Genomics Revolution

Chaired by Sir John Chisholm, former Chair of Genomics England and Professor Sir Mark Caulfield, Chief Executive for Barts Life Sciences

In partnership with PerkinElmer

State of the Nation 2021
About Public Policy Projects
Public Policy Projects (PPP) is a subscription-based organisation which offers practical policy analysis and development in infrastructure, digital technology, economy, health, care, and other public services. PPP is independent and cross-party, bringing together public and private-sector leaders, investors, policymakers and commentators with a common interest in the future of public policy. PPP offers policy analysis and a secure environment for the discussion of ideas.
Executive summary by the Chairs
Recommendations
1) Integrated advanced analytics and infrastructure
2) Pathogen surveillance and infection management
3) Prevention and detection
4) Cancer: early detection and surveillance
5) Functional genomics and therapeutic innovation
6) Ethics: the public and patient perspective
7) Global horizons in genomics
Acknowledgements

The content of this report is based on over 18 hours of discussion among 100 participants drawn from academia, industry, the NHS and wider civil society. We are grateful to each of them for sharing their time, knowledge and expertise. We would particularly like to thank our session Chairs (highlighted) for their contribution.
Public Policy Projects would like to thank and recognise the commitment of Sir John Chisholm and Professor Sir Mark Caulfield in co-chairing and developing this programme of work. Public Policy Projects would also like to thank the session Chairs for contributing their time and expertise.

In recognition of their support throughout this programme of work Anna Dickinson would like to thank the wider Public Policy Projects team. Additional acknowledgements of gratitude are extended to Eleanor Murray, Policy Analyst, Public Policy Projects; as well as former Policy Director Daniel Male and former Senior Policy Analyst Taome Jennings.

Public Policy Projects would like to thank Perkin Elmer and Roche, for providing overall sponsorship for this programme of work. Roche has had no involvement in the authoring or development of this report - Public Policy Projects has retained full editorial control throughout the programme of work.

This report was designed by Joe Everley, 19-ninety.com, and was subedited by Vicky Burman.
The Covid-19 pandemic has been the most disruptive event in global affairs for half a century. A virus for which there is no known cure has swept the planet and revealed to the world's population the fragility of health in today's globalised high-tech world. But science has come to the rescue. With astonishing speed, entirely new vaccines have been developed, the evolution and the transmission of the virus has been tracked and therapies have been repurposed. Underpinning all those developments is the science of genomics.

The UK has been able to play an important role because it is well placed in this science. Not only are UK research institutions rated among the best in the world, but the far-sighted investment over the past decade in programmes and infrastructure such as the 100,000 Genomes Project,1 UK Biobank,2 the National Institute for Health Research (NIHR) BioResource,3 Health Data Research UK (HDR UK)4 and the NIHR Clinical Research Network5 have opened the door to high-quality large-scale data resources. Furthermore, NHS England has instituted a first-in-the-world Genomic Medicine Service,6 incorporating seven state-of-the-art Genomic Laboratory Hubs7 directly connected to the point of care via matched Genomic Medicine Service Alliances,8 with similar initiatives underway in the devolved nations.9, 10 In England this service has been standardised through a dynamic National Genomic Test Directory that evolves the test portfolio on an annual basis,11 taking advantage of the rapid rate of discovery in genomics. Most important of all has been the public realisation, through the effort to combat Covid-19, that the NHS, academia and industry acting in a coordinated manner can marshal health data to enable scientists and clinicians to find accelerated insights that deliver enormous human benefit.

Executive summary
by the Chairs

The Covid-19 pandemic has been the most disruptive event in global affairs for half a century. A virus for which there is no known cure has swept the planet and revealed to the world's population the fragility of health in today's globalised high-tech world. But science has come to the rescue. With astonishing speed, entirely new vaccines have been developed, the evolution and the transmission of the virus has been tracked and therapies have been repurposed. Underpinning all those developments is the science of genomics.

The UK has been able to play an important role because it is well placed in this science. Not only are UK research institutions rated among the best in the world, but the far-sighted investment over the past decade in programmes and infrastructure such as the 100,000 Genomes Project,1 UK Biobank,2 the National Institute for Health Research (NIHR) BioResource,3 Health Data Research UK (HDR UK)4 and the NIHR Clinical Research Network5 have opened the door to high-quality large-scale data resources. Furthermore, NHS England has instituted a first-in-the-world Genomic Medicine Service,6 incorporating seven state-of-the-art Genomic Laboratory Hubs7 directly connected to the point of care via matched Genomic Medicine Service Alliances,8 with similar initiatives underway in the devolved nations.9, 10 In England this service has been standardised through a dynamic National Genomic Test Directory that evolves the test portfolio on an annual basis,11 taking advantage of the rapid rate of discovery in genomics. Most important of all has been the public realisation, through the effort to combat Covid-19, that the NHS, academia and industry acting in a coordinated manner can marshal health data to enable scientists and clinicians to find accelerated insights that deliver enormous human benefit.
This project was conducted at the same time as the Government’s recently published Life Sciences Vision, but focuses particularly on the opportunities afforded by genomics, building on and developing the strategy outlined in Genome UK: the future of healthcare. The work was segmented into seven key themes, and each was chaired by experts in that field.

1) Integrated advanced analytics and infrastructure

Unpicking the insights from genomics to deliver human benefit requires matching the phenome (clinical characteristics, multimodal imaging and digital pathology) to the genome on a colossal scale. Most countries in the world do not have the combination of the population-scale data, the skilled health infrastructure, and the deep scientific understanding required to undertake this. In principle, the UK has all three, but delivering on the potential across multiple lines of control while meeting understandable privacy concerns is challenging. That said, during the pandemic the UK has demonstrated an extraordinary and world-leading capability to pull together unique platform trials such as RECOVERY (Randomised Evaluation of COVID-19 Therapy), and large-scale sequencing of the virus through the COVID-19 Genomics UK Consortium (COG-UK).

The UK must seize the moment to realise its combined health and genomic data potential, capitalising on the latest technological advances such as long-read sequencing and providing a means by which researchers can apply their skills and artificial intelligence (AI) tools to the phenomenal data resources that exist. For maximum benefit this must stretch well beyond genomic data and into multimodal clinical data and diagnostics from all disciplines. Past and current investment offers a treasure trove of data assets in the UK, but to achieve maximum value there needs to be an infrastructure of federated connections to enable seamless analysis. Naturally, data linkage must be subject to access and privacy controls.

2) Pathogen surveillance and infection management

The pandemic has shown the importance of being able to quickly scale sequencing in response to a major outbreak while also responding locally. The UK was among the first countries in the world to develop a capability to track SARS-CoV-2 through its genomic signature and use that to shape policy and operational responses. This builds on existing and routinely used capabilities pre-pandemic that generated sequence data to detect and investigate foodborne-associated and other outbreaks (such as MERS coronavirus, Ebola virus and Zika virus), and predict resistance in pathogens such as Mycobacterium tuberculosis (the cause of TB) and HIV. We must build on existing capabilities to provide a high-quality infrastructure that integrates pathogen genomics, longitudinal clinical data, and host sequencing, including data from the severely ill. A combination of centralised and local sequencing is likely to be increasingly needed as technologies are developed that support local and near-patient sequencing, and the use of sequencing as a diagnostic test. Sequencing will remain important for the monitoring of SARS-CoV-2 variants, including variants of concern, but could bring benefit to other infectious challenges, including seasonal influenza and antimicrobial resistance (AMR). By building a national and global network for pathogen surveillance it will be possible to detect national and globally emerging infectious threats and act to reduce the impact of these on human health.

3) Prevention and detection

NHS policy rightly looks to pivot towards prevention. We believe that genomics has a key enabling role in that. Early hopes that the first sequencing of the human genome would reveal an abundance of specific genes linked to common conditions were frustrated by the phenomenal complexity of gene interactions.
With the development of data assets such as the 100,000 Genomes Project and UK Biobank, the UK is placing itself in the forefront of a global movement to routinely link an individual’s genomic data to their health data and make that available for the citizen to be aware of and progressively modify their life choices. The UK already has important programmes in this area (the Our Future Health genomic risk score pilot, diversity programmes and a newborn screening pilot). Screening in early life for disorders such as familial hypercholesterolaemia, where there is an intervention that can avoid disease or reduce the consequences, offers very important opportunities. Now is the time for health policy to encourage the emergence of an ecosystem that will identify the increasing insights from research for prevention and disease management linking major datasets and make them quickly available to citizens through a national genomic research library linking major datasets.

4) Cancer, early detection and surveillance

Cancer is a disease of disordered tumour genomes and can be triggered by inherited predisposition to malignancy. There is growing evidence that genomic data might be amenable to tailored patient care via precision treatment selection and is therefore especially suitable for the use of new data and new technology to deliver benefit. The role of whole genome sequencing, especially newer technologies and deeper ‘omic’ characterisation of tumours e.g. measuring RNA, and integrating this data with digital pathology and imaging could enable a comprehensive multimodal picture of cancer that helps optimise therapy and chance of cure. Furthermore, a benign consequence of the success of vaccine platforms in tackling the pandemic is the advance by many years of the proof in principle that detection of the DNA sequence of a key protein in the disease pathway can be rapidly translated into a vaccine aimed at stimulating a vigorous immune response. This offers yet another opportunity for the UK to exploit its advantages.

It has now been shown that an individual’s blood can pick up traces of mutated DNA (circulating tumour DNA, or ctDNA) from affected cells long before the disease has become visible by other means. A trial of GRAIL’s Galleri technology by the NHS is currently underway. This is a very promising direction, and the UK should now move to create an adaptable national ‘technology agnostic’ platform for evaluation and optimisation of the most promising ctDNA technologies across the course of the disease to inform the NHS about optimal technologies and sampling strategies. The UK is ideally placed to conduct an adaptive trial in which emergent technologies could be optimised and tested within a similar platform to that of the RECOVERY trial. This will be a step towards the NHS Long Term Plan ambition of dramatically improving cancer survival, partially by improving the proportion of cancers diagnosed early.

5) Functional genomics and therapeutic innovation

With the colossal data resource that will become available, the time is ripe for a significant initiative to bring together, in a collective effort, the research communities from academia, clinical practice and industry to unpick the complex paths by which an individual’s DNA gets translated into a phenotypic outcome. This is the understanding that will deliver therapeutic benefit. The UK has significant skill sets, infrastructure, and tools in the three communities that we recommend are drawn together to accelerate the translation of genomic knowledge into function and therapies for patients. A coordinated consortium of centres with a major UK funding stream, which could be stimulated by the new Advanced Research and Invention Agency (ARIA), could enable a step change...
in the biological understanding of rare and common diseases. This could create a strong standardised and compliant platform to fuel and accelerate therapeutic innovation in both conventional therapies and new advanced cell and gene-based therapies. The UK response to the pandemic demonstrated the enormous power of researchers, industry and regulators all working together with the NHS to deliver benefit to patients in record time. We believe this can and should become the norm.

6) Ethics, the public and patient perspective

Genomics data has always been regarded as particularly sensitive because it might have risk implications for the individual and their family. Projects such as the 100,000 Genomes Project\textsuperscript{23} consulted widely to adopt a rigorous ethics protocol. Building on that experience and recognising the enormous expansion of genomic testing as a consequence of the pandemic, the time is probably right to establish a national framework to enable experience to be shared and new developments to be given rapid and informed consideration. This report recommends such a structure. Although there is, as yet little evidence of deliberate misuse of genomic data, we suggest consideration be given to enacting legislation specific to genomic data to recognise the importance of public trust and the seriousness with which any attempted breach would be taken.

