



Unlocking the promise of UK health data

Supporting innovation in the
development of new medicines

Spring 2020



Contents

Executive summary	4
Introduction	6
Health data and the development of new medicines	10
Industry perspectives on building trust and confidence in the use of health data	14
1 Greater public understanding in how health data is used by biopharmaceutical companies	15
2 Robust protections for people's private information and how it is shared	15
3 Clear processes to share the benefits of research using health data	16
Harnessing the opportunities of UK health data	18
1 The data landscape is fragmented	22
2 Processes to access data are inefficient	24
3 The NHS is not sufficiently digitised to allow data to be linked and accessed readily	26
4 The quality and accessibility of the UK's health data resources is opaque	28
5 Regulatory and reimbursement processes do not recognise the potential that developments in health data can offer	30
Enhancing the UK health data environment	32

Executive summary

Patients, the NHS and the biopharmaceutical industry should be able to use the increasing amount of digital health data in the UK to improve outcomes for patients, reduce costs and enhance efficiency in the NHS, and discover, develop and deliver new treatments.

However, the full potential of health data cannot be realised until structures and processes enable straightforward accessibility. The industry has been handling health research data generated in clinical trials securely for decades and we can work with UK policymakers and the NHS to enable efficient, legitimate health data access for research and care.

This report outlines how the industry uses health data to develop new medicines, gives ABPI members' views on how to build trust with patients, details the industry's current views on opportunities and challenges with UK health data, and concludes with a proposal for ABPI members to work with the Government and the NHS to realise the potential of the UK's health data.

Health data is used throughout the process of developing new medicines, from understanding disease and defining unmet need to proving the value of a medicine in the real world.

This includes data on clinical history, diagnosis, genetics, current and experimental treatments, effectiveness, prices, costs and long-term consequences. This data can contribute value, and companies are investing hundreds of millions of pounds to develop, share and access it.

However, the proportion of these investments being made in the UK could be increased, attracting more global funding to help improve NHS data, develop the latest innovative therapies in the UK and demonstrate their value to UK patients.

This will require better public understanding of what health data is used for by companies, confidence in the strength of privacy protection laws and enforcement, and clarity on how the value from health data research is shared.

The recently developed 'Guiding principles on the NHS's use of health data' are welcome, but they are understandably high-level principles. Further fragmentation of health data must be avoided, approaches to sharing value must be worked out, and central guidance on practical implementation is needed. All stakeholders must support the highest standards of governance to ensure that trust amongst patients and the public is generated and maintained.

We gratefully acknowledge the contributions of our member companies in the process of compiling this report.

The UK's theoretical potential for nationwide longitudinal health data is well known. The UK has many excellent datasets and globally leading data scientists. We share Health Data Research UK (HDR UK)'s aims to increase the diversity, breadth and length of health datasets.

However, fragmentation of data sources, unclear and inconsistent access processes, incomplete digitisation, opacity of quality and accessibility, and regulatory acceptability of both novel outcome measures and real-world data all hold back practical data use.

The most significant issue faced by pharmaceutical researchers is the unpredictability and inefficiency of the data-finding and access process in the UK, which drives them to work with other countries. These are challenges that the ABPI wants to work with the UK Government, its agencies and the NHS to solve.

We describe a set of specific action areas to address the major challenges, where the ABPI is offering to add support and resources to existing Government and NHS commitments. These action areas aim to:

- Reduce fragmentation across the UK.
- Increase efficiency of data access processes.
- Enable clinical trials.
- Enhance transparency.
- Harness data to demonstrate value.

Industry supports initiatives aimed at improving the interoperability and accessibility of health data in the UK and is committed to identifying opportunities for collaborative projects, particularly with HDR UK and NHS services across England, Northern Ireland, Scotland and Wales.ⁱ



We believe that when these organisations achieve their aims for data management, we will be well on the way to unlocking the promise of UK health data. We call on the Government and its agencies to consider these action areas and actively support collaborations, perhaps in further 'Industrial Strategy' sector deals, that can accelerate the desired improvements.

This will enable UK health data to be used more widely to improve patient outcomes, enhance the efficiency of the NHS, and attract commercial data science and investment in clinical trials for the development of novel treatments in the UK.

ⁱⁱ This report is intended to help position the UK health data offer in an international context; the intent is to embrace all UK-wide NHS services and patients. Hence, although some of the organisations and initiatives cited may not be UK-wide, references in this report to 'NHS' typically refer to all the UK health services in England, Northern Ireland, Scotland and Wales.

Introduction

Over the last thirty years or so, advances in digital technology have enabled exponential growth in the amount of data we can collect, store, manage, analyse and share.

At every contact between a patient and the NHS, wherever it happens, information is generated. If this data could be systematically collected, stored and linked for each patient across the NHS in a way that supports straightforward access and analysis, the UK would have a unique opportunity to interrogate large-scale, detailed, longitudinal datasets.

The potential of research based upon health data at this scale (see Box 1) to help advance our understanding of disease, improve the way we manage patients and save lives is well recognised.

UK health data supports three goals:

- ◆ **Improved patient outcomes**
- ◆ **Increased efficiency within the NHS**
- ◆ **The development of effective new treatments**

Box 1: Our definition of health data

'Health data' as used here includes all information that could or should be included in every patient's health record (ideally held electronically) – for example, clinical examinations, signs, symptoms and diagnostic tests including scans and laboratory tests, treatments prescribed, records of vaccination, procedures undertaken and outcome measures, as well as similar information generated during the conduct of a clinical trial. Different subsets of this data across a group of patients will be relevant to different research projects.

The biopharmaceutical industry recognises the potential of harnessing UK health data combined with advanced analytics to achieve three main goals:

- To improve patient outcomes.
- To increase the efficiency of the NHS.
- To support the development of effective new treatments.

While the third is of particular interest to industry, all three goals are interlinked. Analysis of data linking patient outcomes to new approaches to treatment provides evidence to help refine optimal disease management guidelines and to support better clinical decision-making within the health service.

Understanding and implementing the optimal patient pathways for each disease, and at each stage of disease, can help health service leaders transform the quality of health and care services and reduce their cost, unlocking productivity benefits estimated to be worth up to £10 billion a year across the NHS in England.¹

Analysis of data on patient responses provides improved understanding of how diseases begin and progress, and together with the stratification of patients through genomic analysis and biomarkers, supports the development of new interventions to prevent, treat and perhaps cure disease.

The UK should be well placed to deliver on the promise of health data, given the perception of the NHS as a single organisation. While recent initiatives funded through the Industrial Strategy Challenge Fund via HDR UK, as well as NHS England's plans to capture health data digitally at a population-wide level (see box 2) are designed to improve the utility of health data, management of healthcare (and hence collection of health data) is devolved across the UK and the theoretical benefits of the NHS for health data research are not yet reflected in the practical realities.

During 2019, the Medicines Discovery Catapult undertook three studies to understand the health data needs of the life sciences sector, supported by the ABPI, HDR UK and the industry's Pistoia Alliance.

These studies (structured workshops, in-depth interviews and an online survey) enabled the identification of six themes around the potential use of health data:²

- Breadth, depth and scale of health data.
- The need for a single, easy-to-use route for access.
- The need for high-quality data.
- The need for expertise in areas such as artificial intelligence (AI) and analytics.
- Public trust and the need to return benefit from analyses and use of data to the NHS and the public.
- Cost-effectiveness of data access for all sizes of organisation.

The biopharmaceutical industry has been a pioneer in health data science, through the evolution of the design and conduct of clinical trials, and remains in the vanguard today.

Improved understanding of disease, and innovation in approaches to both diagnosis and treatment, inevitably lead to the development of new tests and outcome measures initially used in clinical trials that gradually become accepted as part of routine data collection in the NHS.

In addition, the industry has always been committed to the highest standards of governance of patient data collected in clinical trials, with decades of experience in ethical approval, patient consent and anonymisation of results reporting.

To make the most of UK health data, researchers, data custodians and others will need to be able to not only store, access and analyse data, but to do so under consistently high standards of governance in order to generate and maintain the trust of patients, the public and other stakeholders.

