductive importance, or personal utility; may or may not be clinically actionable but may be valued by participants with return

[148] Genomics England. The National Genomic Research Library protocol v5.1. 2020. Section 6.2, p.46. Available from https://www.genomicsengland.co.uk/national-genomic-research-library/

[149] Ibid.

[150] Kalia SS, Adelman K, Bale SJ, 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. *Genetics in Medicine*. 2016; 19: 249–255.

[151] van El C, Cornel M. & Borry P et al. Whole-genome sequencing in health care. *European Journal of Human Genetics.* 2013; 21; 580–584. https://doi.org/10.1038/ejhg.2013.46

offering net benefit from their perspective; and only offered when return comports with law and the participant has been given a choice about whether to receive the result, or category of results.[152]

Prior to a decision to report health information via the platform, GEL, in partnership with relevant experts, the host and participant panel, should go through a process of scrutinising whether medical and non-medical reports should be returned, exploring how this information might be received, and considering what systems would need to be put in place to mitigate potential harms. The thresholds for medical and non-medical results should be carefully deliberated and more granular categorisation within these groups could be useful. For example, it may be appropriate to report the results of a medical survey which identified potential health implications (albeit with insufficient certainty to constitute a diagnosis) provided that participants are aware of the limited nature of the information and have a choice of whether they would like to know the results. However, this should be subject to careful consideration including how information may be used by participants and potential downstream impact on the healthcare system.

> Consideration: The return of results that have potential health implications should go through a process of approval and scrutiny by multidisciplinary committees which include patient representatives.

The delivery and communication of medical reports

The delivery and communication of medical results, especially genomic results, is an active area of discussion. In the context of the 100,000 Genome Project, and indeed more widely, medical results are returned to patients/participants by their clinical team, who tailor delivery in terms of the volume, complexity and tone of the information, to suit the individual before them and can respond to questions that arise. Commentators argue that this is particularly important in the context of genomic information, which is complex and susceptible to being misconstrued. Genomic information is also unique insofar as its implications can extend beyond the individual to other family members. Therefore, the involvement of the clinical team (which often include genetic counsellors) is crucial for the effective and ethical delivery of genomic results, helping the patient to understand the implications for themselves and others, and reducing the risks of uncertainty and confusion. It is thought that including genetic counsellors in this process in particular, attends to both the educational as well as emotional and psychosocial components of delivery of genomic information.[153]

[152] Wolf SM, Branum R, Koenig BA, et al. Returning a Research Participant's Genomic Results to Relatives: Analysis and Recommendations. *The Journal of Law, Medicine & Ethics*. 2015; 43(3): 440-463.

[153] Patch C & Middleton A. Genetic counselling in the era of genomic medicine. *British Medical Bulletin*. 2018; 126(1): 27–36.

Concern has been raised, however, about the clinical workforce capacity due to the limited number of geneticists and genetic counsellors and the much larger number of patients requiring genetic services in the future.[154] Increasingly, healthcare professionals are building their expertise to enable them to have conversations with patients about genomics results and evidence is being gathered as to the impacts and effectiveness of this strategy.[155] The feeding back of genetic results without offering genetic counselling or a conversation with a healthcare professional, as is often the case with direct-to-consumer genetic tests, has been heavily criticised[156] and tailored counselling and support is widely seen as a necessary part of the delivery of genomic information.

In the context of the participant platform, the delivery of results based on genetic data in the form of medical reports should abide by the same best practice standards, and be delivered by a professional with appropriate clinical expertise. The delivery of reports digitally via the online platform benefits from being less resource intensive, circumventing the bottlenecks caused by clinical workforce capacity. Remote contact with a healthcare professional may improve patient access to results and be more convenient for both parties. Indeed, alternative methods for reporting genomic results are being investigated by several groups using web-based informatics technology for patient education,[157] and use of telemedicine using video or telephone for counselling.[158] However, comfort with web-based communication will vary across participants, as will access to and ability to use digital methods; a barrier that equally applies to the use of the platform as a whole.

It is intended that medical reports will be 'added by a medical professional',[159] but the extent of the interaction between the participant and healthcare professional, and how this will occur (e.g. phone call, email via the platform, video call or a combination) is unclear. Although there is limited literature assessing the benefits and harms of communicating *genetic related* information online through different media, evidence suggests that digital communication can affect the patient clinician interaction both positively and negatively, encouraging openness in relation to embarrassing or sensitive issues, but sometimes at the cost of the relational connection that can be fostered during a face-to-face interaction. Some evidence suggests that the removal of the patient 'being seen' alleviates the feelings of embarrassment, social disapproval

[154] Hoskovec JM, Bennett RL, Carey ME, et al. Projecting the Supply and Demand for Certified Genetic Counselors: A Workforce Study. *Journal of Genetic Counselling*. 2017; 27: 16–20.
[155] Vassy JL, Davis JK, Kirby C, et al. How Primary Care Providers Talk to Patients about Genome Sequencing Results: Risk, Rationale, and Recommendation. *Journal of General Internal Medicine*. 2018; 33: 877–885.
[156] Hock KT, Christensen KD, Yashar BM, et al. Direct-to-consumer genetic testing: an assessment of genetic counselors' knowledge and beliefs. *Genetics in Medicine*. 2011; 13(4): 325-332.
[157] Biesecker LG, Lewis KL, Umstead KL, et al. Web Platform vs In-Person Genetic Counselor for Return of Carrier Results From Exome Sequencing: A Randomized Clinical Trial. *JAMA Intern. Med*. 2018; 178: 338–346.
[158] Voils CI, Venne, VL, Weidenbacher H et al. Comparison of Telephone and Televideo Modes for Delivery of Genetic Counseling: A Randomized Trial. *Journal of Genetic Counseling*. 2017; 27: 339–348.
[159] Sano participant portal progress update and MVP overview, shared with PHG Foundation in December 2020.

and stigma leading some people to be more likely to be self-disclosive and ask questions.[160] However it can also be viewed as 'impersonal'[161] and a review by Ong et al. suggested that face to face consultations were essential for communication about emotional states.[162]

> Consideration: Consideration should be given to how to facilitate the interactions between participants and healthcare professionals so that participants are sufficiently informed about the implications of their results. This may require different modes of communication tailored to the individual's personal preferences and the level of support that they need.