7) Global horizons in genomics

Across the globe, genomics is seen as a key to the future of health. The UK is recognised as a leader and is particularly respected for bringing its capability to bear to help combat the pandemic. Collaboration both at the transnational and the binational level will be important both for giving the UK access to more diverse data sets and to ensure we participate fully in the emergence of new global standards.

Conclusion

The UK is strong in genomics, and this is crucial to the future of health. We fully endorse the Government's Life Sciences Vision,\textsuperscript{24} and to realise that potential we need to seize the advances we made in combating the pandemic and embed them in routine processes in the NHS, the research community, regulators and industry. There remains a long way to go in discovery, translation, industrialisation and adoption to put in place the world-leading ecosystem that the UK is possibly uniquely able to become. The Government could have in the National Genomics Board\textsuperscript{25} the basis of the machinery to oversee the planning and execution to deliver that vision. Moreover, it now has the Advanced Research and Invention Agency (ARIA),\textsuperscript{26} for which this world-leading challenge will be an excellent fit.

This report supports the proposition that the UK has, within its borders and with its international connections, the raw materials to build the world's best genomic ecosystem, delivering both superior healthcare and economic performance. Our recommendations provide signposts as to how that could be achieved.
Recommendations

Integrated advanced analytics and infrastructure

1: The diversity of genomic data must be improved to avoid the risk of health data poverty, and by extension, health inequalities. This might require targeted community engagement, alongside patient information and education.

2: Building on public interest in health data, there should be an open and transparent conversation with the public about genomic data, clearly setting out the risks, but also explaining the benefits to individuals and wider society. This cannot be a one-size-fits-all approach, rather it should be tailored to individual groups, reflecting their specific concerns. Particular steps should be taken to include socio-economically deprived groups and the digitally excluded.

3: Genomic data should only be used when the intent is one of public good and improving health and wellbeing. If necessary, this should be written into legislation. Benefits from the use of healthcare data should accrue to healthcare organisations and patients. As the use of genomic screening increases, there must be investment in support and resources, such as genomic counsellors.

4: Functionality on the NHS App could be enhanced to allow people to consent to the use of their healthcare data for research or for participation in a clinical trial. It should be simpler and faster for patients to 'opt in' to be contacted about research.

5: The extensive potential of UK health and care data should be unlocked. The Government should set an ambitious target to integrate and annotate datasets, cohort studies and bioresources within the next two years, building on the UK Health Data Research Innovation Gateway to accelerate efficient data sharing and the highest calibre research.

6: As recommended by the Royal Society and HDR UK, there should be a standardised qualification for data access, equivalent to a ‘data driving licence’ that would demonstrate trustworthiness and ensure that qualified experts are authorised and authenticated to get rapid access to different data types within a standardised legal and ethical framework.

7: The Government should invest in a broad-ranging birth cohort, bringing together genomics and other molecular measurements with environmental and lifestyle factors.

Pathogen surveillance and infection management

8: The UK must build an end-to-end pathogen diagnostics and surveillance system for the next decade and beyond. This must be flexible enough to accommodate the likely growth in pathogen sequencing and balance the need for both centralised and near-patient testing with robust quality assurance, standardised reporting and a centralised data repository.

9: There should be a presumption towards pathogen genome data sharing for research. This should be done within a robust framework, based on risk, to enable the use of data for research, while respecting individuals’ views on data usage.

10: There should be sufficient surge capacity built into public health agencies, the NHS and academia in the event of another significant pandemic. This requires intelligent design and commissioning that should be regularly tested against a range of scenarios.

11: The UK Government should incentivise the development of sequencing technologies to ensure a vibrant genomic ecosystem attractive to investors willing to develop innovations for the benefit of the patient and the sequencing pathway.

12: The Government should invest in research and development to establish the role of metagenomic services in the UK, including impact on patient care and outcomes.

13: The Government should establish a genomics observatory service bringing together metagenomics, environmental health, animal health, public health and the NHS to track and identify potential outbreaks of infectious disease or AMR.

Prevention and detection

14: Opportunities for the use of genomic-based risk stratification for screening should be explored and integrated with existing screening programmes where necessary. This should be supported by a national genomics research library linking major datasets to identify the insights from research.
15: There should be robust investigation of the role of long- and short-read whole genome sequencing alongside multi-omics to understand how these can add biological insights and health gain. The potential utility of shallow whole genome sequencing (sWGS) and where it could have value merits exploration. Consideration should be given to how ctDNA and sWGS could be combined for those individuals at high risk of diseases caused by known genetic variations.

16: Polygenic risk scores could be routinely used in primary care to identify those patients at highest risk from cardiovascular conditions and other disease. This will necessitate professional engagement and involvement to establish decision support tools and mainstream this in the NHS.

17: There should be a pilot of whole genome sequencing (WGS) in newborns, conducted in line with the recommendations of the Genomics England and UK National Screening Committee public dialogue.

18: Circulating tumour DNA (ctDNA) has the potential to be used throughout the clinical pathway, including screening and diagnosis, as well as triaging patients and identifying relapse or recurrence. There should be a clear roadmap via a technology agnostic platform to wider testing and, where appropriate, adoption of ctDNA technologies. This could include further exploration of the benefits and risks, and the use of randomised controlled trials when necessary. Such a platform could enable accelerated evaluation and potential adoption onto the National Genomic Test Directory of proven ctDNA technologies to allow for greater use in clinical practice.

If these technologies are adopted, then early consideration should be given to standardising approaches by genomic tumour boards.

19: The future of cancer care is expected to include multimodal imaging, digital pathology, multiple biomarkers and genomic sequencing. This could include accelerated integrated pathology processes, the rapid application of multiple biomarkers including ctDNA, tumour proteomics, metabolomics and long-read tumour DNA sequencing with real-time analysis enabling earlier precision cancer care. A first-in-class global multimodal test platform could transform cancer detection and migrate later-stage presentations to earlier detection, which would make the UK an extremely attractive test bed for life sciences investment.

Functional genomics and therapeutic innovation

20: The ambition must be for a nimble and collaborative ecosystem with a joint ethos, infrastructure and goals. To do this, organisations, including academia, industry and the NHS, should combine skills and expertise and work together on large-scale projects harnessing their collective strengths. This will require the establishment of a collaborative network, bringing organisations together and encouraging an exchange of expertise through a portfolio of large-scale collaborative and interdisciplinary projects.

21: Every doctor should be trained in data analysis. They should be introduced to the basics of data analysis and coding as part of the medical undergraduate curriculum. A substantial subset of doctors should be able to undertake master’s degree level training in genomic and/or data science as applied to health, to ensure we mainstream the benefits of genomic healthcare.

22: Technology in the field of functional genomics is moving at great speed and it is challenging to predict where the field will be in five years. The UK needs to continue to invest in basic scientific research and technology development to enable engineering and testing of the role of human variants at scale.

23: The UK Government should make a major investment over the next five years in a platform to support more rapid progress from patients to gene identification to function to treatment, which would ultimately bring the benefits of genomics to every patient in the NHS.

24: The Government must invest in making the UK a world leader in multimodal data for health, including next generation digital pathology, bringing together tissue samples, genomics and imaging across all disease areas.

Ethics: the public and patient perspective

25: There should be a new human genetics commission, responsible for public and patient deliberation on the ethical development and application of genomic science by setting the framework for engagement and good governance and ensuring timely advice to government on policy. The first task of the new human genetics commission should be to convene a public deliberation that brings a wide range of the public, patients and experts together to articulate and validate a framework of principles relating to genomics and the use of genomic data that is consistent with the relevant human rights and addressed to anticipated applications and scientific advances.

26: There should be a genomics watchdog or ombudsman for the public to contact if they have concerns about the use of their genomic data.
Global horizons in genomics

27: Building on the work of the International HundredK+ Cohorts Consortium, the UK should promote the development of a global initiative to support the collection of a more diverse range of samples, supported by truly federated databases to inform scientific understanding and therapy development.

28: There should be large-scale international collaboration and greater investment in programmes to move from variant to function to therapy to patient. This initiative would focus on functional characterisation and understanding, and how that can be taken forward into functional genomics and therapeutic uses.

29: There should be greater investment in global genomic surveillance to identify emerging pathogens. To allow for rapid sharing of information about potential pandemic threats, there should be international best practice minimum standards for pathogen sequencing that all hospital, commercial and academic laboratories should be encouraged to adopt. There should also be international protocols to promote rapid sharing of data for surveillance and analysis.
The UK is fortunate to have some of the best genomic data, health service data and cohort data in the world. Few other countries have systems that combine high-quality data, consistency and national coverage on the same scale as the UK, or with the same potential to link patient data to enable patient-based analysis and follow-up.

Combatting the Covid-19 pandemic has hugely accelerated our capability to collect, link, access and use health data for planning and research. It has allowed the NHS and researchers to identify and protect millions of people at high risk from Covid-19, to deliver and monitor the safety and effectiveness of the Covid-19 vaccination programme, and to identify life-saving treatments for Covid-19, including dexamethasone.

These benefits must not stop with Covid-19. We can, and need, to go much further. Beyond the prevention of future pandemics, we must capitalise on this once-in-a-generation convergence of analytical technologies, big data management, scientific breakthroughs in disease prevention and the arrival of new tools to enable ‘citizen power’. With these new assets, we are reaching a tipping point where the opportunity to reset the agenda for personalised medicine and to make a historic step change in population health will define the next era of health and care for all. The opportunities afforded by these seminal innovations extend to people living with non-communicable diseases, (such as mental illnesses, cancer, heart disease and diabetes), to people with inherited or genetic conditions and rare diseases, and to people at high risk of a shortened lifespan.
The tipping point is real. Methods have recently been developed for combining genomic and health service datasets with other data sources such as the educational, environmental data and population surveys. We are beginning to understand more about the drivers of the social determinants of health, such as pollution, and their impact on disease prevalence. We are beginning to understand the power of ‘massive’ data to help guide the search for new drugs, design new NHS services (such as social prescribing), simplify complex patient pathways and identify people at risk.

We need to invest, as a country and as the NHS, in innovating and scaling new capabilities.

- New data linkages, and the prospect of linking genomic and other polygenic and molecular data to a patient record, offer the prospect of much earlier interventions to prolong life and deliver patient benefit. This will allow us to measure long-term outcomes in clinical trials, assess the safety of new medical interventions and understand patterns of health and illness across the whole population.
- Currently, health data is fragmented across a complex institutional landscape, stored in legacy silos and often hard to link or extract because our systems are not interoperable. We must invest in data mobility and data accessibility.
- Health and care data is of variable quality. We must invest in curation and standards to make it easily readable, easily manageable and fully longitudinal.
- Patient trust in how we use and disseminate data is paramount to progressing this new frontier of population health and wellbeing. We must invest in communicating with the public about their rational concerns and respond with new principles governing safe, ethical and timely use of data.
- There is therefore an opportunity to develop and maintain the integrity of a trusted data research infrastructure that makes a wide range of multidimensional data available, linkable and ‘FAIR’ (Findable, Accessible, Interoperable, Reusable). This will provide a platform for developing and deploying AI-enabled tools and technologies, including through developing at-scale genomics, imaging and citizen-generated wearable and sensor data.