Looking to the future, the life sciences sector will need an increasingly large workforce trained in the ability to manage data within the governance requirements.

Recognising and welcoming that the UK Government as well as the leaders of the health services in England, Northern Ireland, Scotland and Wales increasingly understand and are seeking to unlock the potential of the UK's health data (see box 3), the ABPI has developed this report with our members, building on the themes identified

in the Medicines Discovery Catapult studies,² to update the industry's perspectives on where improvements can be made – and to describe what support we and our members can offer to help make them happen.

The report sets out:

- 1 A summary of the ways in which the biopharmaceutical industry uses health data to research and develop new medicines.
- 2 Our members' perspectives on the prerequisites for building trust amongst patients on the appropriate use of health data, informed by our experience of working with health data.
- 3 The opportunities that the UK's health data landscape offers to biopharmaceutical industry researchers, and the challenges that the industry currently faces in accessing and using health data for research in the UK.
- 4 The ways in which we and our members propose to work in partnership with the UK Government and NHS leadership to address these challenges, recognising that better data offers shared benefits and that improving health data should therefore be a shared endeavour.

We gratefully acknowledge the contributions of our member companies in the process of compiling this report.

Each year in England there are
400 million
GP appointments and
1.5 billion
diagnostic tests



Box 2: NHS England's ambitions on technology and digitally-enabled care are set out in its Long Term Plan⁴

"[The Long Term Plan] will result in an NHS where digital access to services is widespread. Where patients and their carers can better manage their health and condition.

"Where clinicians can access and interact with patient records and care plans wherever they are, with ready access to decision support and AI, and without the administrative hassle of today.

"Where predictive techniques support Local Integrated Care Systems to plan and optimise care for their populations. And where secure linked clinical, genomic and other data support new medical breakthroughs and consistent quality of care."

Box 3: the scale of NHS data generation

During every contact between a patient and the NHS (with general practitioners, nurses, emergency services, hospital specialists etc) information – health data – is generated. At the scale the NHS operates, the volume of data routinely generated and collected is enormous. For example, the NHS:

- ◆ Sees one million patients every 36 hours.³
- ◆ Provides 400 million GP and outpatient face-to-face appointments each year.⁴
- ◆ Undertakes 1.5 billion diagnostic tests each year.⁴

In addition, new technologies are being adopted which facilitate the collection of a wider variety of data through genetics and advanced imaging, and by patients themselves through wearable devices.

Health data and the development of new medicines

Figure 1a. The linear stepwise process of developing a new medicine



The discovery and development of new medicines has been a great triumph of science, helping people extend and improve their daily lives, with about 40 new medicines approved for use each year.⁵

However, it is a long and expensive global undertaking, with more than \$1.5 trillion invested by the biopharmaceutical industry in R&D across the world over the last decade – an annual spend of \$179 billion in 2018.⁶

Health data is used by the biopharmaceutical industry at all stages of the discovery and development process, as set out in the following chart and described in more detail (see Figure 1).

The routine collection of **health data** about patients helps doctors, the NHS and the biopharmaceutical industry understand the effectiveness of current treatments in routine use and the progression of disease under current treatment pathways, as well as supporting **pharmacovigilance** – the monitoring of the safety of new medicines after their initial introduction.

It helps us learn more about the underlying cause of diseases, how to detect them earlier and how they progress – and the impact this has on patients, and the costs for healthcare systems.

Health data can help **patient stratification** into more meaningful groups, leading to better understanding of why some patients respond better than others, and to the refinement of disease management guidelines, treatment algorithms and patient pathways.

This patient stratification, together with historic outcomes data, can help the biopharmaceutical industry identify and define the areas of **highest unmet need**. This shows the industry where it should focus its research investments, and how to define and measure the success of new medicines.

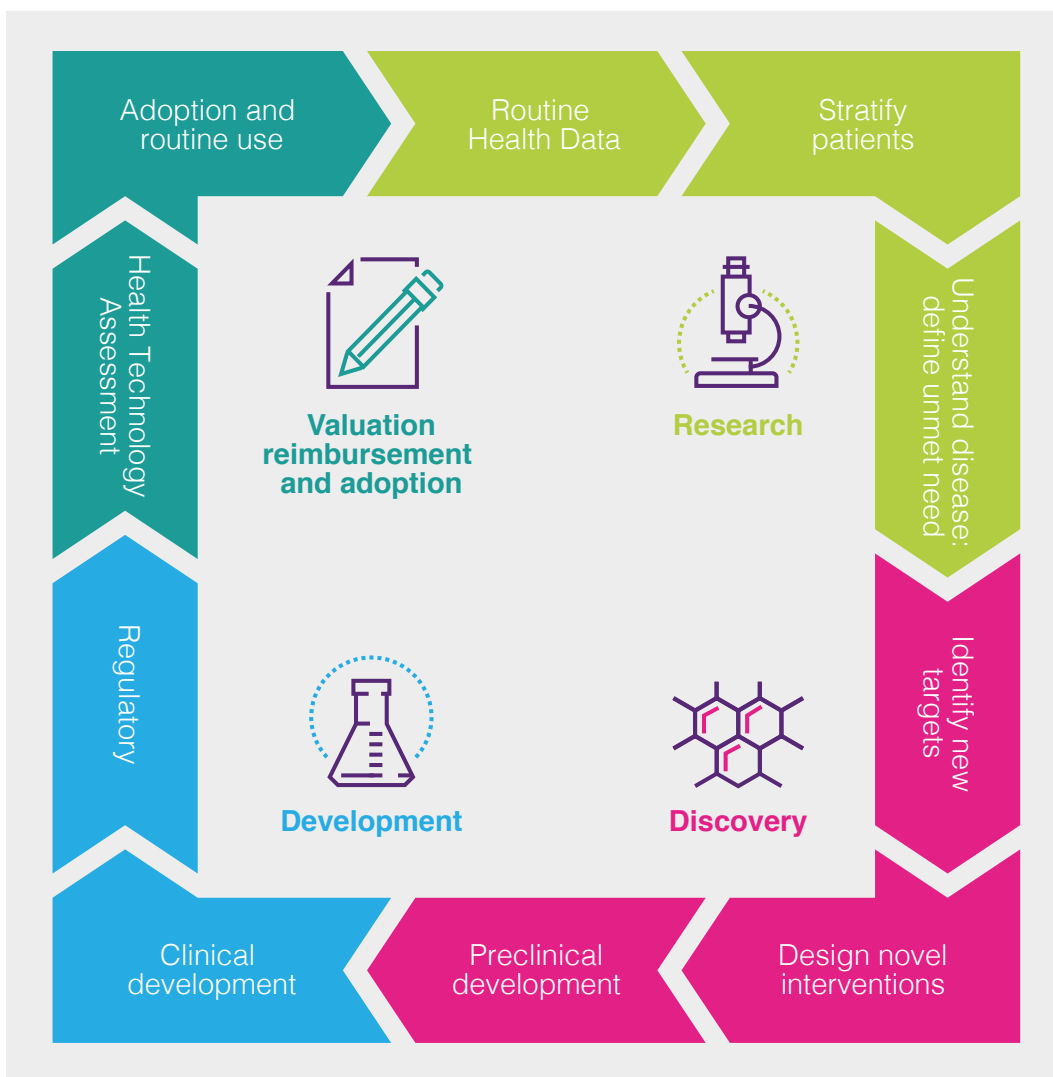
The global biopharmaceutical industry spends

\$179billion

a year on researching and developing new medicines.



Figure 1b. The cycle of research and development of new medicines; high quality health data can support each stage



Data from laboratory tests and from genomic and proteomic analysis combined with clinical data, can lead to greater understanding of disease processes. It can also help researchers to discover **new targets** and design **new interventions** to affect these targets and interfere with the disease processes.

Understanding the biochemistry of disease helps researchers to find biomarkers to indicate those patients most likely to benefit from treatments. Pre-clinical investigation will indicate whether a new treatment can be reliably administered to patients, and whether it will likely be tolerated and effective.

Once new experimental treatments have shown pre-clinical promise to enable the treatment of symptoms, the inhibition of the disease process or to offer the possibility of cure, health data can support the process of further **clinical development**.