One of the most promising applications of online and web-based healthcare interaction is to supplement face to face interaction. Evidence suggests that when used in this way, digital tools can actually improve communication.[163] However, this must be caveated with the requirement that the results being returned should not pose a serious threat to health. A randomised clinical trial compared the return of carrier status results via a web-based platform versus a genetic counsellor. It showed that these results can be returned via a web-based platform, conveying relevant information with sufficient gains in knowledge and no evidence of adverse psychological well-being.[164] The patient group in this randomised controlled trial were well-educated, healthy, post-reproductive adults and although effective in this particular cohort, this mode of communication may not be appropriate for all populations and subsets of test results. This trial also noted that in person genetic counselling should be reserved for individuals receiving results that are a greater threat to their health. Thus, where the communication of information is likely to have a significant effect on patient care, there is broad agreement that a face-to-face encounter is necessary.

It is unclear whether the healthcare professional delivering the results will be part of the participants clinical care team or a private 'in house' clinician. The former may be preferable if participants are more likely to place trust in someone with whom they are familiar, but comes at the cost of placing additional demand on the NHS. An 'in house' medical professional would alleviate this concern, but generates other

[160] Huxley CJ, Atherton H, Watkins J, et al. Digital communication between clinician and patient and the impact on marginalised groups: a realist review in general practice. *British Journal of General Practice.* 2015; 65(641): e813-e821.

[161] Ignatowicz A, Slowther AM, Elder P, et al. Ethical implications of digital communication for the patient-clinician relationship: analysis of interviews with clinicians and young adults with long term conditions (the LYNC study). *BMC Medical Ethics*. 2018; 19(1): 11.

[162] Ong LM, de Haes JC, Hoos AM, et al. Doctor-patient communication: a review of the literature. *Social Science & Medicine*. 1995; 40(7): 903-918.

[163] Ignatowicz A, Slowther AM, Elder P, et al. Ethical implications of digital communication for the patient-clinician relationship: analysis of interviews with clinicians and young adults with long term conditions (the LYNC study). *BMC Medical Ethics*. 2018; 19(1): 11

[164] Biesecker LG, Lewis KL, Umstead KL, et al. Web Platform vs In-Person Genetic Counselor for Return of Carrier Results From Exome Sequencing: A Randomized Clinical Trial. *JAMA Intern. Med*. 2018; 178: 338–346.

considerations such as whether the absence of a personal healthcare relationship might affect the quality of counselling or assessment of the relevance of the result to an individual and their family.

> Consideration: Reports containing medical results (based on genetic data) that are clinically significant should be returned face to face by a health professional who has appropriate clinical expertise, or be accompanied by the offer of in person counselling.

Implications for family members

As has been raised previously, one of the unique implications of genetic results is that they hold implications for family members as well as the individual to whom they relate. Guidance on confidentiality from the General Medical Council (GMC) suggests that it is good practice to inform a patient about the relevance of information to their relatives and advise them to disclose risks to family. This is at the discretion of the patient/participant, although GMC guidance also suggests it may be justified to disclose directly to relatives, even against the patient's express wishes, in some extreme circumstances.[165,166] Where participants choose to do so, most genetic services offer "family letters" for patients to pass on to relatives, but these are also not always effective.[167] The platform therefore presents a novel mechanism through which healthcare professionals can directly contact family members who are also registered to the platform, potentially streamlining this process.

Consent, recontact and participant expectations

If this hypothetical application is implemented as part of the platform, the intention that medical (and non-medical) reports be returned to participants will need to be incorporated into the consent form, and the participant should have the opportunity to decide whether they would like to receive these results. Consent could either be sought via the platform, or, if it is foreseeable that harms might arise as a result of receiving unexpected or expected results in medical reports, more personalised conversations may be required.

[165] Communication of information given in confidence is generally permitted only if the patient consents or if each of three criteria are met: the patient refuses to inform others, an identifiable person (relative) is at serious risk of harm, and such harm might be prevented by disclosure

[166] General Medical Council. *Disclosures for the protection of patients and others.* Available from: https://www.gmc-uk.org/ethical-guidance/ethical-guidance-for-doctors/confidentiality/disclosures-for-the-protection-of-patients-and-others

[167] Dheensa S, Lucassen A, Fenwick A. Limitations and pitfalls of using family letters to communicate genetic risk: a qualitative study with patients and healthcare professionals. *Journal of Genetic Counselling*. 2018; 27(3): 689-701.

It is also important to note that asking participants to consent to medical reports will affect participant expectations about the platform. Individuals and families taking part in GEL research are primarily acting altruistically to help generate medical insights. Whilst desirable, a clinical diagnosis is not guaranteed and, for the 100,000 Genome Project participants, this is emphasised in the information sheet and patient consent form. Transparency around what participants can expect from medical reports should they choose to receive them and the possible benefits and harms of these additional insights will be essential. Facilitating realistic expectations is also important to avoid participants feeling that they are missing out on information pertinent to their condition or are in some way disadvantaged if they do not join the platform or opt into receiving data reports.

Consideration: Participants should be asked explicitly whether they wish to opt out of receiving non-medical and medical reports.

Recontact is a common challenge raised in the context of managing communications between clinicians/researchers and patients/participants. Concerns around recontact often arise due to (a) the discovery of new and potentially meaningful information that the participant is not expecting, (b) after a significant period of time has elapsed during which an individual's preferences may have changed dramatically, and (c) information being fed back by clinicians or researchers with whom the participant is unfamiliar.

The implications for participants of results returned in medical reports may change over time (as is the case with all results based on genetic data), however many of the traditional concerns highlighted above are alleviated in the context of the participant platform. The process of variant classification has been described as 'more of an art than a science' and variants of uncertain significance in particular are routinely reclassified as more data is gathered - they may be 'upgraded' to pathogenic or 'downgraded' to benign (the latter being more common).[168] When this occurs, laboratories generally issue amended reports to healthcare providers to, in turn, disclose to their patients. Participants of the platform will still need to be made aware when the meaning of their results evolves or changes but the nature of the platform is such that the channels of communication between participants and researchers remain open, enabling an ongoing relationship. Should the participant's preferences change, they can be relayed back to the platform, circumventing many of the challenges that are encountered in relation to recontact following participation in research.

[168] Hoffman-Andrews L. The known unknown: the challenges of genetic variants of uncertain significance in clinical practice. *Journal of Law and the Biosciences*. 2017; 4(3): 648–657.

3.6 Longitudinal data collection

Another exciting enhancement that a platform can bring to research is the ability to collect new data from individuals, engaging with participants in an ongoing manner and enhancing research datasets for new discoveries. Such longitudinal data collection has far-reaching potential to greatly enrich the data about each individual participant and create a powerful research resource by linking this to the data held by GEL.[169] Where this new data is made available to healthcare professionals and patients, it is also possible that it could be used to improve individual healthcare and provide an example of the learning healthcare system in practice.