To achieve this the UK needs to do the following:

- Provide responsible stewardship of a federated data infrastructure that enables open science through a trustworthy system of data governance and rules for data access by positioning ethics, human rights and public engagement at the heart of what we do.
- Power up the UK’s major research investments by enabling infrastructure and interoperability that supports the discovery, linkage and analysis of large-scale, diverse data that represents the whole population. This will enable Trusted Research Environments (TREs) to work together to build a robust cloud-enabled service to curate and scale up to multi-petascale analyses.
- Educate our citizens in the significance of this historic moment in population health and the benefits that can be safely delivered to transform our solutions for healthy ageing, living with chronic conditions and early prevention.
- Build and sustain the confidence of patients, healthcare professionals and policymakers in our ability to safely capitalise on innovations in data by adhering to the principles of full transparency, clear communication and ongoing engagement, particularly about how the data will be accessed and disseminated. It is vital that healthcare planning and research includes and represents the full spectrum of citizen perspectives so that we find treatments, improve care and deliver positive outcomes for everyone.
Genomics is one of the most exciting technologies of the 21st century. It has the potential to change the world, just as microelectronics did in the 20th century. During the Covid-19 pandemic, genomics and health data have been in the public eye more than ever before, but for this technology to achieve its potential it must have continued investment and, importantly, public support. Only then can genomic technology reach a point where it can offer life-changing benefits to the global population.

For genomics to achieve its potential there must be access to high quality health and care data and our national assets such as UK Biobank, the 100,000 Genomes Project and Our Future Health should be linked so that we realise maximal potential of these major investments. For maximum benefit this must stretch well beyond genomic data and into multimodal clinical and data and diagnostics from all disciplines. The UK has made some good progress in this area, with datasets that can be used as a basis for great science, but there remains much to be done. Genomic data needs to be combined with NHS data in a structured, curated and accessible format. Underpinning this there must be a strong governance framework built on public understanding and trust.

Access to curated data is not a challenge that is unique to the UK; there was a commitment to greater international health data sharing at the recent G7 summit in the UK. Many countries are building good datasets, but face difficulties over how that data can be used and shared, both within and between countries. This challenge is particularly acute in the field of genomics, as the current tools, privacy architecture and information governance (IG) frameworks were not built for sharing data around genotyping and phenotyping.

The UK and other countries need to use the pandemic as a catalyst to develop a framework and infrastructure that allow for the appropriate use and sharing of data for research and clinical practice in a way that maintains public support.

Patients and healthcare data

Since the start of the pandemic there has been a growth in digital technologies in all sectors. In the healthcare space that has been particularly notable; prior to the pandemic approximately 90 per cent of GP consultations were face to face, by July 2020 90 per cent were remote (primarily phone or video). Many of the assumed barriers about patients not being willing or able to use the technology seem to have been overcome.

Independent charity Future Care Capital recently undertook a survey on public attitudes to the use of data. It found that 27 per cent of respondents were more open to the idea of the NHS sharing some of their personal health information with third parties than they were prior to the pandemic, and 49 per cent said their views on the sharing of health data had stayed the same. Similar research from Savanta ComRes and MHP Health showed that 57 per cent of people said they were comfortable or very comfortable sharing their data for the purpose of developing new treatments or vaccines, and 33 per cent were more comfortable sharing data with the NHS since the start of the pandemic.

While this shift in public attitudes towards data sharing might be welcome, a separate survey found that the potential of data-driven technology was not being fully realised, with just 42 per cent of respondents saying that digital technology was making the pandemic situation in the UK better. As a result there might be an ‘opportunity gap’ between technology’s potential and the perceived reality of how it is being used. This creates an opportunity for greater use of data to help
catalyse health and economic benefits. However, it is important to be cognisant of the ‘data divide’. Evidence from the Ada Lovelace Institute has found that during the pandemic data-driven technologies have not been experienced equally, and the digital divide has shaped a ‘data divide’, based on access to technology as well as knowledge, awareness and skills. The report found that: “A large proportion of the public lacks awareness of the existence of, and the potential to use and adopt, some of the technologies [available].”

The data that is being collected needs to be representative of the whole population, and efforts must be made to actively reach out to under-represented groups to explain the benefits and risks and invite them to participate in research. Much of the current data has significant biases that need to be addressed, otherwise there will be consequences for health outcomes. This is particularly important as progress is made towards a greater use of AI in health, where any biases in data could be exacerbated, leading to ‘health data poverty’.

**Recommendation 1:** The diversity of genomic data must be improved to avoid the risk of health data poverty, and by extension, health inequalities. This might require targeted community engagement, alongside patient information and education.

One of the significant differences between standard healthcare data, such as a GP record, and genomic data is the predictive power of genomic data, and the insights it can provide into an individual and their relatives. Genomic data, such as that derived from whole genome sequencing, is very personal, and there are questions about whether it can be completely de-identified. Therefore, there might be a need for even higher standards of ethics, access considerations, use cases and privacy protection.

Another challenge of genomic data is that the answer is rarely binary, and as the science advances understanding and interpretation might change. Currently, the expectation from a patient is that a test is either positive or negative, but polygenic risk scores can lead to new uncertainties. This can be particularly problematic when the interpretation changes, or when the ‘risk’ turns out to have occurred, such as when an individual who was categorised as ‘low risk’ still develops the condition. It is therefore important to set clear expectations with the public about what can be learned from genomic information now, and what could be discovered in the future, as the science advances.

Individuals will also need to understand that for their own risk score to be developed it will need to be combined with those from other people to build the breadth of dataset needed to draw useful conclusions. This solidarity approach to healthcare relies on individuals recognising that by sharing their data it improves the overall dataset from which everyone can benefit. As part of this collectivist approach, it will be necessary to ensure that any risk stratification tool captures the broad range of the population, with a representative cross-section based on ethnicity, socio-economic background and other factors. A priority must be to minimise gaps and biases in the data and build datasets that are truly representative of society.

At the beginning of the Covid-19 pandemic, NHS Digital built a risk stratification platform. Based on the QCovid® model, the Covid-19 Clinical Risk Assessment Tool is designed for use by clinicians during a consultation with a patient or as an aid to support direct care. It can also be used by clinicians to review patients added to NHS Digital’s Shielded Patient List as a result of the Covid-19 Population Risk Assessment. It has initially focused on risks related to Covid-19 but could be expanded to other disease areas.
The pandemic has awakened a lot of interest in health data among the public, and the UK should capitalise on this opportunity to engage the public about genomic data, its current role and future potential. It is vital that any such public conversation considers how to reach those who are socio-economically deprived or digitally excluded. This conversation must also be underpinned by a strong ethical framework (see Section 6 on ethics). Any such campaign should clearly set out the risks and limitations of genomic data so that individuals can assess the risk themselves, based on clear explanations.

One approach could be to link the benefits to individuals and society to advances in healthcare treatments, to help patients understand that sharing their data could contribute to advances in treatments that could ultimately benefit them or their loved ones. An alternative might be to frame it as ‘what can patients do for the NHS?’ During the pandemic people have been willing to take individual actions, such as wearing masks and being vaccinated, for the common good. The NHS has broad public support, thus framing the sharing of patient data as helping the NHS might appeal to some individuals. Many of the concepts around genomics are intangible, and clear, real-life case studies demonstrating clinical utility and how patient data has contributed to advances in treatments could contribute to building support for data-sharing frameworks.

Whichever approach to public conversation is chosen, it must be underpinned by the ‘five safes’ (people, projects, settings, outputs and data), so that individuals are confident that their information will be protected and safeguarded in an appropriate manner.

Recommendation 2: Building on public interest in health data, there should be an open and transparent conversation with the public about genomic data, clearly setting out the risks, but also explaining the benefits to individuals and wider society. This cannot be a one-size-fits-all approach, rather it should be tailored to individual groups, reflecting their specific concerns. Particular steps should be taken to include socio-economically deprived groups and the digitally excluded.

However, if this hard-won public trust is to be maintained, then it is important that data is only being used where the intent is one of public good and improving health and wellbeing. Transparency on which organisations are using patients’ data and for what purposes is also essential for public understanding. Finally, it is important that patients and healthcare organisations are compensated for the use of this data, and benefits must accrue to those organisations that provided the data.

One significant challenge when it comes to phenotypic information is that patients do not always understand or act on the information they receive about their current or future health. There needs to be further exploration of the behavioural science around how to activate citizens once they are equipped with phenotypic information.

Individuals must also be supported and have access to appropriate resources when they are given information about their genetic profile and potential risks. One example is Lynch syndrome, where it is thought that about one in 280 people carry a pathogenic variant in a mismatch repair gene. This would equate to around 175,000 people in the UK. Despite national guidance, currently only around 6,000 patients have been identified by the NHS. Were all these individuals to be identified overnight there would be significant additional pressure on the NHS for screening and other services such as genetic counsellors. In 2019, there were approximately 310 genetic counsellors in the UK. As the amount of genomic screening available increases, this will need to be accompanied by significant additional resources in other services, such as genomic counsellors.
Recommendation 3: Genomic data should only be used when the intent is one of public good and improving health and wellbeing. If necessary, this should be written into legislation. Benefits from the use of healthcare data should accrue to healthcare organisations and patients. As the use of genomic screening increases, there must be investment in support and resources, such as genomic counsellors.

Evidence suggests that the pandemic has had an impact on the public’s attitude to research, with 27 per cent of respondents to an NIHR survey saying they were more likely to participate in research.46 This public enthusiasm should be harnessed through longitudinal research, as the FinnGen initiative is doing in Finland.47 One way to do this could be by offering everyone the opportunity to ‘opt in’ and be contacted about suitable research based on their health data. This could be done by enhancing the capabilities of the NHS App to engage citizens with the preventative health agenda.48 This improved NHS App could allow people to consent to data sharing, invite participation in clinical trials, or allow individuals to enter their own patient-reported data, which would be a valuable asset for research.
**Recommendation 4:** Functionality on the NHS App could be enhanced to allow people to consent to the use of their healthcare data for research or for participation in a clinical trial. It should be simpler and faster for patients to ‘opt in’ to be contacted about research.

**Data linkage**

Genomic data on its own has limited value. For real benefits to be recognised, genomic data needs to be linked to other datasets such as electronic patient records, and then curated, aggregated and analysed using machine learning and AI tools. Only when this process has been completed – which is costly and time consuming – can the data generate real value, both for individual patients and for wider health research.

There are increasing efforts to link genomic information and bio-tissue samples to ‘traditional’ datasets such as electronic patient records. This type of linkage affords the opportunity to follow up with patients over a life course through a cohort. Such data linkage would also allow for the identification of potential participants in trials based on specific genetic variations. However, too often this data linkage is not possible, and the information is fragmented, and the data captured in different formats.

For the greatest impact, organisations need to effectively share the data and ensure the metadata and the annotations that are associated with a dataset are coherent. AI only works when data is well annotated, and so it is necessary to annotate and share data in an efficient and effective manner. Alongside data annotation, steps need to be taken to standardise data tools, methodologies and ontologies. Once this has been done, sharing data and metadata is more straightforward. Important progress has been made with the development of the UK Health Data Research Innovation Gateway, with over 650 datasets and associated tools and resources available.\(^\text{49}\) This data aggregation and annotation not only benefits academic research, but it also makes the UK an attractive investment proposition for life sciences companies, which are interested in large, well-labelled datasets to support drug development programmes.