Well-curated data can speed up the process of identifying and recruiting patients into clinical trials and can also help define new measures of success and trial endpoints.

For very rare patient groups or areas where it is unethical not to offer patients a new treatment, historic health data can allow researchers to find information on outcomes for untreated patients for comparison.

Data from these trials – increasingly a combination of routinely collected data and specific clinical and laboratory measures defined in the trial protocol – contributes to the **regulatory submission** that demonstrates that the balance between efficacy and safety warrants a marketing authorisation, allowing the medicine to be sold for use.

Further, more specific data on outcomes, costs and savings can then be used to support **health technology assessments** to demonstrate that a treatment represents value for money.

Health data is then used to monitor uptake of a new treatment and thus support its **adoption** in the health service and eventually (completing the cycle) its **routine use** – when the health data collected will help monitor ongoing safety, as well as evaluations of efficacy in further different patient groups.

Health data is therefore important at all points in the cycle of drug development, and the move to health data digitisation and sharing has created the potential for analysis of larger health datasets that will allow us to identify safe and effective drugs earlier in the development process, and therefore reduce the waste associated with work on drugs that subsequently are not approved.^{ii 7}

Furthermore, recent developments in routine genomic profiling have attracted significant industry investment,⁸ with the hope that the resulting data will support the identification of new drug targets and the efficient delivery of stratified and personalised medicines (see Box 4).

ii The Office of Health Economics reports that the probability of a drug's success at Phase I, Phase II and Phase III to be between 49% and 75%, 30% and 48%, and 50% and 71% respectively⁷

Box 4: Investments by the global biopharmaceutical industry in health data resources

- ◆ In 2018, GSK invested in US company 23andMe and its database of genetic and phenotypic data.⁹
- ◆ In 2015, Roche acquired a majority stake in molecular and genomic analysis business Foundation Medicine,¹⁰ and took full ownership in June 2018.¹¹
- ◆ In 2012, Amgen acquired DeCODE Genetics which held genetic and clinical data on the Icelandic population. In June 2019, DeCODE announced a major collaboration with US-based healthcare delivery network Intermountain Healthcare, which aims to analyse the genomes of 500,000 people from Intermountain's patient population in Utah and Idaho.¹²

Given that well curated health data can support each step in the process of developing new medicines, and given the scale of health data generation and collection in the NHS, the UK should be well placed to attract a larger share of the annual \$179 billion⁶ global industry investment in clinical development and early research – contributing to the Government's overall ambition for R&D spending to reach 2.4% of GDP by 2028.¹³

However, despite the UK's promise, evidence suggests an increasing proportion of global R&D investment is being placed elsewhere. Although the UK retains a long-standing global reputation for high-quality research in both academic institutions and biopharmaceutical companies,¹⁴

- ◆ The National Institute for Health Research (NIHR) has warned that as expertise and capabilities emerge around the world – and in particular, in China and Brazil – biopharmaceutical companies are increasingly looking beyond traditional markets when making R&D investment decisions.¹⁴
- ◆ In each of the last three years the UK's share of patients recruited to global clinical studies has fallen year-on-year and is now just 3%¹⁵ – and just 2% in the early-stage Phase II trials, in which the UK should excel.¹⁶

The NIHR has noted that the UK's declining share is due to a number of factors, including the size of the potential patient pool and the burden of setting up research studies¹⁴ – which are both factors that can be addressed by unlocking the promise of UK health data.

Improving the UK's health data environment has therefore been a priority of the UK Government and the industry for many years. In 2013, for example, the ABPI developed its Big data road map,¹⁷ and in 2017 our members helped support the Life Sciences Industrial Strategy, which pressed for the development of platforms to enable health data to be appropriately shared for the research and development of new technologies.¹⁸

The following chapters explore our perspectives on this process, and what more needs to be done.

Industry perspectives on building trust and confidence in the use of health data

Health data is collected from patients and members of the public, and – in order for it to be used to support the development of new treatments – patients and members of the public must have trust and confidence that their personal information will be used only in the way that they choose, it is securely protected when it is used, and that its use will be carefully governed.

This trust and confidence is already high in the UK: members of the public report much greater confidence about their data being held by healthcare providers than in other countries,¹⁹ perhaps on account of the high level of trust in the publicly-funded NHS.²⁰

The biopharmaceutical industry has decades of experience in handling health data, collecting, managing, analysing and reporting on specified clinical data patient-by-patient, in the process of carrying out clinical trials.

The industry routinely provides this data to regulatory authorities as well as publishing it in aggregated, anonymised form. The biopharmaceutical industry is also increasingly making health data from trials available for others' research efforts.

For example, the open source platform, tranSMART²¹ originally developed in 2009 by scientists working in Johnson and Johnson's R&D division and Recombinant Data Corporation, is a repository for data that supports feasibility queries, exploration and analysis of clinical, translational and genomics data.

Since then, more than 100 other corporate, non-profit, academic, patient advocacy and government organisations have joined tranSMART, which is now overseen by a public-private partnership, the tranSMART Foundation.

More recently, in 2013 representatives of the global industry agreed and published Principles for Responsible Clinical Trials Data Sharing to support enhanced data-sharing with researchers whilst safeguarding the privacy of patients.²²

By virtue of this expertise and experience, the biopharmaceutical industry sees three prerequisites for generating and maintaining trust in a vibrant, well-supported health data research environment here in the UK.



In England, the national data opt-out rate is stable at less than

3%



1. Greater public understanding in how health data is used by biopharmaceutical companies

When people participate in clinical trials, they are aware that information on their experiences and outcomes is being used for the development of new medicines.

However, when information is collected routinely by the NHS in the course of patient care, there is a low level of awareness amongst the public about how this data could be used to help others.

Given that three-quarters of the public support the sharing of health data when it is used for medical research,²³ building greater public understanding will help build a more supportive UK health data environment.

The biopharmaceutical industry has taken a number of steps to help build public understanding, including through the ABPI's continuing commitment to the annual Pioneering Partnerships conference in collaboration with the Academy of Medical Research Charities and the NIHR,²⁴ and through the industry's membership of the European Commission's PARADIGM project – which is identifying the tools needed by both industry and patient and medical research charities to improve public and patient involvement in R&D.²⁵

Important steps have also been taken by others: the Wellcome Trust has established an independent patient data taskforce – 'Understanding Patient Data' – to provide evidence on the use of data for research, to help people understand data choices, and to champion the responsible use of data.

In addition, organisations like UseMyData work with patients' representatives to help build confidence and understanding amongst patients in how health data is used,²⁶ and their views are complemented by other organisations such as National Voices.²⁷

2. Robust protections for people's private information and how it is shared

The public is acutely aware of the risks associated with the electronic storage of personal data: the 18 largest data breaches since 2000 have included governments, healthcare providers, banks and tech companies.²⁸

Nevertheless, patients participating in biopharmaceutical companies' clinical trials can have confidence that their information will be protected, and this leads to high levels of explicit consent to the use of their data for the purposes of research.

Obtaining consent from people to use their routinely collected health data for research can be more challenging. Examples such as 'care.data' (a program to extract data from general practice to a central database) show what can happen when mechanisms to obtain consent for data-sharing are poorly designed and communicated.²⁹

A number of initiatives have been put in place to address these challenges. For example, the creation of the National Data Opt-Out has established a robust, legally-valid model of consent for the NHS to share patients' data for research purposes, and the opt-out rate has been stable at 2.7% for the last 10 months.³⁰

In addition, a new statutory 'National Data Guardian' has been appointed with a remit to help ensure that the public can trust that health and care information is securely safeguarded and used appropriately.³¹

New technologies are allowing these initiatives to be built on: for example, NHS Digital in England is using a de-identification product – Privitar Publisher – to enable the safer sharing and linkage of data between authorised parties.³²

3. Clear processes to share the benefits of research using health data

The benefits of access to and analyses of high-quality health data can support:

- Direct improvement in clinical decision-making, leading to better outcomes for patients.
- Refinements in patient pathways, resulting in efficiencies in NHS services; and
- Innovation at various stages in the process of discovering and developing new interventions.