Longitudinal data collection is already an aspect of the GEL framework as the research environment incorporates longitudinal data sources such as data held by NHS Digital or the Clinical Practice Research Datalink[170] which are brought into the research environment and updated at intervals.[171] Where the platform could add significant value is in the further collection of wider data, often referred to as 'Real World Evidence' or RWE.[172] This has been defined as '[h]ealthcare information derived from multiple sources outside of typical clinical research settings, including electronic medical records (EMRs), claims and billing data, product and disease registries, and data gathered by personal devices and health applications'.[173] The manner and form of data collection via the platform is an open-ended and potentially wide ranging landscape; from short and engaging 'pulse surveys'[174] to more extensive patient reported outcome measures (PROMS) such as medical questionnaires, symptom tracking surveys or even follow up interviews or other in person forms of data collection (perhaps alongside further sample collection). The rapidly developing technology landscape also creates an opportunity for data to be streamed and uploaded to the platform from wearables or other devices. We consider the additional issues such device integration may give rise to in the following section. In this section we discuss overarching issues of how such data

[169] Particularly in areas where the collected phenotypic information was not extensive: Kirkpatrick BE, Riggs E R, Azzariti DR et al. GenomeConnect: Matchmaking Between Patients, Clinical Laboratories, and Researchers to Improve Genomic Knowledge. *Human Mutation*. 2015. 36(10),); 974–978. https://doi.org/10.1002/humu.22838

[170] CPRD is jointly sponsored by the Medicines and Healthcare products Regulatory Agency and the National Institute for Health Research (NIHR), as part of the Department of Health and Social Care. Available from https://www.cprd.com/

[171] There are expectations to collect data from a range of NHS and other national sources set out in the NGRL protocol at section 8.5.7: Genomics England. The National Genomic Research Library protocol v5.1. 2020. Available from https://www.genomicsengland.co.uk/national-genomic-research-library/

[172] Real world evidence is one of the four themes of commercial research included in the NGRL protocol (at 5.2) and the import of additional datasets to enable further research is recognised as one of the potential routes to enrich the existing dataset and satisfy the NGLR commercial research aims.

[173] U.S. Food and Drug Administration. Use of Real-World Evidence to Support Regulatory Decision-Making for Medical Devices,": Guidance for Industry and Food and Drug Administration Staff. U.S. Food and Drug Administration. 2017. Available from https://www.fda.gov/media/99447/download

[174] Sano Participant portal progress update and MVP overview, shared with PHG Foundation, December 2020.

collection may be classified, how it may fit with GEL's existing ethical and governance framework, compliance with data protection and confidentiality and the potential challenge of 'incidental findings' from longitudinal data collection.

Is the data collection 'research'?

An initial consideration for some of these forms of data collection is how they should be categorised—as research or as some other form of data collection. While the genomics field is familiar with the boundary between clinical activities and scientific research (or the debate about the presence or appropriateness of such a boundary[175]) some forms of 'data collection' via the platform bring into question where the divide lies between research and other activities.

Short online surveys can be viewed mainly as a way of engaging users/participants and, depending on the questions and topic, be seen simply as an enjoyable activity or means of gathering information about the use of a website or preferences for the development of the platform.

The distinction between research and non-research activity is not entirely straightforward. The UK Policy Framework for Health and Social Care Research defines research as the: 'the attempt to derive generalisable or transferable new knowledge to answer or refine relevant questions with scientifically sound methods'[176] and, explicitly includes activities that are carried out in preparation for or both interventional and non-interventional research that aim to generate hypotheses, methodological research and descriptive research.[177] In the social sciences, consideration of ethical implications and review is also recommended for all research involving human participants.[178] In recent years there have been efforts to improve the consideration of research ethics in computer science and areas which may not have well developed training or emphasis on ethical frameworks for human subject research.[179] This includes developing good practice for internet-mediated research.[180] However, private as opposed to academic or

[175] Bertier G, Cambon-Thomsen A & Joly Y. Is It Research or Is It Clinical? Revisiting an Old Frontier through the Lens of next-Generation Sequencing Technologies. *European Journal of Medical Genetics*. 2018; 61(10): 634-641.

[176] At section 3.1. NHS Health Research Authority. UK Policy Framework for Health and Social Care Research. Last updated 10 March 2021. Available from https://www.hra.nhs.uk/planning-and-improving-research/policies-standards-legislation/uk-policy-framework-health-social-care-research/uk-policy-framework-health-and-social-care-research/#scope.

[177] Ibid.

[178] E.g. the Economic and Social Research Council's framework for research ethics. Available from https://esrc.ukri.org/funding/guidance-for-applicants/research-ethics/

[179] Editorial, Cambridge Analytica controversy must spur researchers to update data ethics. *Nature*. 2018; 555; 559-560.

[180] UK Research Integrity Office. Good practice in internet-mediated research. UK Research Integrity Office, December 2016. Available from http://ukrio.org/publications/guidance-notes/

scientific actors are not necessarily subject to the same policies and professional oversight,[181] and commercial online surveys are unlikely to be subject to independent scrutiny or require ethical approval.[182]

What does this mean for the platform?

Applying this analysis to longitudinal data collection via the platform, surveys which gather health information in order to derive generalisable new knowledge are highly likely to be considered research.[183] This applies whether or not the results are to be anonymised or maintained in identifiable form (although that impacts data protection requirements as discussed below). If surveys are purely for fun, with no aim to collect meaningful information this may not be within the scope of research. However, the intention to develop a database of responses which could be used or mined in future could bring the data collection within the scope of research according to some frameworks. For example, The Council for International Organizations of Medical Sciences' (CIOMS) 2016 International Ethical guidelines for health-related research involving humans consider the collection of data that may be mined for health-related research. They advise that, even if the health-related data are not collected deliberately (citing the examples of search engine queries or consumer choices on websites), the entities collecting the data should strive for 'governance structures and mechanisms to obtain authorization for future use of these data in research'.[184] One such mechanism for authorisation for future use in the UK context is the possibility of voluntarily applying for ethical approval for the creation and maintenance of a research database of individual-level information.[185] This could provide a 'generic ethical approval' for secondary researchers to undertake projects using the database (provided they have ethical approval and potentially further consent if requiring access to identifiable information or further contact with data subjects) within the terms of the approval.[186]

---

[181] Although there are some relevant codes such as the Market Research Society's code of conduct which sets ethical standards of behaviour for its members. The code is available from https://www.mrs.org.uk/standards/code-of-conduct

[182] Indeed, it may not be straightforward for private companies to obtain ethical approval for research outside the biomedical setting: Pater J. Discussion: Cambridge Analytica and Research Ethics. Cognitive Science at UMass Amherst (blog). April 25 2018. Available from https://blogs.umass.edu/cogsci/2018/04/25/discussion-cambridge-analytica-and-research-ethics/