**Recommendation 5:** The extensive potential of UK health and care data should be unlocked. The Government should set an ambitious target to integrate and annotate datasets, cohort studies and bioresources within the next two years, building on the UK Health Data Research Innovation Gateway to accelerate efficient data sharing and the highest calibre research.

**Data governance**

In recent years, the UK has built a panoply of data architecture in the field of genomics, phenotyping, and polygenic risk scores. Yet the science is moving rapidly. The governance architecture around the use of data is overextended and has been regularly amended and updated. It is underpowered to deal with the challenges of the future, particularly relating to genomics and personalised medicine.

To date, the Trusted Research Environment (TRE) system has been effective in controlling access to datasets for use in research. However, as the amount of available data is likely to grow significantly in the coming years, and as there will be a need for linkage among a growing number of datasets, this infrastructure might no longer be fit for purpose. Alternative models of data sharing should be considered, to make access to the data fast and reliable and reduce the number of processes that researchers must complete to access data.
**Recommendation 6:** As recommended by the Royal Society\(^50\) and HDR UK, there should be a standardised qualification for data access, equivalent to a ‘data driving licence’ that would demonstrate trustworthiness and ensure that qualified experts are authorised and authenticated to get rapid access to different data types within a standardised legal and ethical framework.

**Government investment in genomics**

The UK Government has a track record of investment in genomic research and technologies. Much has been written about the potential of AI to transform our understanding of disease and its potential applications in genomics are widespread. There is significant investment in both genomics and AI by large-scale health and technology companies, but there should also be some public spending and some public–private partnership to ensure public good remains at the forefront.

One option for UK Government investment could be a birth cohort on the scale of UK Biobank.\(^54\) The data within UK Biobank is likely to be the foundation for lots of preventative medicine, but primarily for late-onset disease rather than life outcomes. The UK could explore the option of a large-scale birth cohort capturing DNA and genomics, but also proteomics, metabolomics, transcriptomics and other scalable molecular measurements. These should sit alongside information about the social determinants of health, including environmental factors. This would enable the identification in early life of children who could benefit from interventions that will improve their life outcomes. Such a scheme would be a unique global asset of measurable economic benefit to the country.

**Recommendation 7:** The Government should invest in a broad-ranging birth cohort, bringing together genomics and other molecular measurements with environmental and lifestyle factors.
Introduction by Professor Sharon Peacock, University of Cambridge, and Professor Paul Kellam, Imperial College London

The UK has a strong track record of translating basic science into policy and practice relating to pathogen genomics in diagnostic and public health microbiology. Prior to the pandemic, the UK’s public health agencies routinely sequenced numerous pathogens to support public health. These included Mycobacterium tuberculosis (the cause of TB) and pathogens associated with foodborne and healthcare-associated infection outbreaks. The UK has been a leading light in sequencing being used during the ongoing treatment of viral infections, including sequencing of HIV to detect variants encoding drug resistance and internationally during virus outbreaks such as those caused by MERS coronavirus, Ebola virus and Zika virus.

The Annual Report of the Chief Medical Officer 2016: Generation Genome described the benefits of sequencing M. tuberculosis and bloodborne viruses, and specifically described the importance of sequencing in epidemics and pandemics. Covid-19 has been a major catalyst for the further development and implementation of pathogen sequencing in the UK. A national SARS-CoV-2 sequencing network, the COVID-19 Genomics UK (COG-UK) Consortium, was formed in April 2020. This distributed network of numerous academic laboratories, the four public health agencies, and the Wellcome Sanger Institute have generated almost a million genomes at the time of writing.

COG-UK was developed as an emergency response, and a sustainable pathogen sequencing network is required for the future. Scotland, Wales and Northern Ireland have national sequencing capabilities in place (with further expansion planned), while NHS Test and Trace is leading the development of a network of
seven Public Health England laboratories (from 1 October 2021, UK Health Security Agency), which will offer a national pathogen sequencing capacity. This will provide quality-assured public health genome sequencing capabilities that will become embedded in existing national infrastructure.

We now need to think beyond the present day and consider the future innovations required over the next decade, and what impact these could have on how we deliver pathogen sequencing that provides actionable information.

Ongoing improvements in sequencing technologies could result in rapid, affordable sequencing that is performed closer to patients and healthcare settings, and lead to a shift in the current model of pathogen sequencing. Innovation in medical technology (‘MedTech’), particularly at, or close to, the point of care, will be vital for this change to happen, and private sector innovation and partnership with the NHS as a validation service could speed up adoption. But sequencing that moves closer to patients will require considerable research and development to establish the evidence for use cases, based on the benefits created for patients and public health. Any ambition to sequence primary patient samples (which might contain both the infecting organism plus organisms from the microbiota) will also require research into how metagenomics approaches can be used, and how they impact on patient care. The generation of genomic data linked to patient information and opportunities for increasingly advanced analytics could be the biggest potential win if planned and supported effectively. But important questions need to be asked, including how we make the most of what the NHS can provide to capture a national strategic advantage. Innovative design and co-creation will be required for automated analysis pipelines to reduce time to action, and accessibility to genome data by healthcare providers, public health specialists and researchers will maximise the value of data generated. We must drive innovation and translation so that distributed sequencing leads to even greater detection, prevention and control of infectious diseases, making every pathogen genome count towards better public health.

The UK has a strong track record for the application of pathogen genomics to matters of public health. In early 2013, three working groups (on rare diseases, cancer and infectious diseases) met and made recommendations on the priorities for the 100,000 Genomes Project in the UK. The infectious diseases group prioritised sequencing of Mycobacterium tuberculosis, Hepatitis C, and deep sequencing of HIV (to detect minor variants encoding drug resistance, which can emerge to become the dominant variant during treatment). The 2013 report also proposed the need for a devolved network of pathogen sequencing laboratories. The recommendations were adopted, and in 2017 Public Health England announced the launch of a national mycobacterial reference whole genome sequencing service and that it would offer HIV and hepatitis C sequencing through its antiviral unit. The UK was one of the first adopters of routine TB isolate sequencing, which is used to rapidly predict drug resistance and provides additional information on genetic relatedness to support outbreak investigations. Numerous other pathogens are now routinely sequenced to detect and investigate foodborne outbreaks.

In 2016, the Chief Medical Officer’s annual report, Generation Genome, articulated the many benefits of sequencing of TB and bloodborne viruses, and specifically described the importance of sequencing in epidemics and pandemics using several case studies, including foot and mouth disease and Ebola. This continued to raise the profile of pathogen sequencing for disease control. UK scientists and public health agencies were highly active during the
Ebola outbreak in West Africa, the MERS coronavirus outbreak in Middle Eastern countries and the Zika virus outbreak in the Americas, where sequencing played an important role in eradicating the disease.

Covid-19 has provided a major catalyst for the further development and implementation of pathogen sequencing in the UK. The COG-UK Consortium was instigated when there were fewer than 90 diagnosed cases of Covid-19 in the UK, with funding from the UK Government (through the Covid-19 ‘fighting fund’ set up by Sir Patrick Vallance and Professor Chris Whitty) and from the Wellcome Sanger Institute. COG-UK has become one of the most successful SARS-CoV-2 sequencing operations in the world. It has developed methods, tools, logistics and operations, networks and ways of working that are being used and followed globally. To date, the UK have generated almost a million SARS-CoV-2 genomes.

COG-UK’s rapid genome sequencing programme has been one of the big successes of the pandemic. As a result of the consortium’s foresight, the UK has, to date, sequenced one of the largest numbers of SARS-CoV-2 genomes worldwide. From a standing start, and by adopting an inclusive model, it has been able to scale quickly and collect and sequence large numbers of SARS-CoV-2 from people with Covid-19 in hospitals and the community to give vital insights into the pandemic. In particular, it has enabled identification of SARS-CoV-2 ‘variants of concern’ that can be monitored and tracked, contributing to their risk assessment. Achieving this has been a highly collaborative effort, with multiple providers contributing to the development of a consistent approach to SARS-CoV-2 sequencing.

The COG-UK operating model is one of distributed sequencing, followed by centralised data repository into the CLIMB-COVID computing infrastructure, with data access then provided to a range of organisations dependent upon need. This structure has been effective during the pandemic, but consideration should be given to the most appropriate model for future needs. When designing the most appropriate system for pathogen sequencing, it will be necessary to plan beyond SARS-CoV-2 and consider all potential future needs, including risks from a range of pathogens across bacteria, fungi and other viruses. Any system must also continue to enable ongoing surveillance for diseases such as influenza and TB, foodborne pathogens such as Salmonella and E. coli, and, increasingly, drug resistance in a range of pathogens. This is challenging, as system and sample flow needs will vary, and ethical considerations will vary dependent on the pathogen.

As the science continues to advance and more use cases are adopted for pathogen sequencing, there will also be increasing overlap between sequencing for public health surveillance and disease control and for diagnostic and patient-focused purposes. Use as a diagnostic would increase demand and expand datasets that are available for use for public health surveillance but will require strategic planning supported by funding to be prepared in advance for such changes. Furthermore, the speed of innovation in the field of genomic sequencing is likely to be rapid, and it is likely that the sequencing technologies of the future have yet to be developed.

Any new operational model will need to consider and balance the benefits and challenges of centralised and near-patient sequencing. The former enables efficiencies of scale and simpler data flows, with oversight and surveillance of infectious diseases such as TB and foodborne-associated pathogens. Near-patient testing allows for easier access by healthcare professionals and can facilitate rapid local outbreak detection and personalised medicine approaches to care. Over
the next decade it is anticipated that a blended model will emerge that combines central reference sequencing with smaller-scale local hospital and community-based sequencing, automated sample preparation and standardised outputs.

Moving sequencing closer to the patient will not be without difficulties. There are likely to be computational challenges because sequencing generates large quantities of data that need to be transferred, processed and stored. In many parts of the NHS, particularly away from large laboratory centres, data infrastructure and bandwidth are not sufficient to enable large-scale data transfer. To address this there must be investment in gigabyte networks across the NHS and access to appropriate computational infrastructure.

A second challenge of moving sequencing closer to the patient is training sufficient staff in sample collection and processing. There are already workforce challenges in academic and central NHS laboratories, with many of the staff working on SARS-CoV-2 sequencing having been redeployed from other research projects. Any move towards near-patient sequencing will require a significant increase in trained staff.

Thirdly, it is vital that all staff involved in pathogen sequencing follow the same quality assurance and protocols, but with near-patient sequencing it could be more challenging to retain high standards of quality assurance across significant numbers of sites.

Finally, issues remain around availability of automated tools that analyse the sequence data and create reports on the sequencing outputs. Currently, sequencing reports are not produced in a consistent manner, which can result in variations in interpretation and understanding among clinicians. If pathogen sequencing is to become more widespread, standardised reports need to be developed that can be understood by clinicians from a range of fields and specialisms.

Irrespective of the model that the UK adopts, it is vital that public health agencies are fully integrated into the systems and processes, and that the protocols allow for rapid sharing and analysis of data and information for surveillance and research.

Any future system must also allow for international and cross-functional data sharing. The pandemic has demonstrated that pathogens do not respect borders or species, and rapid international data sharing and collaboration are vital to monitor and control outbreaks of disease. However, for this international collaboration to be most effective there need to be standardised methods for conducting pathogen sequencing as well as the analysis and sharing of data (see Section 7 on global horizons). The UK Government has an important role in ensuring that scientists, laboratories, systems and structures are able to handle the next pandemic, irrespective of its origins, while simultaneously delivering on an expanding remit for pathogen sequencing.