The biopharmaceutical industry recognises that the consistent collection, storage, curation and management of routine health data require resources and incur a cost. Therefore, it is expected that contracts with custodians to enable access and analysis will reflect this.

Efficient access to the right data requires both a straightforward technical solution and, ideally, a simple standardised contracting process.

There needs to be clarity around the purpose for which users are requesting access to the data, the fees paid for access should contribute directly to the costs of collecting and managing the data, and there should also be recognition that the potential benefits of the results of the data analysis should be made available throughout the NHS – to improve patient outcomes, enable efficiencies and deliver innovative new interventions.

While data is being collected in increasingly large quantities, and there are already many different types of commercial access, there are public concerns over the sharing and use of data.³³

Nevertheless a study by IpsosMORI commissioned by the Wellcome Trust suggests that a majority of the public (61%) would support data-sharing to support research, rather than lose out on the benefits that research involving commercial organisations can bring.³²

There are examples of international best practice frameworks to guide how data is shared and used, such as the Global Alliance for Genomics and Health's Framework for Responsible Sharing of Genomic and Health-Related Data.³⁴

We and our members support – and have actively engaged in – work led by the Office for Life Sciences (OLS) to establish guiding principles to help ensure the NHS delivers benefits for patients and the public when health data is shared with researchers.

The latest iteration of these principles is set out in box 5, alongside our perspectives on how they might be further developed. Although these principles have been developed by institutions with jurisdiction over England, it is important that there is consistency in the application of these principles across the UK.

In summary, when arranging access to health data, it is important that the purpose is clear, that contracting arrangements are straightforward and standard, that technical access is simple and functional and that when benefits arise, they are made available across the health service.

Box 5: Guiding principles on the NHS's uses of health data – our perspectives

Principle (July 2019 iteration) ³⁵	Our perspectives
<p>Any use of NHS data, including operational data, not available in the public domain must have an explicit aim to improve the health, welfare and/or care of patients in the NHS, or the operation of the NHS. This may include the discovery of new treatments, diagnostics, and other scientific breakthroughs, as well as additional wider benefits.</p> <p>Where possible, the terms of any arrangements should include quantifiable and explicit benefits for patients which will be realised as part of the arrangement.</p>	<p>We strongly support this principle – given the biopharmaceutical industry exists to improve the health, welfare and/or care of patients – and aim to work with the NHS to design ways of quantifying the benefits for patients that the analysis and use of health data can deliver through the development of new medicines.</p>
<p>NHS data is an important resource and NHS organisations entering into arrangements involving their data, individually or as a consortium, should ensure they agree fair terms for their organisation and for the NHS as a whole. In particular, the boards of NHS organisations should consider themselves ultimately responsible for ensuring that any arrangements entered into by their organisation are fair, including recognising and safeguarding the value of the data that is shared and the resources which are generated as a result of the arrangement.</p>	<p>We recognise the good intent of this principle and support the concept of fair terms. However, there is no definition of 'fair', and there is a risk that this principle could be interpreted to mean that every NHS organisation will need to set up data commercialisation resources with specialist experience and expertise, and therefore that the UK data landscape could become more fragmented.</p>
<p>Any arrangements agreed by NHS organisations should not undermine, inhibit or impact the ability of the NHS, at national level, to maximise the value or use of NHS data. NHS organisations should not enter into exclusive arrangements for raw data held by the NHS, nor include conditions limiting any benefits from being applied at a national level, nor undermine the wider NHS digital architecture, including the free flow of data within health and care, open standards and interoperability.</p>	<p>We support this principle and would like to see it more patient-centred. Through further discussion with patient groups, we hope the principle can be further refined to reflect the fact that the sources of data are patients themselves, and therefore to reflect the importance to many patients that no party should lay claim to the exclusive use of data for research.</p>
<p>Any arrangements agreed by NHS organisations should be transparent and clearly communicated in order to support public trust and confidence in the NHS and wider government data policies.</p>	<p>We support this principle but note practical experience that the transparency of agreements reached to date has varied across NHS organisations. Central guidance informed by patients on what 'transparency' means in practice is needed.</p>
<p>Any arrangements agreed by NHS organisations should fully adhere to all applicable national level legal, regulatory, privacy and security obligations, including in respect of the National Data Guardian's Data Security Standards, the General Data Protection Regulation (GDPR) and the Common Law Duty of Confidentiality.</p>	<p>We support this principle, and – given our members' experience of both national and international obligations and data platforms and systems – commit to observe it and to support other stakeholders in the UK health data landscape in doing so.</p>

Harnessing the opportunities of UK health data

The biopharmaceutical industry engages in a variety of ways with UK health data, some examples of which are set out in box 6, and works with a number of different datasets, of which some of the larger ones are set out in box 7.

In building opportunities to support more health research, HDR UK has identified three dimensions in which health datasets can be strengthened:³⁶

- **Breadth/scale:** the populations that health data cover must be sufficiently wide to enable and support clinical research.
- **Depth:** within each health dataset, the type and volume of information consistently captured about patients must become more diverse and varied.
- **Follow-up duration:** the datasets must routinely collect follow-up data wherever possible, so that changes in patient outcomes over time – and what might have led to them – can be investigated.

The UK's data landscape is illustrated in the diagram (Figure 2). It is worth noting that different research projects at different stages of the new medicine development process require different mixes of the parameters of breadth, depth and follow-up duration.

We welcome the initiatives that are under way to make progress towards HDR UK's ambitions. For example, the UK Government is building on its world-leading 100,000 Genomes Project with ambitions to sequence five million genomes within five years³⁷ – supported by the multi-million-pound contribution that industry is making to UK Biobank.

Northern Ireland, Scotland and Wales have all contributed to the 100,000 Genomes Project³⁸ and the devolved nations maintain their role in the UK-wide ambitions.

In addition to improving the quantity and quality of UK health data in these ways, however, steps must also be taken to improve its accessibility.

As the Life Sciences Industrial Strategy noted, and formal feedback from the biopharmaceutical industry has made clear (see figure 3 and box 8), although the UK possesses a number of data sources which offer significant potential to researchers, the UK does not offer the deep, near-real-time access to data across multiple care settings which would allow the UK to offer health data resources comparable to the best in the world, such as Flatiron (see box 9).¹³

Furthermore, steps need to be taken to facilitate the use of health data captured during routine clinical care by bodies such as NICE, to enable better information on the potential value of a new medicine in the NHS.

We see five characteristics of the health data landscape that need to be addressed to unlock the opportunity that exists, set out below – alongside corresponding priority action areas, where industry can work with Government to deliver the required improvements.

11 million

**patients are registered
on the Clinical Practice
Research Datalink**

Box 6: Examples of the ways in which biopharmaceutical companies use health data for research

- The Salford Lung Study was a community-based, real-world Phase III randomised controlled trial (RCT) for a new treatment for COPD and asthma, sponsored by GSK. The RCT made use of electronic patient records, which allowed patients to be monitored during 'normal' clinical practice in near-real-time but with much less intrusion into their lives than typical RCTs.¹³ The Salford Lung Study shows the potential of establishing virtual clinical trials using the UK's health data resources.
- Research undertaken by biopharmaceutical companies BioMarin and Alexion using health data gathered through the 100,000 Genomes Project has helped researchers better understand the clinical spectrum of symptoms that people living with rare genetic diseases show – and has also helped diagnose patients unknowingly living with rare genetic disorders.³⁹
- The BSRBR-RA study is a unique collaboration between the University of Manchester, the British Society for Rheumatology and the biopharmaceutical industry. It tracks the progress of over 20,000 people with rheumatoid arthritis (RA) who have been prescribed biologic medicines (including biosimilars) and other targeted therapies.⁴⁰
- AstraZeneca is working with NHS Scotland as part of its Global Genomics Initiative to make use of patients' genetic information to develop new treatments.⁴¹