[183] Simply checking the box for generalisable knowledge is sufficient to qualify as research requiring ethical approval using the HRANHS Health Research Authority's 'Is my study research?' decision tool: available from http://www.hra-decisiontools.org.uk/research/result7.html

[184] Council for International Organizations of Medical Sciences (CIOMS) & the World Health Organization (WHO). International ethical guidelines for health-related research involving humans. CIOMS & WHO. Geneva 2016. Guideline 12, p50. Available from https://cioms.ch/wp-content/uploads/2017/01/WEB-CIOMS-EthicalGuidelines.pdf

[185] See section 11 of the UK Health Departments Research Ethics Service' Standard Operating Procedures for Research Ethics Committees. Version 7.4 June 2019. Available from https://s3.euwest2.amazonaws.com/www.hra.nhs.uk/media/documents/RES_Standard_Operating_Procedures_Version_7.4_June_2019.pdf

[186] Ibid, sections 11.23-26.

Consideration: There may need to be careful thought about the use of surveys or questionnaires on the platform and whether these should be considered 'research' or a form of data collection that may require ethical approval.

Finally, even if some data are not considered to be collected for research purposes, ethical standards still apply to ensure transparency and fairness in the collection and processing of data, and to ensure that data are safeguarded against risks to privacy or unauthorised use.

Is longitudinal data collection within the scope of GEL's ethical and governance framework?

As we have already noted, the NGRL protocol sets out the ambition to link the library to longitudinal data sources (2.1) and that GEL expects to collect other health data on its consented patients into its data centre (8.5.7). The Protocol also anticipates 'patient entry' and 'other secondary data sources' as supplementing the 'refreshable identifiable clinical data' that acts as a central pillar for the Genomics England Clinical Interpretation Operational Plan (Figure, p37).

Moreover, participants are already informed that these external sources of data will be collected and the Genomic Medicine Service Patient Information about the research explains that the aim is for 'lifelong data' and that 'Genomics England is constantly working to identify new sources of health data to include that is important for research'.[187] When participants join the platform, a more specific consent can also be sought to link their longitudinal data collected via the platform with their other data contained in the Genomics England data centre.

Taken together, these factors would suggest that the collection of longitudinal data via a platform and its subsequent incorporation within the NGRL have been anticipated by the approved Protocol and ethical framework. It may be appropriate to provide more specific information about the platform as part of patient/participant information as and when they are revised but the development of a platform for this purpose is something that existing participants could reasonably expect within the information they have already been given.

However, as we mentioned in relation to research catalogue and matching there are a range of information governance and data protection considerations that may influence some of the practical arrangements for the collection, safeguarding and linkage of longitudinal data from the platform.

Data protection and confidentiality

As with data collected at the outset when participants join the platform, there will need to be due regard to ways that the privacy and confidentiality of the data collected from participants can be maintained. There also needs to be legal authorisation for the collection and processing of the data as well as compliance with

[187] GMS Patient Information research v1.1, p5.

the range of data protection obligations set out in the UK GDPR. We have already discussed the general confidentiality and data protection issues for data processing by the platform (section 2.4) but there are specific issues that will need to be considered in relation to any longitudinal data collection.

The common law of confidentiality is likely to require the participant's explicit consent for longitudinal data collection and sharing. There are some limited alternatives that could apply in specific situations. For example, the public interest could become relevant in situations where a disclosure is considered to protect others from significant risk of serious harm—e.g. if a genetic risk is identified in relation to genetic relatives and the participant refuses consent to disclose—[188] but this would only apply to exceptional circumstances and any disclosure should be as limited as possible and as minimally invasive of the individuals rights under Article 8 of the European Convention on Human Rights as possible. In particular, this would require that reasonable efforts are first made to secure consent to disclosure.[189] It is unlikely that an active legal duty to disclose confidential information to at risk relatives will develop to encompass such a platform in the near future. So far, the courts have only recognised a very limited healthcare professional duty to warn at risk individuals against the wishes of their patient where the at-risk individual is well known to the patient and in a relationship of close 'proximity' (circumstances as close to a doctor-patient relationship as possible without actually constituting such a relationship).[190]

Within the health service, confidential medical information is also shared on the basis of an implied consent to support the delivery of healthcare, including among healthcare professionals to support the patient's treatment.[191] However, in the platform context, even if data are to be shared for healthcare purposes it will be less clear that participants have granted their implied consent to the clinical use of their information in the platform context, as it is outside the physical or digital healthcare setting and less likely to fall within the 'reasonable expectations' of a participant.[192]

> Consideration: Explicit consent will be required for disclosure of the confidential information collected via longitudinal data collection.

Because longitudinal data collection may take many forms, there will need to be careful consideration of the appropriate legal bases for each form of data collection and processing carried out by the platform according to the UK GDPR. As discussed in section 2.4, there are a range of options within the law depending on the nature of the processing. Perhaps it most likely is that explicit consent will provide a

[188] *ABC v. St George's Healthcare NHS Trust & Ors* [2020] EWHC 455 (QB), [2020] 2 WLUK 400.
[189] General Medical Council. Confidentiality: Good practice in handling patient information. 2017, 67.
[190] *ABC v. St George's Healthcare NHS Trust & Ors* [2020] EWHC 455 (QB), [2020] 2 WLUK 400.
[191] General Medical Council. Confidentiality: Good practice in handling patient information. General Medical Council 2017. Page 13. Available from https://www.gmc-uk.org/ethical-guidance/ethical-guidance-for-doctors/confidentiality
[192] Taylor MJ & Wilson J. Reasonable Expectations of Privacy and Disclosure of Health Data. *Medical Law Review*. 2019; 27(3): 432-460.

lawful basis (under Article 6(1)(a) UK GDPR) for processing of personal data and satisfy the requirements for the collection and processing of health and other special category data (Article 9(2)(a)) but consent has some drawbacks in the data protection context. One of these is the poor fit between the specific consent required under the UK GDPR and more open-ended research purposes that may develop and would require re-consent. One benefit of seeking ethical approval as a research database (as discussed earlier) is that this may enable the use of Article 9(2)(j) scientific research purposes as a condition for processing. This provides a more flexible legal mechanism than consent if new categories of researcher or researchers may be envisaged.