**Recommendation 8:** The UK must build an end-to-end pathogen diagnostics and surveillance system for the next decade and beyond. This must be flexible enough to accommodate the likely growth in pathogen sequencing and balance the need for both centralised and near-patient testing with robust quality assurance, standardised reporting and a centralised data repository.

Sharing of health data is a complex topic, and it is important to be mindful of public opinion and appetite for data sharing (see Section 6 on ethics). There must be a clear framework around how data can be used, which is understood and supported by the public and balanced with the public good.
For the greatest benefits to be accrued from pathogen sequencing, there needs to be an integrated system that supports rapid linkage with epidemiological data. A minimum dataset will need to be defined that includes information on the patient who was sampled, and the time and place that the sample was obtained. During the current pandemic it is extremely beneficial to understand whether an infected patient has been vaccinated, or whether they have previously been infected with SARS-CoV-2. Linking pathogen information to an individual’s electronic health record and making that information available on an anonymised or pseudonymised basis to specific researchers through a Trusted Research Environment (TRE) or similar would also enhance scientific understanding and research into disease. This will require good data quality, flow, integration and aggregation, together with a robust system to ensure appropriate access for research that is in the best interest of patients (see Section 1 on integrated advanced analytics).

A TRE is not required for all types of data and research. Alongside this detailed repository there is a role for less detailed datasets with metadata (such as time and place of sampling) that is used for rapid research with public health impact. This has been done successfully with CLIMB-COVID, allowing for the low-risk use of time and location data to create a stream of genomic epidemiology for use by public health agencies. Data from CLIMB-COVID is routinely enhanced by public health agencies for more detailed analysis. The infrastructure and learning from CLIMB-COVID as a data analysis platform provides a point of reference for plans for a pathogen genomics infrastructure in UK public health.

**Recommendation 9:** There should be a presumption towards pathogen genome data sharing for research. This should be done within a robust framework, based on risk, to enable the use of data for research, while respecting individuals’ views on data usage.

Whatever pathogen diagnostic and surveillance system is established, this needs to be flexible to respond to new and existing threats. During the Covid-19 pandemic, many academic laboratories have been repurposed to support SARS-CoV-2 sequencing, reducing their capacity for sequencing for academic research. For example, during much of the pandemic, the Wellcome Sanger Institute sequencing capacity has been focused on Covid-19 and not on sequencing for UK Biobank.

Across the entire system, there was an increase in sequencing requirements related to Covid-19, but the commensurate reduction in other demands gave in-built surge capacity to the system. However, this might not be the case in any future outbreak of disease, and there needs to be sufficient surge capacity in academic, public health and NHS laboratories to cope in the event of serious outbreaks without significantly affecting routine requirements.

To ensure that the system can cope in the event of another pandemic, there should be rigorous modelling of a range of potential scenarios to ensure that public health agency, NHS and academic capacity is fit for purpose, without a significant impact on other requirements. It will also require strong and deep relationships between the NHS and academic institutions to enable cross-functional working and shared capacity as required.

**Recommendation 10:** There should be sufficient surge capacity built into public health agencies, the NHS and academia in the event of another significant pandemic. This requires intelligent design and commissioning that should be regularly tested against a range of scenarios.
The pandemic has provided a precedent for technologies such as PCR testing moving from the laboratory to widespread use, delivered by the public health agencies, NHS and private providers. The turnaround time and cost of sequencing is also reducing, and the technology to undertake sequencing is becoming more widely available. This also changes the contact points where samples for sequencing accumulate, including the ability to use residual samples from PCR-testing workflows.

There might come a tipping point at which sequencing replaces or is used alongside many other tests and diagnostic tools. This will require a significant increase in the availability and location of diagnostic sequencing, closer to the patient. However, near-patient sequencing requires innovation across the pathway. As discussed above, there not only needs to be technological progress to enable near-patient testing, it will need to be supported by good storage and transfer bandwidth, and consistently high standards of quality assurance. These improvements need
to start at the front end, with how samples are collected and handled, all the way through to the output and the format of the report, who it reaches and how the data is used for research and surveillance as well as individual health. This will be necessary if the UK is to have a vibrant sequencing technology ecosystem, with a range of interoperable partners drawn from the NHS, academia, and commercial and third sectors.

To ensure the technology meets these needs and addresses the challenges in developing a near-patient system with standardised outputs and aggregated data, the UK Government could consider offering mechanisms to achieve this. This could include a commercial imperative through commissioning frameworks, a challenge prize or other means.

**Recommendation 11: The UK Government should incentivise the development of sequencing technologies to ensure a vibrant genomic ecosystem attractive to investors willing to develop innovations for the benefit of the patient and the sequencing pathway.**

Metagenomics is the study of genetic material from environmental samples and has been used to great effect during the pandemic, most notably in wastewater sampling to detect SARS-CoV-2 and determine the presence and distribution of variants of concern. The wastewater sampling surveillance scheme was effective in tracking local outbreaks of Covid-19 and provided an early warning system of variants of concern.

The wastewater programme will be continued beyond the pandemic, with the aim of tracking other pathogens. This continued programme is welcomed and should be supported by close cross-functional working with environmental health and veterinary professionals to provide an early warning of potential outbreaks of disease or AMR.

Beyond wastewater, metagenomics has been beneficial in sequencing pathogens to distinguish between different types of viral haemorrhagic fevers. For example, Great Ormond Street Hospital for Children (GOSH) in London has developed the first accredited metagenomics service outside the US to identify pathogens in undiagnosed encephalitis.

Metagenomics usage will increase as costs reduce and benefits become more apparent, which will have an impact on pathways and diagnostic workflows. For these benefits to be fully realised there needs to be more research, development and investment to establish the role of metagenomics services in the UK, including delineation of use cases based on impact on the quality of patient care and patient outcomes.

**Recommendation 12: The Government should invest in research and development to establish the role of metagenomic services in the UK, including impact on patient care and outcomes.**

Since the start of the Covid-19 pandemic, across government departments and among policymakers and the public there has been a marked increase in awareness of the potential for pathogen sequencing to support public health. There has also been rapid expansion in genome sequencing capabilities and effective application to generate public health knowledge, including through the sequencing of samples from people with Covid-19 and wastewater.
Further investments are being made to apply sequencing to other areas, including funding to develop an end-to-end system to characterise antimicrobial-resistant pathogens in the food chain ‘from farm to fork’.

Looking ahead, the aggregation of these datasets could result in world-leading surveillance that is unique in terms of its breadth and depth. It should include a national genomics observatory service that brings together the data to provide ongoing UK surveillance and advanced predictive analytics. This surveillance should consider the capabilities required to detect new emerging threats, as well as taking account of the wider public health context, including health inequalities and climate change.

National plans will provide important building blocks for international initiatives, including the proposed Global Pandemic Radar and the new World Health Organization (WHO) Hub for Pandemic and Epidemic Intelligence.

**Recommendation 13:** The Government should establish a genomics observatory service bringing together metagenomics, environmental health, animal health, public health and the NHS to track and identify potential outbreaks of disease or AMR.
The risk for disease that we inherit in our genes is very different from other factors that influence health. Even though it could take decades to manifest, the risk is there from the moment we are born and does not change over time. Moreover, science is becoming ever better at reading that risk, driven by research programmes across the world on the genomics of diseases ranging from early-onset developmental disorders to late-onset disorders associated with ageing. From the perspective of prevention and early detection, such advances offer huge opportunity; those people identified as high risk can either be put into enhanced monitoring programmes to detect disease early and when intervention is most effective, or, where possible, put into preventative programmes that allow individuals to sidestep their risk. This logic is undeniable but also challenging; we have the potential to improve health across a wide spectrum of disease but realising that benefit requires a much wider use and acceptance of genomic data within healthcare and society.

Our roundtable event set out to ask some basic questions about where and how to advance the programme of genomic prevention and early detection. Which areas of healthcare are the most likely to benefit from the widespread use of genomic risk information in the near term? How would risk information be integrated into existing healthcare practice? How should we engage with a population that has limited experience with such information? When in an individual’s life should such information be collected? We assembled a broad group of experts from clinical, population health, academic and commercial backgrounds to consider these questions, with an emphasis on adult cancer screening, patients and the use of data, screening in primary care and newborn screening.
Several recurring themes emerged from these discussions. Firstly, our common view is that there are areas of healthcare – in particular, screening for common adult cancers and preventative medicine in cardiovascular disease management – where genomic risk information would benefit individuals today. Here, we need implementation trials, outcome modelling and health economics to work out how best to make changes happen. Secondly, we agree that genomic prevention and early detection will only ever succeed if we empower individuals to understand genomic risk, manage it and participate in the ongoing programme of evidence generation and data sharing needed to improve and develop the science. A key challenge here is addressing the gaps in data for groups that have historically been under-represented within genomic research. Lastly, we acknowledge there are areas where there is still genuine uncertainty about how best to proceed and where the weaving together of scientific feasibility and ethical imperative is work that is yet to be completed.

The UK has generated some of the most important and innovative programmes in the field of genomics, particularly at the interface between research and clinical practice, such as the 100,000 Genomes Project and Our Future Health. The Covid-19 pandemic has demonstrated this global leadership and accelerated progress on important issues around data sharing and pathogen sequencing, as well as highlighting the need for greater investment to ensure the advances in laboratory genomics are translated into patient benefits.

**Early detection of disease**

With recent advances in genomic screening, there should be consideration of whether the UK’s existing screening programmes offer the best cost–benefit returns. It is now possible to use genomics to stratify individuals based on their risk, which would theoretically enable patients to be called for screening at a frequency linked to individual risk. An ‘SNP’ score can be used to refine risk and there is evidence of success of this approach in breast cancer, whereby those women at highest risk get more robust MRI-based screening, and those at lowest risk get mammograms on a less frequent basis than currently (see box below). Such tools are also moderately well developed for prostate and colorectal cancer.

**SNP for stratification of breast cancer risk in women**

It has previously been shown that panels of the earliest single nucleotide polymorphism (SNP) markers to be identified aid breast cancer risk stratification in white women, which is maintained after accounting for classic risk factors and mammographic density.

To assess whether such techniques could be used more widely, a case–control study of women aged 47 to 73 attending routine screening in Manchester was undertaken. In this study, the predictive ability of a predefined panel of 143 SNPs was assessed after adjustment for questionnaire-risk factors and mammography density. The panel showed substantial improvement in risk stratification in combination with classical risk factors and mammographic density, for both oestrogen receptor-positive and receptor-negative breast cancer.
The study showed that the SNP143 score was a risk factor for oestrogen receptor-positive (ER+) and negative (ER−) breast cancer, and that polygenic risk scores based on a large number of SNPs improve risk stratification in combination with classical risk factors and mammographic density.

Another programme that could demonstrate the effectiveness of genetic-based risk stratification in the short term could be application in populations already identified as having a monogenic disorder, for example Lynch syndrome. Many of these individuals already have experience of genetic screening and could be part of a proof-of-concept study to differentiate screening intervals based on risk.

However, reducing the frequency with which some patients are called for cancer screening could cause concern among patients, as it could be seen as removing a service rather than targeting it at those patients at highest risk of developing disease. A wider discussion about the evidential and ethical basis for reduced screening for individuals without risk factors is required.