Box 7: Examples of the UK's larger health datasets

- The Clinical Practice Research Datalink (CPRD) collects data on patients from a network of GP practices across the UK (including 11 million currently registered patients⁴²).
- Wales's Secure Anonymised Data Linkage (SAIL) Databank holds a wide range of de-identified health and care datasets, from primary care to outpatient data, which can be linked and accessed via a remote gateway for approved research projects.¹³
- The 100,000 Genomes Project combines whole genome sequencing data with medical records from around 85,000 people.⁴³
- England's Hospital Episode Statistics (HES) capture a wide range of clinical information on around 20 million patients admitted to hospital a year.⁴⁴
- The UK Biobank has been collecting increasingly detailed data on 500,000 people since 2006.⁴⁵
- England's National Cancer Registration and Analytics Service (NCRAS) collects data on all cases of cancer that occur in people living in England.⁴⁶
- The Scottish Cancer Registry has been collecting population-based information on cancer since 1958 and now holds over 1.8 million records.⁴⁷
- The National Institute for Cardiovascular Outcomes Research (NICOR) collects clinical data on cardiovascular patients across the UK. It oversees the National Cardiac Audit Programme, which had over 380,000 patient records entered in 2016-17.⁴⁸
- The Systemic Anti-Cancer Therapy (SACT) dataset has been collecting data on the use of systemic anti-cancer therapies across all NHS trusts in England since 2012.⁴⁹

Figure 2. Examples of the three key parameters of UK health datasets – breadth/scale, depth and duration of follow-up

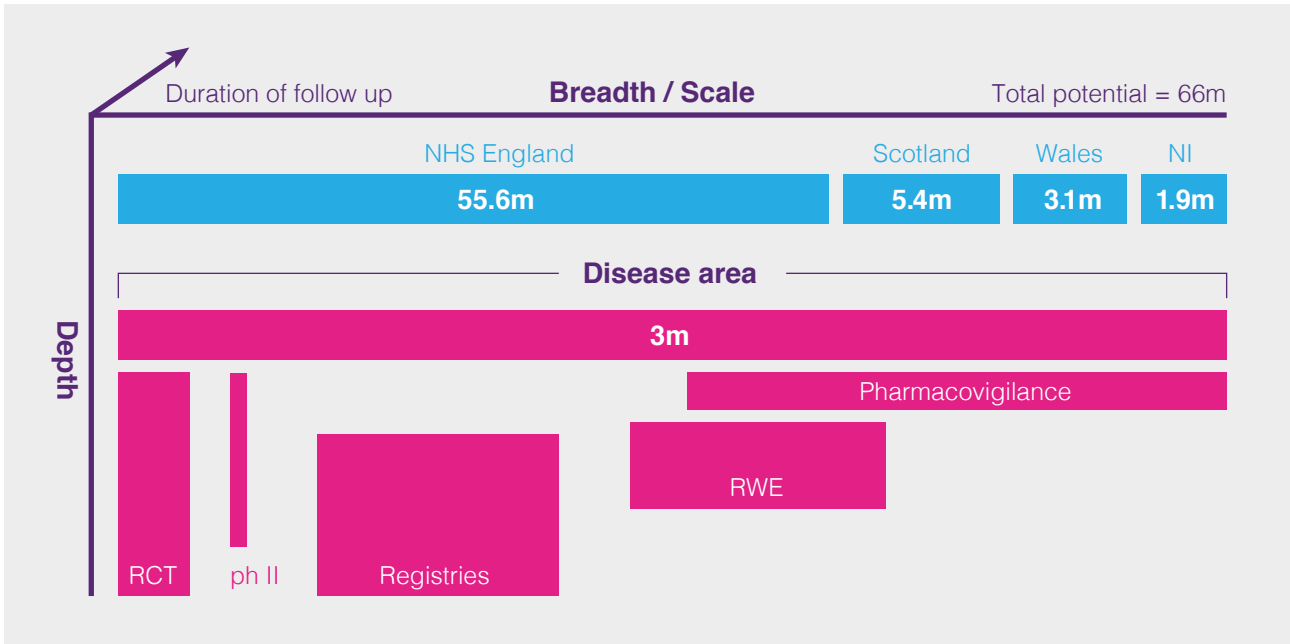
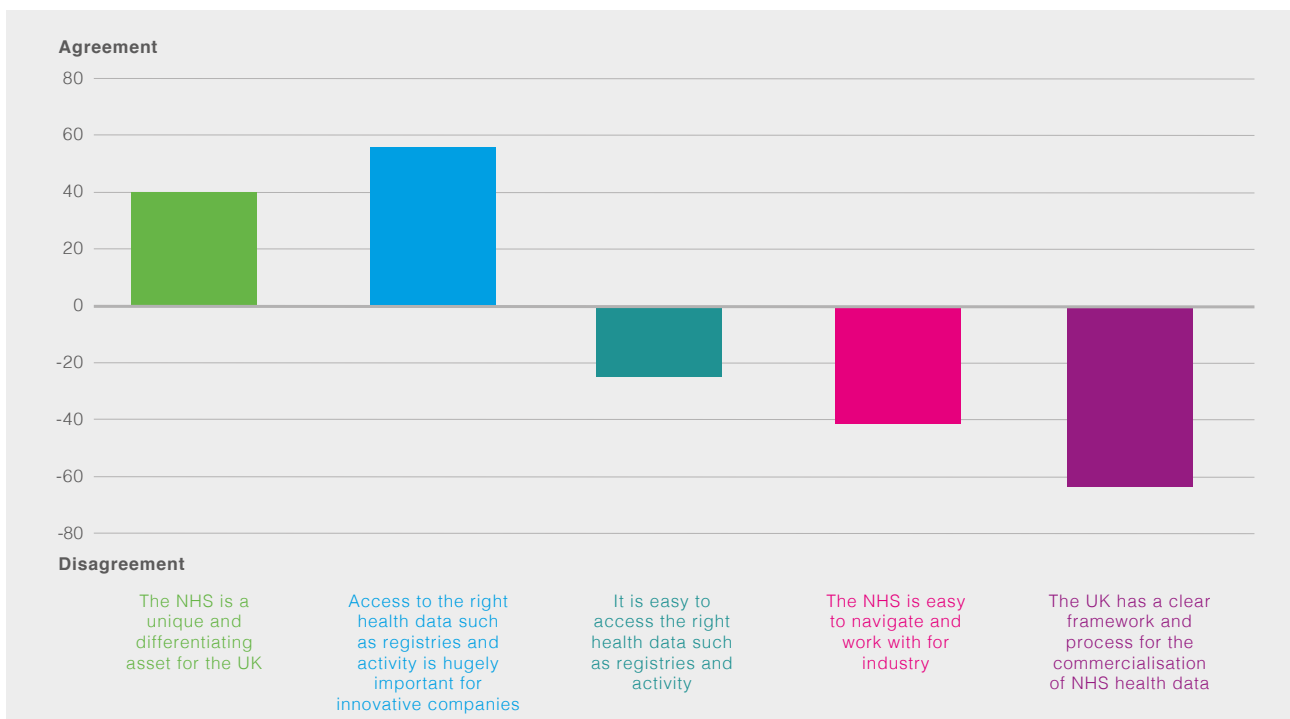


Figure 3. The views of small and medium-sized companies on the NHS data landscape⁵⁰

Respondents were asked: Please could you indicate how much you agree/disagree with these statements about the NHS (net agreement percentages).



Box 8: Feedback collected from the biopharmaceutical industry by Health Data Research UK⁵¹

An engagement process with biopharmaceutical industry representatives led by HDR UK in 2019, in order to inform the specification of the Digital Innovation Hubs (DIHs) programme, collected the following feedback on the UK health data landscape:

- Time delays and unpredictability prevent UK data access for many companies: their priorities are to see transparent, predictable, quick access to data.
- Companies most frequently request health data that can support trial recruitment, help demonstrate value, and understand and stratify disease.
- Companies value health data services that assist with health data discovery, offer quick and predictable access to health data once discovered, provide data curation, and are underpinned by pre-approved contracts and models.
- Gaps in the UK's health data that companies want to see addressed are: direct linkage to secondary care data to understand treatment effectiveness in detail; quick assessments of patients presenting in each site for trial feasibility; and the ability to recruit patients in real time based on automated eligibility checks.