Linking the data to Genomics England datasets

At present, Genomics England's policies and protocol mean there is a one-way valve for identifiable participant data, with data being brought into the trusted research environment periodically and linkage with the rest of the GEL dataset being made within this secure domain.[193] Once in the research environment, data are then de-identified to a vigilant standard against re-identification[194] and access is granted only to approved researchers (through GeCIPs, Discovery Forum and application to the Access Review Committee) to analyse the data within the secure environment. There is some very limited scope for export of data from the Research Environment for further analysis if this cannot reasonably be carried out within the Research Environment but this must be for specific work and data must then be deleted.[195] There is a presumption against export of any individual level data unless they are associated with very rare genetic variants for example, or if consent has been given to include such data in a Case Report.[196] While export of such data can be requested via an 'Airlock application', the Airlock Policy makes clear that it will be considered a breach of the airlock to use this process to reconstruct individual-level datasets outside the research environment.[197]

In practice, this implies that there should be no means for the platform host to be able to identify the subjects of their longitudinal data collection in the GEL datasets. The linkage between longitudinal data and the other GEL controlled data must be carried out by GEL, for example by cross referencing an NHS number provided by the participant with their internal codes. Given our assessment of longitudinal data collection within GEL's ethics and governance framework (above) it is also feasible that no further consent or additional legal basis will be required to incorporate some of the data collected via the platform in the trusted Research Environment. However, it would be good practice in line with the principles of fairness

[193] NGRL Protocol 8.5.7 Genomics England. The National Genomic Research Library protocol v5.1. 2020. Section 8.5.7. Available from https://www.genomicsengland.co.uk/national-genomic-research-library/
[194] Genomics England de-identification Policy version 1.0 (approved 20 Aug 2019).
[195] Genomics England. Airlock Policy, Version: 2.0 Approval Date 16 Jul 2020. Available from https://www.genomicsengland.co.uk/about-gecip/for-gecip-members/documents/
[196] Ibid, section 6.5.
[197] Ibid, section 6.6.

and transparency to ask for participants' consent to such linkages when consent is sought for the data collection by the platform.

Incidental findings from longitudinal data collection

As with much genomic research and discovery, there is the possibility that longitudinal data collection will identify new and unanticipated information with potentially serious implications for individual participants or their families. The management of genomic incidental findings (IFs) has been debated for over a decade[198] and for clinical sequencing it is frequently recommended that as targeted an approach as possible should be adopted to limit the potential discovery of IFs.[199] In genomic research, where it may be less feasible to limit the potential identification of IFs, the most authoritative UK guidance comes from the Medical Research Council and Wellcome Trust, who avoid a 'one-size-fits-all' approach and recommend that each research project should develop a clear policy on health related incidental findings.[200] In 'hybrid projects' which communicate results to participants, it may be more challenging to determine how likely incidental findings should be managed, for example if or when to disclose that an assumed father is likely not the biological father (known as misattributed paternity).[201] Genomics England's policy is not to return incidental or unsolicited findings that occur unexpectedly in the course of research as opposed to diagnostic analysis,[202] unless there is an exceptional reason for doing so.

Although longitudinal data collection could feasibly lead to an unanticipated genomic discovery (for example if phenotypic insight identifies an unanticipated genetic basis for disease in a participant) there is increased scope for such data collection to lead to wider findings that could be important for participant's health and safety. This is particularly the case if longitudinal data collection is enabled through devices and wearables to integrate a real-time digital footprint with biomedical data. This has been called 'deep phenotyping' and as the potential for such integration develops, there is an awareness that deep phenotyping may, or even inevitably will, lead to identification of incidental findings that may be of great

[198] Wolf SM, Lawrenz FP, Nelson CA et al. Managing incidental findings in human subjects research: analysis and recommendations.". *The Journal of* law, *Medicine & Ethics.* 2008; 36(2): 219-248.

[199] van El C, Cornel M. & Borry P et al. Whole-genome sequencing in health care. *European Journal of Human Genetics.* 2013; 21; 580–584. https://doi.org/10.1038/ejhg.2013.46

[200] Medical Research Council & Wellcome Trust, Framework on the feedback of health-related findings in research,. March 2014. Available from https://mrc.ukri.org/documents/pdf/mrc-wellcome-trust-framework-on-the-feedback-of-health-related-findings-in-researchpdf/

[201] Wright CF, Parker M & Lucassen AM. When genomic medicine reveals misattributed genetic relationships—the debate about disclosure revisited. *Genetic Medicine.* 2019; 21, 97–101. https://doi.org/10.1038/s41436-018-0023-7

[202] Genomics England. The National Genomic Research Library protocol v5.1. 2020. Section 6.2. Available from https://www.genomicsengland.co.uk/national-genomic-research-library/

importance to participants.[203] It has been highlighted that the combination of health surveys, GPS tracking and other sensor data in real time could enable researchers to identify events like inebriation while driving.[204] When deep phenotyping is combined with machine learning algorithms (a form of Artificial intelligence) it has been suggested that IFs in this context differ from those produced in next-generation sequencing as they are 'likely to be genuinely unexpected and novel in many cases'.[205] These and other features may mean that guidelines from other biomedical research fields are not a good 'fit' for IFs in deep phenotyping.[206]

Given the potential for identification of incidental findings from longitudinal data collection, the platform host or those responsible for the data collection and processing should ensure they have a policy in place covering this eventuality. These issues are yet to be fully addressed in the context of deep phenotyping but there are some starting points which can be drawn from the ethical and legal framework. If the data collection directly supports healthcare, then this may need to be dealt with within the clinical standard of care (acting in the best interests of the patient). However, for research data collection it is open to researchers to determine whether they should notify participants of some IFs or to exclude them altogether. These decisions will need to be made according to the specific nature of the data collection and the corresponding potential for IFs to be identified in the analysis of the data. Whichever policies are adopted, they should be clearly communicated to participants with a choice to opt in or out of such results where relevant.

> Consideration: Those responsible for collection and analysis of longitudinal data should develop a policy for handling potential incidental findings which is appropriate to the specific nature of the data collection and processing and communicated to participants with a choice about categories of incidental findings where relevant and appropriate.

3.7 Wearables and symptom tracking

As outlined above in section 3.6, the platform may facilitate the collection of longitudinal data to enrich the existing data about individual participants. In the future, a possible source of longitudinal data collection

[203] Kim A, Hsu M, Koire A et al. Incidental Findings in Deep Phenotyping Research: Legal and Ethical Considerations. Petrie-Flom Center 'Bill of Health'. 10th February 2021. Available from https://blog.petrieflom.law.harvard.edu/2021/02/10/incidental-findings-deep-phenotyping/

[204] ibid.