Despite this growing evidential base, there remains a lag between academia and adoption in clinical practice. There needs to be a more dynamic system to encourage real-world studies and promote the adoption of academic discoveries into day-to-day practice. One exemplar for this is how Genomics England completed the 100,000 Genomes Project and supported NHS England in the creation of the new NHS Genomic Medicine Service within six years. The time is right for policy to encourage the emergence of a national ecosystem that will pick up the increasing insights from research and make them quickly available to citizens to improve their lives, and to create a national genomic research library linking major datasets.

**Recommendation 14:** Opportunities for the use of genomic-based risk stratification for screening should be explored and integrated with existing screening programmes where necessary. This should be supported by a national genomics research library linking major datasets to identify the insights from research.

Shallow whole genome sequencing (sWGS) is a mature technology and can be used to support the diagnostic pathway. It is simple and cost effective, although there remain some challenges around sensitivity. It is particularly effective in cases where there is a clear link between an individual's genetic profile and their risk of cancer, such as those carrying BRCA genes elevating the risk of breast cancer. In these cases, it might be inexpensive to find the gene, but consideration needs to be given to the specific liquid biopsy test that would be used and the frequency with which individuals would be tested.

**Recommendation 15:** There should be robust investigation of the role of long- and short-read whole genome sequencing alongside multi-omics to understand how these can add biological insights and health gain. The potential utility of shallow whole genome sequencing (sWGS) and where it could have value merits exploration. Consideration should be given to how ctDNA and sWGS could be combined for those individuals at high risk of diseases caused by known genetic variations.

**Screening in primary care**

One of the longer-term goals of any genomic screening programme must be the prevention of disease, and in particular long-term conditions such as cardiovascular (CV) disease and diabetes. However, once an individual's risk has been identified the
more significant and complex challenge is to how to reduce the risk and prevent
the prediction coming to fruition.

One immediate potential use case could be incorporating polygenic risk scores into
primary care for CV disease. While this might not be able to demonstrate reduced
mortality or impairment in the short and medium term, research could show
whether identification of this risk and discussion with the patient has an impact on
proxies such patients’ behaviour or GP prescribing.

**Recommendation 16:** Polygenic risk scores could be routinely used in primary care
to identify those patients at highest risk from cardiovascular conditions and other
disease. This will necessitate professional engagement and involvement to establish
decision support tools and mainstream this in the NHS.

**Newborn screening**

Currently all newborns are offered the nine-point Guthrie test within the first week of
life.\(^72\) The role of whole genome sequencing (WGS) in early life to diagnose conditions
that would require intervention in childhood is currently being explored.\(^73\)

Offering WGS to newborns allows for the early detection of disease, which in turn
can allow for early intervention to reduce or even avoid disability. However, WGS
for newborns does need to be carefully considered from an ethical perspective.
Many parents might not want to know the risk of all diseases from which their
child is at risk, particularly those that will not manifest until later in life, or those
for which there is no treatment. It is also not always possible to separate screening
for childhood conditions from adult conditions and identifying a condition using
screening in childhood might inadvertently highlight an increased risk in one or
both parents. Moreover, genotype rarely is a deterministic predictor of phenotype
and there can be mutations of unknown significance. One way of addressing this
challenge could be ‘targeted testing’ to conditions where early intervention could
reduce disability or harm rather than reporting on the whole genome.

In the UK, recent public dialogue research from Genomics England and the UK
National Screening Committee found that members of the public are broadly
supportive of the use of WGS in newborn screening. However, this study did
conclude that “proper consideration should be given to designing and planning
any future use of this technology … including involving the public and ensuring
appropriate resources, investment and safeguards are in place”.\(^74\)

**Recommendation 17:** There should be a pilot of whole genome sequencing (WGS) in
newborns, conducted in line with the recommendations of the Genomics England and
UK National Screening Committee public dialogue.\(^75\)
Introduction from Dr Jacqui Shaw, Professor of Translational Cancer, University of Leicester

The advances in genomics and whole genome sequencing (WGS) at scale are now revealing the true extent of human genomic variation. The detection of cancer at an early stage while it is curable by surgical resection is one of the most effective ways of reducing cancer mortality. Liquid biopsy, notably analysis of circulating cell-free DNA derived from tumour cells, termed circulating tumour DNA (ctDNA), is a clinical tool to support earlier detection of cancer, surveillance and personalised medicine approaches. This ctDNA can be differentiated from normal cell DNA by specific genetic alterations, including somatic mutations, methylation patterns, copy number changes, and different fragment sizes. CtDNA testing was first implemented in the UK to detect specific epidermal growth factor receptor (EGFR) mutations in patients with non-small-cell lung cancer (NSCLC).

Implementation of ctDNA testing could also transform the management of cancer, for example for detecting residual disease after chemotherapy/radiotherapy, identifying resistance and relapse earlier, and triaging patients to targeted treatments. The challenge is also to increase earlier detection of cancer where there is a greater likelihood of successful treatment, for example in asymptomatic population screening. To this end, the GRAIL Galleri™ test will now be trialled in the NHS for early detection of cancer. In support of this, an initial attempt of its clinical validation was recently reported in Annals of Oncology. 76

This group was brought together to consider the key features and infrastructure required to develop and deliver a world-leading national early detection and monitoring strategy to transform healthcare in the future. Questions for discussion in the session included:

4) Cancer: early detection and surveillance
1) How do we best capitalise on the leading position of the UK in genomic research, healthcare data and clinical trial methods to increase earlier detection of cancer, when there is greater likelihood of successful treatment?

2) What ctDNA technologies are ready for clinical implementation, if any? What actions should the UK Government take to support the use of liquid biopsy technologies for early detection and prevention of cancer?

Our considerations were complementary to other excellent publications, such as Cancer Research UK’s Early detection and diagnosis of cancer: A roadmap to the future, and supportive of its recommendations.

The UK’s response to the SARS-CoV-2 pandemic has demonstrated the rapid developments the scientific community can deliver when it collaborates. But the pandemic has also shown how quickly standard care pathways can fall apart in the face of a global health emergency. The NHS and other health systems globally should use the pandemic as a chance to reflect on current practice and consider whether this is an opportunity to reorganise clinical pathways for the benefit of patients and the wider health system, taking advantage of new innovations.

Cancer is a disease of disordered tumour genomes, and in some cases it might be triggered by inherited predisposition to malignancy. There is growing evidence that genomic data from sequencing can enable precision cancer care. The role of newer technologies such as long-read sequencing and deeper ‘omic’ characterisation of tumours, for example, measuring RNA and integrating this data with digital pathology and imaging, requires further research. This might enable a comprehensive multimodal picture of cancer that helps optimise therapy and chance of a cure. Therefore, it is important we expand the cancer evidence base to fuel transformed cancer care.

One of the most exciting innovations of recent years is liquid biopsy, notably circulating tumour DNA (ctDNA), which is a clinical tool to support earlier detection of cancer, surveillance and personalised medicine approaches. CtDNA testing was first implemented in the UK to detect specific epidermal growth factor receptor (EGFR) mutations in patients with non-small-cell lung cancer (NSCLC). CtDNA testing is now being used as part of the GRAIL Galleri™ blood test in the NHS, which aims to detect cancer earlier than conventional methods.

The current usage of ctDNA is to increase detection of cancer at a point when there is a greater likelihood of successful treatment. However, ctDNA can be used in a number of ways, including:

- screening of an asymptomatic population
- diagnosing disease
- detecting low levels of disease (i.e. after chemotherapy/radiotherapy)
- identifying recurrence/relapse earlier
- triaging patients (within the 62-day wait for treatment after diagnosis target) to prioritise those at greatest need of treatment.

CtDNA and liquid biopsy technologies are not without their challenges. Recently a large-scale trial for the early detection of ovarian cancer, the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS), looking at the CA125 protein blood test, failed to reduce mortality when compared to the control group. While disappointing, these results should not signal a wholesale move away from the
use of ctDNA techniques; it should instead focus efforts on the best and most targeted use of this technology to improve patient outcomes.

As highlighted above, there are various areas in which ctDNA could be used to support and enhance clinical practice. One of the advantages of the technology is that, unlike some other diagnostic techniques, ctDNA does allow for longitudinal follow-up and repeat testing. ctDNA has been shown to be particularly effective for the identification of minimal residual disease in patients at risk of recurrence or relapse.\(^ {81, 82}\)

Additionally, ctDNA can be used to support patient triage and prioritisation. Even prior to the pandemic, workforce and diagnostic capacity was stretched in many secondary care settings. ctDNA could be used in the diagnostic pathway to triage patients to prioritise the most urgent and those in greatest need of treatment. This would be beneficial for individual patients and would also be more efficacious and cost effective, as well as building resilience into the system ahead of future health crises.

cDNA could in theory be used for routine population screening to identify cancer earlier. However, as has been seen from the UKCTOCS study, this approach is not without its challenges. Widespread community use of ctDNA for screening would also require significant investment in infrastructure, both to deliver the screening tests and to support the investigation and diagnosis of these additional patients entering the system.

While there is a growing evidence base of the effectiveness of ctDNA in clinical practice, in some cases it might be necessary to conduct randomised controlled trials (RCTs) for the validation of assays. In other cases, there might already have been significant technical validation, demonstrating their effectiveness when compared with other techniques, such as tumour assays. In such cases, a full RCT might not be required, but rather a service evaluation or a cost-effectiveness evaluation.

Genomic tests are evolving rapidly, but there will remain significant NHS trust-level testing until the seven Genomic Laboratory Hubs (GLHs) are fully embedded. There is only one mention of ctDNA in the current National Genomic Test Directory,\(^ {83}\) and the pace of adoption of ctDNA technologies onto the test directory has, to date, been slow. There is clear utility for greater ctDNA testing in terms of avoiding repeat biopsies, accessing hard-to-reach tumours, tracking disease evolution and tracking response to treatment. However, it is important that whichever assays are selected for the test directory are useful and usable for clinicians and will have a positive impact on a patient's treatment and care.

Those individuals and organisations developing the test directory should also engage with a full range of stakeholders, including industry, as the directory is updated. It can take time for liquid biopsy assays to be ready for any type of large-scale commercial use, and therefore earlier engagement with those commercial organisations developing the tests would be beneficial.

There is currently significant variation in where ctDNA testing is conducted, and the turnaround time for the results. As the use of these tests becomes more widespread, consideration will have to be given to where these tests are done, and how quickly a result is required and reported. In some cases, for example in screening or early detection, near-patient testing might be more valuable than rapid turnaround. Yet in other cases, such as when used alongside precision medicine, rapid turnaround might be beneficial. An additional bottleneck for
the wider use of ctDNA testing is the interpretation of the clinical data, which is managed through a molecular tumour board.

**Recommendation 18:** Circulating tumour DNA (ctDNA) has the potential to be used throughout the clinical pathway, including screening and diagnosis, as well as triaging patients and identifying relapse or recurrence. There should be a clear roadmap via a technology agnostic platform to wider testing and, where appropriate, adoption of ctDNA technologies. This could include further exploration of the benefits and risks, and the use of randomised controlled trials when necessary. Such a platform could enable accelerated evaluation and potential adoption onto the National Genomic Test Directory of proven ctDNA technologies to allow for greater use in clinical practice. If these technologies are adopted, then early consideration should be given to standardising approaches by genomic tumour boards.