Box 9: Flatiron⁵²

- For maximum utility, cancer datasets need to capture each patient's stage at diagnosis, every treatment cycle (including the specific treatments delivered) and each patient's responses and outcomes. Few health datasets anywhere in the world capture this kind of detail.
- US company Flatiron created a unique dataset of around two million patients with cancer, which was bought by biopharmaceutical company Roche in 2018.
- Flatiron's value was generated not by the sheer volume of information in its database, but instead by the way in which each entry in its database was meticulously curated to develop a clinical research-grade dataset, in an enormously labour-intensive process.⁵²

1. The data landscape is fragmented

The NHS is seen by many as a single, national organisation – but in practice this is not the case for health data.

Health data is collected by a large number of different organisations, including over 200 legally separate NHS trusts across England, autonomous research organisations (including UK Biobank), government-owned companies (such as Genomics England), disease registries (collecting information on, for example, patients with cancer, heart disease and diabetes) and national-level bodies (such as, in England, NHS Digital).

The devolved nature of the UK health system complicates this picture further, with each devolved nation maintaining separate legal and data governance structures for their health systems. They can decide their own approaches to promoting their data and responding to requests from researchers, and the risk is that – as stated above – implementation of the OLS's second guiding principle on health data exacerbates rather than alleviates this situation.

**There are
more than
200
legally separate
data custodians
across England**

The following anecdotal examples illustrate this fragmentation:⁵³

- The UK might have been able to take part in a real-world study in type 2 diabetes (sponsored by a global company), but the available datasets were much smaller than those accessible through US administrative claims databases, which were able to provide data on around 700,000 patients.
- The UK had the opportunity to take part in a multi-country real-world study in inflammatory bowel disease (also sponsored by a global company). However, lack of central data access across the NHS in England meant there was information on too few patients per centre for the UK to be readily included.

We recognise that important steps are being taken to address this fragmentation: for example, the UK Health Data Research Alliance was launched in February 2019 to bring together the many organisations in the UK which hold health data – including academic institutions, NHS England, NHS Scotland, NHS Wales, NHS Digital, Genomics England, Public Health England, CPRD and others⁵⁴ – and to bring about a consistent approach to data access for research.

However, membership of the Alliance does not require specific organisational commitments to data access, and it is not comprehensive: with 18 members as of December 2019, its membership does not include many relevant UK data custodians, including many disease registries and the majority of NHS organisations.⁵⁴

The first priority action area is to address fragmentation.

Shared aim 1: Address the fragmentation in the UK health data landscape; create linkages to enhance scale and depth

Proposed government action	Industry action
<p>Establish (or nominate) a single NHS data access organisation (DAO) which can act on behalf of the whole NHS as a single counterparty to data access agreements. For example, the Health Data Research Innovation Gateway could act as the single NHS DAO.</p>	<p>Pay reasonable costs towards the running of the DAO as part of the data access fee.</p>
<p>Encourage membership of the UK Health Data Research Alliance, and work with HDR UK to standardise access and curation processes within the Alliance.</p>	<p>Provide resources for the co-production of standard commercial models which can be used by NHS organisations.</p>



2. Processes to access data are inefficient

As boxes 10 and 11 make clear, the biopharmaceutical industry encounters processes to access data in the UK which can be slow, bureaucratic and unpredictable, which require multiple applications and agreements, and which are guided by data controllers who are under pressure to be risk averse.

These processes can take months and sometimes years to respond to requests from researchers for health data (see box 10), frustrating the development both of new treatments and new technologies (see box 11). Slow and inefficient processes also impede steps to improve the quality of the UK’s health data, since the longer the time between generating and accessing data (known as ‘latency’), the more difficult it is to make corrections while the clinical team is able to recall the case.

Where processes are more efficient elsewhere in the world, some elements of global new medicine development will be done there rather than in the UK. However, significant improvement could ensure that the UK can attract more R&D investment.

We hope that the Government’s plans to create a new ‘National Centre of Expertise’ in NHSX, to provide commercial and legal expertise to NHS organisations, and tools such as standard contracts and guidance, will help the NHS respond more efficiently to requests for data from researchers.³⁵ In addition, we will continue to support efforts to improve digital and data-handling skills in the NHS as recommended in Dr Eric Topol’s 2019 review for Health Education England.⁵⁵

Thus, the second priority action area is:

Shared aim 2: Increase the efficiency of the UK’s health data access processes

Proposed government action	Industry action
Task the DAO with administering quick and predictable access processes.	Provide resources for the co-production of standard legal agreements to help facilitate commercial access to NHS data on fair terms, following the model of HIPAA BAAs. ⁱⁱⁱ
Establish NHS’s Centre of Expertise in NHSX as rapidly as possible, and make standard contracting for data access simple and quick.	Jointly fund training for relevant industry and NHS colleagues on data governance.

ⁱⁱⁱHIPAA is the Health Insurance Portability and Accountability Act in the USA. It governs healthcare data privacy. The HIPAA defines a business associate as any organisation working with or providing services to health insurers who process personal data. Under Business Associate Agreements (BAAs), certain requirements need to be met concerning the use of the health data and the safeguards which need to be put in place.

Box 10: Examples of inefficient data access processes⁵⁶

- Delays in HES linkage to clinical data for a rare disease specialist centre, requiring further amendments to Confidentiality Advisory Group (CAG) and Health Research Authority approvals, led to an 18-month delay in a project for direct care supporting better disease detection and referrals.⁵⁷
- A global contract research organisation (CRO) reported that it had agreed and executed a data access arrangement in one EU country in eight months, but that the equivalent access arrangement in the UK was still under discussion in the UK two years after the CRO had first sought the data.
- In 2018, a UK SME looking for linked genetic and clinical data to validate a suspected target association and raise funds to develop a new drug found a relevant dataset within two weeks but, after an unexplained delay of three months while the university concerned started the contracting process, had to give up working with that dataset.
- A global company wanted to access national data on outcomes related to current treatment pathways to support a submission to NICE but found there was no way to access data across the country. After discussions with a number of trusts, this eventually resulted in the company conducting a single-centre audit which itself took six months to complete.

Box 11: AI and the need for access to high-quality data

- There is much interest in the promise of AI to improve healthcare decision-taking and improve efficiency.⁵⁸ On 8 August 2019, for example, the Prime Minister announced £250 million of investment to help the NHS become a world leader in its use.⁵⁹
- However, AI tools require access to high-quality data to learn from, and companies investing in AI therefore invest significantly in accessing and improving data – for example:
- A joint report by the Medicines Discovery Catapult and the BioIndustry Association found that 75% of spending by companies in AI is actually on the upstream (often unseen) activities of data access, curation and data labelling, and not algorithm development and improvement.⁵⁰
- IBM has also reported around 80 per cent of the time spent by scientists developing AI technologies is spent finding, cleansing and organising data – rather than in developing the algorithms which actually perform any analysis.⁶⁰

If the NHS's £250 million investment in AI is appropriately allocated, therefore, at least £185 million of the investment may need to be spent on accessing and improving data.

3. The NHS is not sufficiently digitised to allow data to be linked and accessed readily

A particular issue for clinicians and researchers seeking to work with UK health data is that much of the data is unlinked. For example, one global biopharmaceutical company seeking health data to support its cancer research found that the UK's SACT and cancer registry datasets could not help because it was not possible to follow a patient between primary and secondary care.⁶¹

Although commitments have been made to a 'paperless NHS' in England – which would allow data on an individual patient to be linked seamlessly across care settings – the timetable for delivery has repeatedly slipped: the NHS in England was first challenged to 'go paperless' by 2018,⁶² and then by 2020;⁶³ and the NHS Long Term Plan's current target is 2024.⁴

We hope that, in time, the NHS's digital infrastructure across the UK will be sufficiently mature to allow the easy use of health data for the purposes identified in box 11 – for example, trial recruitment, demonstrating value, and understanding and stratifying disease. In relation to the former, this will – for example – create the conditions to allow further trials of the kind exemplified by the Salford Lung Study to be located in the UK. However, in order for the UK to secure such clinical trials investment, researchers must have the confidence that the UK's digital infrastructure will be able to identify patients who might benefit from a treatment under development in near-real time, secure their consent for

participation, enroll them into a trial, and report on results to the standards that traditional clinical trials would offer.