[205] Hallowell N, Parker M & Nellåker C. Big data phenotyping in rare diseases: some ethical issues. *Genetics in Medicine*. 2019; 21(2), 272–274. https://doi.org/10.1038/s41436-018-0067-8

[206] Kim A, Hsu M, Koire A et al. Incidental Findings in Deep Phenotyping Research: Legal and Ethical Considerations. Petrie-Flom Center 'Bill of Health'. 10th February 2021. Available from https://blog.petrieflom.law.harvard.edu/2021/02/10/incidental-findings-deep-phenotyping/

could be continuous monitoring technologies such as wearables devices and sensors. These devices can track a range of biomarkers across prolonged periods of time including blood pressure, oxygen levels, glucose levels, heart rate, sleep duration and quality, physical activity and other health-related factors. They are often connected to apps that allow for the self-reporting of additional symptoms such as diet and mood. The phenotypic, environmental and biological data generated can be combined with genetic data to help provide personalised real time insights into the individual patient. Their use could also extend beyond individuals to population health through providing deeper insights into disease trajectory and how human behaviour and lifestyle can contribute to the health of different disease groups.

Many of the considerations regarding longitudinal data collection raised in section 3.6, such as how this data collection should be categorised and the need to maintain privacy and confidentiality, are pertinent to the use of devices. Building on these considerations, the use of wearable technologies to collect longitudinal data may raise novel challenges around optimising their ability to generate meaningful insights, particularly if commercial companies hold proprietary interests over the data collected on the devices they sell.

Generating 'meaningful' insights

Data from wearables, symptom trackers and other devices can help to generate real time estimates of disease risk and progression, providing a more comprehensive and representative picture of individual health than the 'snapshot' collected intermittently in clinical settings. Integrating multiple sources of data to create multidimensional accounts of an individual's health is also referred to as the 'digital phenotype'.[207] This data rich picture may be useful across many areas of healthcare, and policymakers have expressed interest in its use for prevention strategies[208] and the management and treatment of disease.

The ability of these technologies to help generate meaningful insights depends to some extent upon their accuracy. The accuracy of some features found on direct to consumer (DTC) wearable devices has been contested; for example, a study by Shcherbina and colleagues found that of the seven popular devices they tested, none had an error below 20% for measuring a person's energy expenditure.[209] However, the underpinning technology in many devices is continually improving, and many health indicators, such as

[207] Jain SH, Powers BW, Hawkins JB, et al. The digital phenotype. *Nature Biotechnology*. 2015; 33(5): 462-3.

[208] The potential for data from wearable devices to be used in personalised prevention strategies was highlighted in the Chief Medical Officer for England's 2018 annual report 'Health 2040 – Better Health Within Reach'. For example in conjunction with social, economic, behavioural, biomedical and genomic data, to generate personalised and 'real-time' estimates of disease risk.

[209] Shcherbina A, Mattsson CM, Waggott D, et al. Accuracy in Wrist-Worn, Sensor-Based Measurements of Heart Rate and Energy Expenditure in a Diverse Cohort. *Journal of Personalized Medicine*. 2017; 7(2).

heart rate, steps and sleep duration, have been found to be reliable in popular devices.[210] Furthermore, as long as the results are consistent in their accuracy (i.e. even if they are consistently inaccurate), it is the trends in data and deviations from baseline recorded by these devices that can be most illuminating.[211]

In addition, advances in sensor and computing technology are enabling the development of more sophisticated DTC tools that blur the boundary between lifestyle and medical devices. For example, the latest versions of the Apple Watch (Series 4, Series 5 and Series 6) which have been approved by regulators in the US and UK are equipped with an electrical heart sensor that can take an electrocardiogram (ECG) and show the user their heart trace,[212] a test traditionally only available through the health system.

Not only must they be able to capture data reliably, but critically, participants must be willing to wear and engage with them over a sustained period in order for findings to be meaningful. Potential barriers may include unwillingness or inability to purchase and use wearables. Although participants using the platform are likely to be digitally literate and comfortable with using digital tools, not everyone will have the resources to access devices or may feel uncomfortable with passive continuous data collection, especially if it is automatically uploaded to the platform. Some conditions are also more amenable to regular monitoring than others. Where possible, those who face these challenges should not miss out on clinically relevant insights or on opportunities to improve their health as a result.

> Consideration: Consideration should be given as to how to minimise health inequalities that may arise from the use of wearables for longitudinal data collection. A crucial step will be to gain better understanding of how different factors impact the different levels of engagement with wearables.

Another barrier to the use of wearables amongst participants may be an aversion to collecting data 'for the sake of more data'. Not only is this burdensome (both for the individual who generates it, and for the parties who collect, clean and link it to other data) but excessive data collection may compromise its usefulness, and potentially generate 'noise' from which relevant data will have to be distilled. Instead, evidence suggests that for participants to be motivated to wear a device, they must see a purpose for it.[213]

---

[210] Xie J, Wen D, Liang L, et al. Evaluating the Validity of Current Mainstream Wearable Devices in Fitness Tracking Under Various Physical Activities: Comparative Study. *JMIR mHealth and uHealth*. 2018; 6(4): e94

[211] Cook S, Brigden T, Raza S, et al. Citizen generated data and health: predictive prevention of disease. PHG Foundation. 2020. Available from: https://www.phgfoundation.org/documents/cgd-predictive-prevention-of-disease.pdf

[212] This trace can be used to alert the user to potential abnormalities such as atrial fibrillation, a risk factor for stroke and heart-related disorders

[213] Keogh A, Dorn JF, Walsh L, et al. Comparing the Usability and Acceptability of Wearable Sensors Among Older Irish Adults in a Real-World Context: Observational Study. *JMIR mHealth and uHealth*. 2020; 8(4): e15704.

Using wearables for more targeted data capture over defined periods of time may have greater utility and encourage engagement and adherence.

It is also worth noting that there is a tension between using devices that are minimally disruptive and safeguarding continued participant consent and autonomy. Ubiquitous sensors that integrate easily into everyday living may be preferable as passive data collection requires far less involvement than manually entering symptoms or completion of regular questionnaires, but users might forget that their data is being collected, creating concerns surrounding the ongoing validity of consent.

Consideration: If wearables data are incorporated into the platform, thought should be given to where they could generate most value within different disease cohorts/groups. Targeted use of wearables and symptom trackers focusing on specific biomarkers may generate more meaningful insights and be more acceptable to participants.

Consideration: It should be clear to participants what data is being collected, and for what purpose.

Realising the potential of these data

The commercial nature of many wearables and symptom trackers mean that they are developed and operated by a range of different commercial companies (such as Apple, FitBit) with a variety of business models. Some rely on monetisation of the data generated from these wearables (by selling to third parties); others use the data solely to improve their products or services. Commercial rights over these individual level data potentially place limits on Sano and GEL accessing these data as intended in the platform development plan. One solution could be for participants to exercise their data subject access rights (as discussed in section 2.4) to obtain their personal data and integrate it with the platform.