Over the next two years the NHS will be trialling GRAIL Galleri™, which aims to detect more than 50 types of cancer through a blood test. This is a window of opportunity to gain the right type of evidence and accelerate the use of ctDNA in clinical practice. Genomics England has already conducted a liquid biopsy pilot, and it has shown that some commercially available assays were more effective and sensitive than others. The GRAIL trial could be expanded into a platform trial to enable the community to explore multimodal approaches. A multimodal approach could be particularly beneficial for some cancer types, where it is possible that there will need to be a combination of technologies used to get sufficiently high sensitivity, particularly in early detection. It would also enable parallel studies on patient samples.

**Recommendation 19:** The future of cancer care is expected to include multimodal imaging, digital pathology, multiple biomarkers and genomic sequencing. This could include accelerated integrated pathology processes, the rapid application of multiple biomarkers including ctDNA, tumour proteomics, metabolomics and long-read tumour DNA sequencing with real-time analysis enabling earlier precision cancer care. A first-in-class global multimodal test platform could transform cancer detection and migrate later-stage presentations to earlier detection, which would make the UK an extremely attractive test bed for life sciences investment.
Introduction by Dr Gosia Trynka, Group Leader, Wellcome Sanger Institute

The advances in genomics and whole genome sequencing (WGS) at scale are now revealing the true extent of human genomic variation and its role in human disease. The question is, can academia, industry and healthcare systematically harness this new knowledge and unlock the function of these variants, informing new diagnoses, risk assessment and genomically primed therapies? A therapy is twice as likely to reach the clinic with an underpinning of genetic support. Our response to the pandemic has showcased the rapid developments the scientific community can deliver when it is brought together. We need to apply the same scientific endeavour to defining the functional role of the genome. To address this, the UK Life Sciences Vision and Genome UK strategies propose a nationally defined strategy for functional genomics. This world-leading expert group is being assembled by Public Policy Projects to prepare an authoritative report that could inform government, academia, industry and funding bodies' planning and strategy. There are seven groups advising on the optimal considerations and approaches to implement the Genome UK strategy.

This group will consider the key features and infrastructure required to deliver a world-leading national functional genomics strategy to include:

- establishing the extant human capability, infrastructure, investment and location
- suggesting how these assets could be harnessed
- identifying the new capabilities, infrastructure and configuration required for state-of-the-art functional genomics, now and over one, three and five years
- defining the optimal approaches to high-throughput gene editing and cellular phenotyping

5) Functional genomics and therapeutic innovation

Introduction by Dr Gosia Trynka, Group Leader, Wellcome Sanger Institute
The study of genomics is concerned with the relationship between genetic variation and phenotypic outcomes, and the mechanisms through which these events operate. Functional genomics is translating this knowledge into benefits, by moving from identification of a genetic variant to understanding its function and impact on human biology to translating those discoveries into drug development. Functional genomics will not just benefit the few people who might receive expensive gene therapies; it also has the potential to benefit many via substantial advances in fundamental biological understanding. This might ultimately result in better medicines and treatments for everyone, regardless of their genotype.

As well as its use in drug discovery, functional genomics also has a role in pharmacogenetics, by supporting the targeting of treatment and the prevention of adverse events based on genotype. It is estimated that 99 per cent of people have a variant from the known clinical pharmacogenetic international consensus guidelines that is likely to result in a medicine being ineffective, or worse, causing harm.65

Unlocking the potential of functional genomics will bring benefits for patients and generate value for the UK in return for the investments that have already been committed to genomic research. One of the most significant challenges in this field is translating genomic advances into the clinical domain. To do this it will be necessary to demonstrate both the clinical benefits and cost effectiveness of genomics to the health service and to clinical practice. Only then will it be possible to deploy this technology at scale.

In the UK there are large and well-funded initiatives with an interest in functional genomics as well as world-class universities and a thriving life sciences sector. During the pandemic there have been some excellent examples of cross-functional working, with joint funding calls and integrated cohorts, and this collaborative approach should be incentivised to continue.

To encourage this, there needs to be greater exploration of how existing resources and funding could be better distributed so that the limited financial support available is not divided between multiple similar small projects, none of which alone can be impactful or transformative. Continuing to fragment the available resources could impact on the UK’s competitiveness compared with well-resourced larger countries such as the US.

The UK also has strong clinical research capabilities in the NHS. However, too often the scientific and clinical strengths of the UK work in parallel, rather than collaboratively. There are well-integrated centres, but there are other centres where there are strong capabilities on both the clinical and academic sides, and often shared facilities, but limited integration in terms of collaboration. To address this, funding routes and infrastructure should be designed in a way that encourages collaboration and increases connectivity. Doing this would enable large-scale genomic research projects, including patient recall by genotype.

Over the last decade there has been increasing collaboration between academia and industry, strengthening the UK ecosystem. This collaboration is to be welcomed, but more needs to be done to create an environment that fosters better dialogue, collaboration and exchange of expertise. There needs to be more
partnership working between industry and academia, sharing expertise and facilities. Currently, contracting between academia and industry can be a challenge, with multiple contracts making progress slow and bureaucratic. To address this, standard contracts that allow industry to partner with both academia and the NHS simultaneously would be extremely beneficial.

**Recommendation 20:** The ambition must be for a nimble and collaborative ecosystem with a joint ethos, infrastructure and goals. To do this, organisations, including academia, industry and the NHS, should combine skills and expertise and work together on large-scale projects harnessing their collective strengths. This will require the establishment of a collaborative network, bringing organisations together and encouraging an exchange of expertise through a portfolio of large-scale collaborative and interdisciplinary projects.

However, for the UK to maintain its excellence and world-leading position in the field of functional genomics, efforts are needed to ensure enough talent is entering the field. This is important not only for research but also for clinical practice, as genomics is increasingly a part of patients’ pathways.

Currently, there is only a limited amount of genetics, genomics and data science in the undergraduate medical curriculum and very few genomic medicine MSc or DPhil programmes. There are currently seven UK masters programmes in genomic medicine and six in health data science.

**Recommendation 21:** Every doctor should be trained in data analysis. They should be introduced to the basics of data analysis and coding as part of the medical undergraduate curriculum. A substantial subset of doctors should be able to undertake master’s degree level training in genomic and/or data science as applied to health, to ensure we maintain the benefits of genomic healthcare.

New genome technologies, such as base editing and prime editing to test variants at scale are being developed. Newer methods of gene editing with higher precision are rapidly emerging. Going forward, it is likely that millions of variants will be tested simultaneously, and new technologies will be benefiting patient care in ways as yet unimagined.

Despite much progress, the science is not yet at a stage where all the building blocks for translating genomics ‘to the bedside’ are ready. There is still significant investment required in basic research to streamline the path from a disease genotype to a drug target. Therefore, alongside investment in large-scale translational projects, support for basic science must continue to ensure the UK remains an attractive location to conduct this vital research.

**Recommendation 22:** Technology in the field of functional genomics is moving at great speed and it is challenging to predict where the field will be in five years. The UK needs to continue to invest in basic scientific research and technology development to enable engineering and testing of the role of human variants at scale.

The goal of functional genomics must be to move from gene identification to function to diagnosis to patient therapy as rapidly as possible. These tools should be applied to every single patient in the NHS database by delivering tangible benefits for their health. This might be in terms of rare genetic variants, or through better targeted treatments, or through the development of polygenic risk scores.

There are currently 154 trials of advanced therapy medicinal products (ATMPs) in the UK, the majority employing viral vector mediated gene transfer. However,
in some areas diagnostic yield is very low, and even when a genetic variant is identified it might not result in disease because another variant prevents immediate presentation. To address this there should be investment in a platform to support the identification of variants and the development of treatments, building beyond the work of large UK genomic initiatives such as the 100,000 Genomes Project\textsuperscript{87} or UK Biobank.\textsuperscript{88}

**Recommendation 23:** The UK Government should make a major investment over the next five years in a platform to support more rapid progress from gene identification to function to treatment, which would ultimately bring the benefits of genomics to every patient in the NHS.

To support this 'moonshot ambition' there needs to be investment in next-generation digital pathology, linking imaging, genomics and tissue samples. Currently, there is good sample provision from NHS centres into functional genomic phenotyping pipelines, but it is costly, and requires a significant amount of clinical infrastructure, laboratory capacity, digital storage and analysis.

The UK Government should aspire to digital pathology for everyone who has a tissue sample taken in the NHS in the next decade. Doing this will require investment in storing this data and the associated images as well as in developing a common platform and process for testing and storing the samples. It must be based on an open data access model to allow for large-scale research and analysis. This would be a world-class resource for functional genomics, would be attractive to inward investors, and could generate significant added value for 'UK plc'.

**Recommendation 24:** The Government must invest in making the UK a world leader in multimodal data for health, including next generation digital pathology, bringing together tissue samples, genomics and imaging across all disease areas.
6) Ethics: the public and patient perspective

Introduction by Professor Sir Jonathan Montgomery, University College London

Genomics strategy is not just for scientists and industry, it must also deliver benefits for all members of our society. It is not necessary for everyone to receive precisely the same things, but we must aim to ensure that we will all be better off in ways that we value. This requires extensive public deliberation through which we can explore the opportunities for advancing knowledge and deploying it fairly, along with the concerns that things might go wrong and how we might protect against that possibility.

Such public conversations are not easy to hold. We recommend the establishment of a new human genetics commission to take on the responsibility for ensuring that explorations are inclusive and respectful of the full range of views. That commission must ensure that we can draw on the insights of the best available experts from the very wide range of relevant disciplines. It must create the space in which we can refine the principles on which we should approach the development of genomics. It must also ensure that there are mechanisms for scientists and policymakers to test out the public acceptability of opportunities. It needs to provide assurance that we have an accurate picture of the way in which genomic data is used, including the protections in place for our privacy.

This transparency provides the foundation on which it is reasonable to invite us to place our public trust in the future governance of genomics. There are exciting prospects ahead of us, but we must not risk genomics being used in ways that are exclusive or divisive. That would undermine public confidence in the link between scientific advances and the common good. We must be able to prioritise our impacts, and it is inevitable that we will need to choose some options to pursue over others.
However, those choices must retain close connections with our shared values if we are to make the most of the opportunities that genomic medicine offers us. We propose a new human genetics commission with strong patient and public membership that involves, engages and enables strategy setting by those most affected by illness and could serve as the guardian of that relationship.

Genomics is a rapidly advancing field with discoveries enhancing our understanding of human biology and the world around us and enabling the enormous strides in vaccines and virus tracking that have saved a very large number of lives during this pandemic. It is therefore important that there is an open and honest conversation with the public about genomics and its potential to affect healthcare. The UK is at a crossroads, and failure to consider the ethical implications of these scientific advances could have profound impacts on our health and the UK research community. The science should not be advanced in isolation from consideration of the ethical implications of genomics. But considering the ethical implications should not act as a brake on the science; rather, it should be viewed as a navigation system to lead us in the right direction.

Genomics is an extremely complex and rapidly evolving field that can be hard for individuals to fully comprehend. Often the public debate around genomics is a very deterministic one. Many people understand single gene determinism (e.g. BRCA genes and breast cancer) but are less clear on concepts such as polygenic risk scores, which work in probabilities not certainties. There is also much still to discover about the link between genetics and health risks and outcomes.

To achieve the objective of an open and honest conversation with the public about ethics relating to genomics, there are three principles on which any programme should be built: diverse inclusion, respectful deliberation, and good governance.

**Diverse inclusion**

When having any ethical debate, but particularly on a topic as important as genomics, there must be diverse inclusion of public, patients and experts. It is necessary to be proactive about establishing the legitimacy of the ethical foundations of our approach to genomics. This requires the rich range of expertise that is available from different disciplines to be brought to bear on our understanding of the scientific, moral and social implications of genomics. To establish public acceptability as an essential foundation for sound policy there needs to be investment in an inclusive public and patient deliberation process.