We recognise that the UK is taking important and positive steps to help digitise and link the UK's health datasets, and achieve this goal:

- ▶ NHS England is supporting 'global digital exemplars (GDEs)' in secondary care and 'local health and care record exemplars (LHCREs)' to join together health and care records to test the most efficient approaches to achieving its aim of a 'paperless' NHS.⁴
- ▶ HDR UK's 'Sprint Exemplar Projects' aim to test technologies and methodologies to enable the utilisation of linked datasets,⁶⁴ and include a project led by the University of Leicester to make patient datasets safely linkable and discoverable so that a complete patient profile can be readily located.⁶⁵
- ▶ Seven 'Health Data Research Hubs' were established in September 2019, and offer the prospect of creating rich disease-focused datasets that will enable new clinical trials and real-world evidence studies to be undertaken in the UK.⁶⁶

Whilst we recognise the progress that each of these initiatives will deliver, it is important that timetables for delivery do not slip, and that new initiatives are not created on top of older initiatives, which would further fragment the health data landscape.

While industry continues to support the digitisation of the NHS, the issues of linkage and accessibility are addressed through the first two action areas. The third priority action area for industry is specifically to harness health data to support clinical development of new medicines:

The UK's data access processes need to be **more efficient**



Shared aim 3: Harness UK health data specifically to support the efficient design, feasibility, recruitment and conduct of the full range of clinical trials (from Phase II through to real-world studies)

Proposed government action	Industry action
Support the development of timely access to health data which helps the industry rapidly explore the feasibility of conducting clinical trials in the UK.	Invest in more commercial clinical trials, and a greater share of patients in multi-centre trials in the UK.
Develop processes to transition rapidly from feasibility to recruitment of patients.	Work with the NHS, NIHR and Health Data Research Hubs to enable and co-fund near-real-time recruitment processes in key therapeutic areas aligned with the NHS Long Term Plan.
Develop processes to encourage patients to be given the opportunity to participate in clinical trials using their routinely collected health data.	

4. The quality and accessibility of the UK's health data resources is opaque

Researchers engaged in biopharmaceutical R&D frequently struggle to understand the quality and accessibility of the UK's health data, encountering particular challenges in answering the following common questions:

- Consent: does the patient consent cover the intended use of the health data?
- Permissions: am I legally able to use the dataset in the way I intend?
- Cost/terms: what are the costs or terms of accessing the dataset?
- Time: how long will it take me to get permission and then practically access the dataset?

As a result, researchers look elsewhere in the world for health data to meet their research needs, rather than work in a UK health data environment where the discovery of relevant health datasets and information about their quality is a challenge (see box 12).

We are nonetheless encouraged by a number of initiatives:

- In Scotland, the electronic Data Research and Innovation Service (eDRIS) offers an effective, single point of contact to assist researchers with their data access questions.⁶⁷
- The Health Data Research Innovation Gateway has been established in January 2020, with the aim of allowing data from the Health Data Research Alliance members to be discovered quickly.⁶⁸
- The new National Centre of Expertise being established in NHSX is due to set clear and robust standards on transparency and reporting.³⁵

Therefore, the fourth priority action area is to make it easy for researchers to find out what datasets are available, and what the quality and utility of each dataset is:

Shared aim 4: Enhance the transparency of the quality and accessibility of the UK's health data resources

Proposed government action	Industry action
Monitor and report on data applications, access and turnaround times (TATs).	Capture and collate user experience and TAT from biopharmaceutical companies, through the ABPI.
Ensure the Health Data Research Innovation Gateway's metadata provides information on the quality and accessibility of the data in available datasets (for example through a directory).	Fund and deliver consent codification for priority datasets.

Box 12: Examples of the challenges in discovering UK health data⁶⁹

- In 2017, a global biopharmaceutical company looking for cancer-related health data was incorrectly informed that comprehensive, national health data was available. However, after six months it became clear that the data available was incomplete and low-quality, particularly regarding prescription data. The delay in accessing the data meant that it was nearly impossible to have the quality issues addressed.
- In 2019, a global CRO requested data on the number of specific patients attending UK hospitals so that the UK could be included as a potential site for a global clinical trial. However, the data took so long to arrive that the UK was not included as a possible location.



5. Regulatory and reimbursement processes do not recognise the potential that developments in health data can offer

Many of the challenges that the biopharmaceutical industry encounters in securing approval of new medicines for use in healthcare systems – including the NHS – relate to the methods used by regulatory and reimbursement authorities to assess them. For example:

- The clinical endpoints required by regulatory authorities are naturally based upon historical experience in clinical trials of medicines with established measures of efficacy.

As innovation proceeds, new mechanisms with new measures and endpoints are discovered and novel designs for clinical trials are developed to assess these. The adoption of these new approaches in the regulatory process needs to be accelerated through early dialogue.
- NICE's methods of evaluation rely mainly on the use of research findings in academic journals (or syntheses of these research findings), despite the growing possibility of using broader types of data – including data captured from the NHS during routine clinical care – to inform its assessments. This gap is itself recognised by NICE, which launched a consultation on making use of broader types of data in June 2019.⁷⁰
- Current, international approaches to the regulation of clinical trials are based on standards agreed by the International Council for Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) in 1995, before health data healthcare institutions became digitised.¹³

These characteristics pose particular challenges for the new categories of treatments which the industry is researching and developing. For example:

- The efficacy of a variety of new treatments for cancer may need to be tested on multiple tumour sites, leading to a need for novel trial designs (for example 'basket' or 'umbrella' trials).
- The need to address smaller patient populations due to biomarker-led stratification, personalised treatments or rare diseases also requires novel designs of clinical trials.
- Clear demonstration of the long-term value of emerging curative technologies such as cell and gene therapies is challenging and may require new approaches including long-term follow-up.

We therefore hope to see that the advances being made in biopharmaceutical R&D and healthcare systems' digital capabilities are accompanied by advances in the approaches taken by regulators and reimbursement authorities.

This will enable the use of data captured digitally, allowing the use of real-world evidence to support the approval and adoption of new treatments. As the Life Sciences Industrial Strategy made clear, this will improve the speed and efficiency of regulatory studies, increase the cost effectiveness of trials, and reduce the cost of developing medicines.¹³

The UK share of global biopharmaceutical R&D spending is, at 7%, relatively small⁷¹ – but the UK has an excellent reputation for regulation and health economic assessment. The UK could lead the way in the development of a regulatory and reimbursement environment which harnesses the potential improvements that health data can deliver – as the MHRA and NICE are already considering.⁷²

The rapid evolution of clinical trial designs and endpoints, coupled with the potential for health data to be accessed and analysed in near-real time, leads to the suggestion that regulatory and health

technology assessment authorities will in future have a wider variety of data to consider – some of it eventually enabling more precise estimation of both the clinical and economic value of an intervention.

At the same time, novel treatments that promise to provide long-term remission or even cure after a single (or only a few) administrations will need specific programmes of long-term data collection,

to demonstrate or confirm their value and perhaps to underpin new approaches to reimbursement (for example outcomes-based payments).

Therefore, industry’s fifth priority area is to work with the authorities to understand new types of health data, how they can be factored into evaluations, and how standards can be developed.

Shared aim 5: Broaden the data considered to help demonstrate value

Proposed government action	Industry action
Task the MHRA, NICE and the NHS with developing clear guidance on the increasing variety of data they use to support their regulatory, value assessment and reimbursement processes.	Work with the MHRA, NICE and the NHS to embed new UK data standards for approval, valuation and reimbursement processes.
Work towards developing standards for different types and sizes of patient population, and a wider variety of outcome measures.	Support the dialogue about new approaches to payment models and the generation of data that can underpin them.
Lead the dialogue about new approaches to payment contracts and the data required to support them.	



The quality and accessibility of UK health data resources must be **more transparent.**

Enhancing the UK health data environment

The UK health data environment offers much promise to both UK-based and global researchers. For the reasons set out above, however, the UK's promise does not yet translate into reality, and the current value of the UK's health data tends to be overestimated by policymakers.

Indeed, the difficulties of accessing UK health data have led to the formation of specialist private companies – such as Sensyne Health (see box 13).