Even if these data can be accessed, ensuring interoperability of different data systems is highly problematic. Device developers often use distinct, proprietary and closed communication methods making it difficult for communication and data transfer to occur between devices and external systems and databases. Intellectual property rights protecting proprietary systems and the lack of system interoperability could make it difficult for Sano to pull continuous streams of data from different devices onto the platform, and ultimately into the NGRL and the individual patient profiles managed by GEL.

Finally, as discussed in section 3.6, the potential for unexpected or incidental findings arising from the integration of wearables (for example, sleep and activity tracking identifying symptoms of poor mental health) requires careful consideration prior to implementation on the platform.

Consideration: Addressing system interoperability, patient privacy, and the potential for data overload will be critical to the incorporation of data from wearables and symptom trackers into the platform.

4. Discussion

4.1 Wider contextual issues

If the expectations of the project team are met, this resource has the potential to facilitate safer, more effective research at a number of points in the research process. The platform may integrate seven different applications, each focusing on various aspects of the research landscape and pathway. At a population scale it could potentially streamline research processes through facilitating more targeted recruitment, potentially resulting in safer, more effective interventions and fewer adverse events. This could result in less attrition during the research process and higher rates of adherence and compliance during research. More personalised and timely communications from researchers might also result in a more motivated and engaged patient population. Ultimately the use of the platform could help to facilitate improved population health as these benefits are reflected in advances in diagnostics, therapeutics and patient management.

From an individual participant's perspective, the use of a platform could enable personalised choices to join the platform and to participate in different functions according to individual values and preferences. The platform will also be physically accessible for many participants since the interface for the platform will be hosted via an app on the participant's computer. However, if the development of the platform proceeds as planned, incorporating all seven applications which have been identified in participant engagement activities, there could be points of friction with internal and external ethical, legal and regulatory frameworks. These issues will need to be resolved if the platform is going to be developed and implemented in ways that satisfy existing ethical and legal/regulatory requirements.

4.2 Key findings from our analysis and potential mitigations

In previous sections of this report, we have described our evaluation of cross-cutting themes impacting on the development and use of the platform (section 2); and the iterative accumulation of ethical and legal/regulatory challenges that could potentially be generated as various applications within the platform are rolled out (section 3). In order to make these findings more concrete, and to be able to frame possible mitigations, we have distilled an overarching issue from each topic in these sections. This 'distillation process' draws on our assessment of the significance and the likelihood of these issues arising.

For each issue, we have suggested a potential mitigation or topic for further exploration. These are not exhaustive or conclusive, but point to a potential action or area of policy that could be explored further by

the project team. In practice, the ethical and legal/regulatory challenges that are encountered will be heavily dependent on the granular details for implementation, and, in part, on the wider policy context which might prevail at the time the platform is rolled out.

For ease of reference, the key composite issues and their commensurate mitigations are summarised in the form of a table below.

Table 2. Summary of potential ethical and legal issues arising and potential mitigations

Area/application	Potential issue/challenge arising	Potential mitigation	Section number in report
Transparency	Clarity about the involvement of all parties involved in the development and operation of the platform.	Maintaining transparency through comprehensive and accurate branding on patient facing materials.	Section 2.1
Inequity and inequality	Ensuring that the platform operates in ways that minimise potential inequality and inequity, and that the existence of the platform does not, in itself, aggravate existing inequalities.	Supporting as many people as possible to be able to use the platform and actively engaging with those who cannot or choose not to as part of wider research and policy development.	Section 2.2
Consent/capacity	Ensuring individual choices about how their data are used and enabling meaningful participation by those who lack capacity.	Develop systems for active and personalised communication with participants and to identify individuals where capacity might be lacking (potentially by using automated methods), with appropriate support to facilitate their participation where possible.	Section 2.3
Privacy, data protection and confidentiality	Maintaining robust de-identification of data, establishing clear approaches to legal bases and developing procedures to fulfil data subject rights.	To transparently set out any differences in how data are used or safeguarded by the platform compared to existing GEL approaches and to develop procedures to facilitate data subject rights relating to data collected by the platform.	Section 2.4
My contribution	Manage inadvertent disclosure of unwanted health data.	Develop opt-out systems for communication around specified topics.	Section 3.1

Area/application	Potential issue/challenge arising	Potential mitigation	Section number in report
Patient voice	Actively listening to those who use the platform and those who don't.	As above, to promote active engagement with those who cannot or those who chose not to use the platform.	Section 3.2
Research catalogue	To ensure that all potential users of the platform have consistent and realistic expectations and being clear about any differences between studies on the platform and GEL approved research.	Align materials where possible for all current and future cohorts or participants in the NGRL and ensure transparent and comprehensible information about the governance of all studies hosted by the platform.	Section 3.3
Research matchmaker	To adopt a proportionate approach ensuring that individuals are only offered participation in research (including being a control) where the potential benefits outweigh the potential harms or, at a minimum, are commensurate with each other.	Be transparent about the potential risks and benefits of new research opportunities and ensure choices about any studies that may be based on risk derived from genotypic or phenotypic attributes that participants are not aware of.	Section 3.4
Data and reports	To ensure effective communication and sufficient clinical support for returning medical and non-medical findings to patients/participants, whilst minimising the potential for this to be a barrier to feedback.	Undertake research on novel ways of communicating medical/non-medical results of varying levels of clinical significance. This could support the evidence-based development of a multi-stakeholder system that balances the potential benefits and risks for feeding back medical/non-medical reports.	Section 3.5

Area/application	Potential issue/challenge arising	Potential mitigation	Section number in report
Longitudinal data collection	To build processes which promote participant participation and maximise the utility of information that is collected whilst maintaining the best interests of participants.	Evaluate the potential utility of these datasets in informing diagnosis, management and treatment in clinical care.	Section 3.6
Wearables and symptom tracking	Participants must be willing and able to wear and engage with these devices in order for meaningful findings to be produced	Undergo a process of co design in collaboration with participants to understand where wearables can generate the most value for different disease cohorts and how to reduce potential barriers to participant engagement.	Section 3.7

4.3 Additional overarching challenges

In addition to the specific ethical and legal/regulatory challenges addressed in this report which are relevant to the development and the implementation of the platform, there are some overarching challenges which seem likely to influence the operation and use of the platform. Deep phenotyping and data integration from the research data as well as from data collected from the platform could generate an unprecedented scale of potentially clinically significant findings requiring investigation. It will be important that sufficient resources are available to investigate and interpret these findings, to assess their potential clinical significance. Once a finding has been found to have clinical significance, robust and adequately resourced pathways and infrastructures need to be put in place to return those results to clinicians and to participants. As the volume of these findings increase, it is important that health care professionals have the requisite expertise to answer questions about the findings and their significance and to direct the participant into appropriate onward pathways for management and care.