Too often these public deliberation processes are dominated by those who can afford the time, money, and energy to be involved. A wide range of methods now exists to support public debate, including social media and virtual meeting platforms. These can enhance and extend the opportunities for inclusive deliberation for many but might exclude those without access to the technology or infrastructure to participate.

If engagement is considered an important component of ethical governance, then consideration must be given to the accessibility of these structures to different groups. If an individual wants to know something about genomics, wants to say something about genomics or has a concern or hope they want to raise about genomics, what channels are there for facilitating these conversations? And how should the outcomes of such conversations feed into governance?
It must therefore be made easier for those who are not participating in the formal engagement processes to share their views, either solicited or unsolicited. There should also be social media listening exercises to understand what is being said via those channels.


“Effective democratic deliberation calls for inclusive and respectful debate and depends on collaborative decision making. Stakeholders with a range of perspectives are encouraged to present their views, seek common ground whenever possible, and maintain mutual respect even when irreconcilable differences among viewpoints remain.”

Respectful deliberation

It is also important to consider the objectives of public and patient engagement. Is the public a source of legitimacy (e.g. social contract) or are they advisory (to learn what people think and adjust behaviour accordingly)? Or is public and patient engagement a cover for educating people into the ‘expert’ way of thinking? Too often public and patient engagement is used as a cover for public relations, steering people towards already pre-decided priorities or views. Efforts must also be taken to be attuned to the path of dependencies driven by early decisions, so as not to inadvertently steer participants towards one outcome.

Explaining genomics is difficult, and the data is complex, and articulating it in a way that builds and retains trust will be challenging. This is exacerbated by the continually advancing science changing our understanding of genomic information.

It is therefore important to invest time and resources in building potential participants’ understanding of genomic ideas and concepts so that they can fully engage with the conversation and provide an informed opinion. Equally, it is vital to pay attention to the views of actual participants so that the debate is grounded in experience.


“Democratic deliberation is a method of decision making in which participants discuss and debate a question of common concern, justifying their arguments with reasons and treating one another with mutual respect, with the goal of reaching an actionable decision for policy or law, open to future challenge or revision.”

Good governance

There is a clear link between trust and transparency. Consequently, one way to build trust is by increasing transparency about where the genomic information comes from, how it will be used and by whom.
However, in the field of genomics traditional types of consent can be ineffective, and fully informed specific consent is hard to achieve and the attempt might have perverse consequences. To address this, it might be necessary to consider a more open-ended type of consent, which will evolve alongside understanding of the science. It could be the case that not everyone wants to understand every aspect of consent, they might just want to trust the people to whom they are giving their genetic material.

To ensure an ongoing conversation with the public and patients via a range of channels there should be a new organisation. This forum should be responsible for ensuring that the research and clinical communities do not move too far away from the ethical baseline that has properly balanced the inputs of public and patients. This could follow principles and practices such as those of the Irish Citizens’ Assembly, which has considered issues such as abortion, gender equality and climate change.\textsuperscript{91}

Too often the public conflates ethics and regulation, and people are disempowered from ethics as they think it is regulation. There needs to be engagement with a set of ethical principles separate from the governance framework under which things operate. The original Human Genetics Commission, an advisory non-departmental public body on the ethical and social aspects of genetics was created in 1999 and abolished in 2010. It offered some principles and terms of reference (see box below) that could be revisited for any new human genetics commission. We propose a national group combining citizens, including a majority drawn from public and patients, to advise on strategy and implementation of genomics in healthcare, supported by professionals. This could act as a ‘citizens jury on ethics’ so public and patient views can be solicited on a range of topics relating to genomics, continuously reviewing some aspects of genomics as they evolve, such as data sharing and the use of AI.

**Recommendation 25:** There should be a new human genetics commission, responsible for public deliberation on the ethical development and application of genomic science by setting the framework for engagement and good governance and ensuring timely advice to government on policy. The first task of the new human genetics commission should be to convene a public deliberation that brings a wide range of the public, patients and experts together to articulate and validate a framework of principles relating to genomics and the use of genomic data that is consistent with the relevant human rights and addressed to anticipated applications and scientific advances.

---

**Human Genetic Commission (1999 to 2010) terms of reference\textsuperscript{92}**

- To advise Ministers in the UK on the potential ethical and legal implications of human genetic knowledge and its applications for health, economic and social wellbeing, including:
  - the need to fill any knowledge gaps
  - the development of national policies
  - the requirement for legislative action.
- To encourage and promote public awareness and understanding of human genetic knowledge and its applications.
- To work in partnership and exchange information with other relevant organisations in the pursuit of the above.
- To consider specific issues related to human genetics and related technologies as requested by Ministers.
- To work in accordance with best practice for public bodies with regard to openness, transparency, accessibility and timeliness.
There are clear differences between Anglo-American ethical frameworks and European ones, for example on solidarity, reciprocity and dignity, which are more prominent in European frameworks than in the US and UK. Anything developed in the UK sits within a broader global context, and ideally, some elements of genomics frameworks should be tackled at a global level.

Effective data governance needs an underpinning of public consent. Governance only allows for processes to be replicated in a transparent and standardised manner; it cannot build public support where none exists. Therefore, as well as updating data governance frameworks and processes there must be significant effort to build public understanding and confidence in the data governance framework.

As well as building public support through education and debate, there must be an organisation for individuals to contact if they have concerns about the use of their genomic information. This body must have the statutory power to act and issue sanctions in the event of the misuse of genomic information.

**Recommendation 26:** There should be a genomics watchdog or ombudsman for the public to contact if they have concerns about the use of their genomic data.

---


Key principles:

**Genetic solidarity and altruism.** “We all share the same basic human genome, although there are individual variations which distinguish us from other people. Most of our genetic characteristics will be present in others. This sharing of our genetic constitution not only gives rise to opportunities to help others but it also highlights our common interest in the fruits of medically-based genetic research.”

**Respect for persons** “affirms the equal value, dignity and moral rights of each individual. Each individual is entitled to lead a life in which genetic characteristics will not be the basis of unjust discrimination or unfair or inhuman treatment.”

Secondary principles:

- privacy
- consent
- confidentiality
- non-discrimination
7) Global horizons in genomics

Introduction by Peter Goodhand, Chief Executive Officer, Global Alliance for Genomics and Health and Simon Linnett, Chair, Bedfordshire Hospitals NHS Foundation Trust

We come to the world of genomics from different backgrounds and countries but are equally persuaded of the immense health benefits that can derive from linking genetic data to clinical experience. Those benefits in the short term will cover the spectrum from more accurate and predictive diagnosis to more targeted and effective therapies. In the global pandemic we are seeing on a daily basis the importance of genomic analysis of both the pathogen and the host. In the longer term, with more comprehensive and nuanced understanding of the complexities and diversity of genomes representing all humanity, we will develop meaningful strategies to lower individual risk of serious disease. As genomics enters mainstream healthcare in areas such as pharmacogenomics and non-invasive prenatal or newborn screening, it will touch all our lives. In an industry worth £140 billion annually (in the UK alone), the opportunities are immeasurable.

The challenge will be, in many healthcare systems (even in the ‘joined-up NHS), finding ways of realising these benefits, particularly that of prevention. To do that will require significant structural changes so that learning health systems effectively implement health and wellness strategies at both population and individual levels.
As the Covid-19 pandemic has demonstrated, healthcare operates on a national or local basis, but human health is a truly global concern. The 100,000 Genomes Project\textsuperscript{94} led by Genomics England laid the foundation for UK leadership in the world of genomics. However, as the potential of genomics becomes more apparent, many more countries and private organisations are launching genomic and precision health initiatives. The next step is translating these scientific advances into frontline healthcare, and once again the UK, and the NHS, have been at the forefront of this advance.

The UK must now balance its leadership in this field with learning from, and including the contribution of, other countries and organisations. It will be important to understand how this balance between leadership and participation should be achieved, not just in science and healthcare but in innovative technology and its applications.

In addition, the UK will need to consider the balance between protecting patient privacy and the financial and other benefits that can accrue from using data for research. It is imperative to consider how to support innovation without prejudicing personal data. Finally, how can a system be designed to reward the innovators while ensuring fair return to the organisations and individuals who provided the data on which those innovations were based?

One of the most significant challenges facing global genomics is that while one million whole genomes have now been sequenced, 80 per cent of genetic research is done on Western Europeans.\textsuperscript{95} As a result, many of the findings might not be relevant to large sections of the global population. Often this is a result of genomic initiatives being funded by central government organisations, which might be restricted to funding in-country initiatives. These funding bodies might also prioritise research investment in the diseases of the developed world, such as cancer, cardiovascular disease and Type 2 diabetes. As such, many of the therapies and products that are being developed based on genomic research are targeted at higher-income countries.

Failure to include a more diverse range of samples risks limiting our perspective. This is particularly important when large datasets are being used to train AI algorithms, as unless the dataset is truly representative of the global population there is a risk of data bias in the result. It is well understood that there is normal variation in different populations, but more needs to be done to understand which variations are impactful on human biology and which are benign. Efforts must be made to take a more holistic approach to genomic research with a more representative selection of participants from across the globe as well as more genomic research focused on the infectious diseases that have the greatest impact in the lower and middle income countries.

Another challenge is around data, and the significant variance in levels of data maturity between countries. The International HundredK+ Cohorts Consortium\textsuperscript{96} has made some advances in linking cohorts around the world, and this progress is to be commended. However, more progress needs to be made in setting a minimum standard for data quality, and building interoperable global systems for data collection, storage and sharing, while maintaining high standards of information governance. This will build a platform of federated databases from which research can be conducted on a broad range of genotypes and phenotypes.

**Recommendation 27:** Building on the work of the International HundredK+ Cohorts Consortium, the UK should promote the development of a global initiative to support the collection of a more diverse range of samples, supported by truly federated databases to inform scientific understanding and therapy development.
Increasingly, technology is not the limiting factor to generating genomes. Analysing the genomic reports and understanding the function of genes is the key limitation to deployment in medicine. Failure to understand the influence of a particular gene limits our understanding of biology, which impacts on the ability to develop treatments and potentially even cures. To date there are some well-understood mutations from which treatments can be developed, but this number needs to increase if the true potential of genomic sequencing is to be realised.

To address this there needs to be a concerted global effort to speed the science from variant identification to treatment. The International Common Disease Alliance (ICDA)\textsuperscript{97} has begun this work, but there needs to be greater global collaboration and investment.

**Recommendation 28:** There should be a large-scale international collaboration and greater investment in programmes to move from variant to function to therapy to patient. This initiative would focus on functional characterisation and understanding, and how that can be taken forward into functional genomics and therapeutic uses.

The pandemic has demonstrated that pathogens do not respect borders or species, and global surveillance, rapid international data sharing and cross-border collaboration are vital to monitor and control outbreaks of disease. However, for this international collaboration to be most effective there need to be standardised methods for conducting pathogen sequencing as well as the analysis and sharing of data.

**Recommendation 29:** There should be greater investment in global genomic surveillance to identify emerging pathogens. To allow for rapid sharing of information about potential pandemic threats there should be international best practice minimum standards for pathogen sequencing that all hospital, commercial and academic laboratories should be encouraged to adopt. There should also be international protocols to promote rapid sharing of data for surveillance and analysis.
References