As well as managing the return of clinically significant results, clinical services also have responsibility for undertaking capacity checks for existing 100,000 Genomes Project participants every five years. We have highlighted some of the uncertainties relating to these checks in section 2.3 of the report. Determining capacity may be highly dynamic for individuals within the 100,000 Genomes Project and other cohorts potentially using the platform. However, changes in capacity also apply to the cohort as a whole: those

participants who were recruited as children may now be capable of making their own decisions about their health; participants who entered the project as adults may have lost capacity through progressive illness or aging; others may encounter fluctuating capacity. Thus, managing changing capacity for contributors to the NGRL is likely to be a continuing ethical issue.

Developing appropriate policies and approaches to these challenges is complicated by the way in which a digital participant platform blurs boundaries between research, technology innovation and healthcare. While the blurring of research and healthcare has been much discussed, this project has additional complexity incorporating and reconciling any clashes between the norms and practices of digital innovation and those of scientific research or healthcare. As with the wider GEL research endeavour, the thorough deliberation and development process carried out by the partners in this project means that it could become an exemplar for hybrid practice in this new space.

One recurring theme encountered during this analysis is the applicable governance that should apply to communications via the platform. Given that researchers might increasingly use this for research matchmaking functions from which participants could infer individual risks, and for the return of clinically significant results, a further question is whether these researchers should be regarded as having a duty of care to participants and patients that is akin to the legal duty owed by clinical professionals to their patients. This question is not merely an ethical or legal/regulatory question, but also a normative issue, which will need to be resolved by a wider debate encompassing all the relevant stakeholder groups. The expectations of participants surrounding the use of the platform will be a key consideration. It is worth noting that evidence suggests that many participants have poor recollection of their consent choices and what these mean for them personally and indeed may not have read the material provided to them.[214] Generating empirical evidence on how the platform is used, the expectations of participants and other stakeholders, the perceived and actual benefits and harms experienced by users, and the emergence of applicable best practice advice will all be relevant to future policy development in this area.

In additional to these broad policy drivers, the ethical and legal/regulatory challenges will be influenced by the strategic decisions taken by the collaborating partners and other key stakeholders such as NHS England. These could include the extent of data integration between all aspects of the platform and the NGRL (e.g. establishing identity and security credentials; additional data collected during registration; and data collected through the operation of the platform in surveys, questionnaires etc.). To the extent that a separate database could be established which runs in parallel with the NGRL, it could be desirable to explore separate research ethics approval for those aspects of the platform, if that database could operate

[214] Dheensa S et al Fostering trust in healthcare. Participants' experiences, views and concerns about the 100,000 genomes project. EJHG 62(2019) 335-341

independently from the NGRL as a de facto biobank. The existence of an independent research ethics approval for the platform could provide additional reassurance about the governance of these aspects: however, it could also raise concerns about future divergence of these elements.

When translating these broader questions into future health policy, continuing to ensure that there are robust mechanisms in place for capturing patients' and publics' voices will be key.

Conclusions

The aim of the collaboration between Sano Genetics, Genomics England and Zetta Genomics is to develop technology enabling a participant engagement platform for population-scale genomics. The success in securing funding from Innovate UK as part of its Digital Health Technology Catalyst Round reflects the novel and ambitious objectives of this project. By enabling a robust and secure mechanism for data and information to flow between research participants and researchers this novel platform potentially offers unprecedented opportunities to enrich the datasets held within the NGRL and related datasets and in the process, build engagement from participants and patients.

Our report describes the plans for developing the platform, and examines these in terms of four cross-cutting themes: participant expectations and transparency, inequality and inequity in health, consent and capacity and privacy, data protection and confidentiality (section 2). In section 3 the report evaluates how incorporating various features within the platform might generate additional ethical and legal/regulatory challenges.

Given its ambitious objective and potentially wide-ranging impact, it is not surprising that our analysis has generated a comprehensive list of issues and challenges which could potentially be generated as the platform is rolled out. For many of these challenges, our analysis has led us to a 'consideration' to be borne in mind in future policy development. These considerations are highlighted in bold throughout the report. This list is not exhaustive as issues raised will depend heavily on the policy context that prevails at the time the platform is implemented. As described in the Discussion (section 4) for each set of challenges, we have suggested a composite challenge and possible mitigation.

None of these issues are unexpected; nor are they necessarily insurmountable. Together these mitigations suggest that three overarching themes could emerge for taking forward the development of the platform. The first is the need for *clarity* about the wider benefits, burdens and risks associated with the platform. In part, this flows from transparency about how the platform will operate, including the nature of the data that will flow between participants and researchers and vice-versa. But in order to make sense of these potential risks and benefits there also needs to be clarity about wider operational issues including the interface with clinical services, patient pathways and other forms of support.

A second theme suggests the need for increased *personalisation.* The ability to tailor the functionality and content of the platform according to the personal preferences of participants potentially minimises the likelihood of a lack of congruence between individual patient expectations and how the platform is used by researchers and administrators. Providing personalised approaches that take account of personal

preferences might safeguard privacy, data protection and confidentiality in ways that still allow participation in research whilst minimising the potential harms from planned or inadvertent disclosure of health data and research results.

A third and final theme highlights a commitment towards *engagement*: meaningful and sustained engagement is needed with potential participants in order to circumvent or minimise potential ethical and legal/regulatory challenges that might arise, for example through disenfranchising participants who could use the platform if sufficient support were provided. This includes engaging with those who may seem ineligible to use the platform on grounds of fluctuating capacity or a lack of capacity. It will also involve a continuing dialogue with participants as features of the platform evolve over time or as the external environment changes.

Together these three themes could provide a blueprint for realising the promise of this novel technology whilst minimising potential associated harms.

Appendix - Interviewees

The desk-based research underpinning our analysis was supplemented by six semi-structured interviews. We thank the following interviewees for their time, engagement, and expert insights:

Name	Job title	Organisation
David Birkinshaw	Senior Information Governance Manager	Genomics England
Jamie Ellingford	Health Education England Postdoctoral Research Fellow	The University of Manchester
Jillian Hastings Ward	Chair of the Participant Panel	Genomics England
Michael Parker	Chair of the Ethics Advisory Committee	Genomics England
Christine Patch	Clinical Lead for Genetic Counselling	Genomics England
William Spooner	Co-Founder	Zetta Genomics


phg
foundation
making science
work for health