The Government has recognised the need to improve the system. The importance of the sector to the UK was captured in the Life Sciences Industrial Strategy and subsequent sector deals,⁷³ and associated funding has enabled the establishment of HDR UK and the triple structure of the Health Data Research Alliance, Health Data Research Hubs and the infrastructure for the Health Data Research Innovation Gateway (see box 14).

The biopharmaceutical industry leads the global undertaking of discovering new medicines, developing them and delivering them to patients.

Through our members in the UK, we support the ambition of the Government to increase the proportion of R&D spending towards 2.4% of GDP.

Success in unlocking the promise of UK health data can help attract more of the global R&D effort to the UK, particularly in the area of clinical development: investment in clinical trials is already estimated to bring £192 million worth of value to patients within the NHS,⁷⁴ as well as providing valuable experience for the healthcare professionals involved.

This will help to ensure that the NHS is ready to make the most of new medicines as and when they become available.

In summary, this report sets out five priority action areas where shared aims should lead to further collaboration with government agencies, perhaps under memoranda of understanding or sector-deal type projects in order to maintain progress in improving the management and accessibility of UK health data:

- 1** Reduce fragmentation in the UK health data landscape
- 2** Increase efficiency of data access processes
- 3** Enable design, feasibility, recruitment and conduct of all clinical trials
- 4** Enhance data transparency
- 5** Harness health data to demonstrate the value of interventions

While focusing on these areas, industry is also committed to work with Government and all stakeholders to ensure that consistently high standards of governance are established and maintained. Looking to the future, it will be also be important to work to ensure the availability of a data science equipped workforce, developing and attracting the right skills into the life sciences sector.

Together, we can unlock the promise of UK health data to support research that will improve patient outcomes, enhance the efficiency of the NHS and support the development of new medicines.

Box 13: Sensyne Health

- ◆ Sensyne Health is a unique partnership (initiated in February 2017) with a small number of NHS trusts which develops digital health products and enables companies to analyse anonymised data. The data remains in NHS ownership.⁷⁵
- ◆ Each analysis of anonymised patient data is pre-approved for each programme on a case-by-case basis by the relevant NHS trusts. This is to ensure that the purpose of the anonymisation and the proposed analysis are subject to appropriate ethical oversight and information governance, including conformance with NHS principles, UK data protection law and applicable regulatory guidance.⁷⁶
- ◆ In August 2018 Sensyne Health floated on the London Stock Exchange.⁷⁵

Box 14: HDR UK and the components of data access

1 The UK Health Data Research Alliance

The Alliance was established in December 2018 to bring together leading healthcare and research organisations and health leaders to establish best practice for the ethical use of UK health data at scale.

2 Seven Health Data Research Hubs

In September 2019, the following Hubs were announced:

3 UK Health Data Research Innovation Gateway

The purpose of the Innovation Gateway is to provide services to the Hubs, and others, so that data from the Alliance can be discovered and accessed safely and responsibly. The first phase went live in January 2020, includes a catalogue of metadata to facilitate discovery of relevant datasets.

Name	Focus	Aim
DATA CAN	Cancer	Enable UK-wide high-quality cancer data access to improve care, diagnosis and research.
INSIGHT	Eye health	Use data, analytics and AI to develop insights into eye disease and wider health.
Gut Reaction	Inflammatory bowel disease	Use data to better stratify Crohn's Disease and ulcerative colitis patient responses.
PIONEER	Acute care	Use linked data to enable companies to develop acute care products and services.
NHS Digital	Clinical trials	Increase opportunities for patients to participate in clinical trials.
BREATHE	Respiratory	Improve the lives of people with conditions such as asthma and COPD.
Discover-NOW	Real-world data	Understand, develop treatments and prevent long-term conditions such as T2 diabetes.

References

1. Nuffield Trust, Will technology save the NHS £10 billion?, 30 June 2015
2. Medicines Discovery Catapult and ABPI, Use of health data by the life sciences industry: a UK perspective, October 2019
3. NHS Confederation, NHS statistics, facts and figures, accessed 2 July 2019
4. NHS England, NHS Long Term Plan, January 2019
5. European Medicines Agency, Human medicines: highlights of 2018, 4 January 2019
6. EvaluatePharma, World Preview 2019, Outlook to 2024, June 2019
7. OHE, The R&D cost of a new medicine, December 2012
8. EY, Realising the value of health care data: a framework for the future, July 2019, unless otherwise stated
9. GSK, GSK and 23andMe sign agreement to leverage genetic insights for the development of novel medicines, 25 July 2018
10. Reuters, Roche in \$1 billion-plus Foundation Medicine deal to boost cancer efforts, 12 January 2015
11. Roche, Media Release, 19 June 2018
12. Amgen, Intermountain Healthcare and deCODE genetics Launch Groundbreaking DNA Study of 500,000 People to Find New Links Between Genetics and Disease, 11 June 2019
13. Life Sciences Industrial Strategy, 30 August 2017
14. National Institute for Health Research, Impact and value report, 12 July 2019
15. Office for Life Sciences, Life Science Competitiveness Indicators, June 2019
16. NESTA, All together now: Improving cross-sector collaboration in the UK biomedical industry, March 2011
17. ABPI, Big data road map, 21 November 2013
18. Life Sciences Industrial Strategy, 30 August 2017
19. OpenDataInstitute, Who do we trust with personal data?, 5 July 2018
20. HRA, Sharing anonymised patient-level data where there is a mixed public and private level benefit – a new report, 1 September 2019
21. tranSMART Foundation, Our history; accessed 29 July 2019
22. EFPIA and PhRMA, Principles for Responsible Clinical Trial Data Sharing Our Commitment to Patients and Researchers, July 2013
23. Understanding Patient Data, Public attitudes to patient data use: a summary of existing research, July 2018
24. AMRC, Pioneering Partnerships 2019: Putting Patients First; accessed 29 July 2019
25. ABPI, Working with patients and patient organisations, June 2019
26. See for example UseMyData's work in Understanding Patient Data, Guest blog: Patient data saves lives, let's acknowledge it, 18 September 2017
27. National Voices, About us, accessed 24 September 2019
28. CSO Online, The 18 biggest data breaches of the 21st century, 20 December 2018
29. BBC, Care.data: How did it go so wrong?, 19 February 2014
30. NHS Digital, National Data Opt Out – March 2019, 19 March 2019
31. DHSC, Dame Fiona Caldicott appointed as the first statutory National Data Guardian for Health and Social Care, 11 March 2019
32. The Royal Society, Protecting privacy in practice: the current use, development and limits of Privacy Enhancing Technologies in data analysis, 26 March 2019
33. IpsosMORI and the Wellcome Trust, The One-Way Mirror: Public attitudes to commercial access to health data, March 2016
34. HDR UK, Principles for participation, May 2019
35. Department of Health and Social Care, Creating the right framework to realise the benefits for patients and the NHS where data underpins innovation, 15 July 2019
36. HDR UK, One Institute Strategy 2019-20, 9 April 2019
37. Genomics England, Secretary of State for Health and Social Care announces ambition to sequence 5 million genomes within five years, 2 October 2018
38. Genomics England, Wales joins the 100,000 Genomes Project, 16 February 2018
39. Genomics England, Industry collaboration already benefiting participants of 100,000 Genomes Project, 17 April 2019
40. University of Manchester, BSRBR-RA study, Monitoring the long term safety of drugs to treat rheumatoid arthritis, accessed 3 July 2019
41. University of Glasgow, Scotland and UOFG at forefront of global genomics initiative, 26 October 2016
42. Clinical Practice Research Datalink, accessed 2 July 2019

43. Genomics England, The 100,000 Genomes Project, accessed 19 July 2019
44. NHS Digital, Hospital Episode Statistics; accessed 29 July 2019
45. Biobank, About UK Biobank, accessed 2 July 2019
46. NCRAS, About the National Cancer Registration and Analytics Service; accessed 19 July 2019
47. ScotPHO, Scottish Cancer Registry, accessed 30 August 2019
48. NICOR, National Cardiac Audit Programme Report 2016/17, accessed 30 August 2019
49. SACT, About SACT, accessed 30 August 2019
50. Medicines Discovery Catapult and BioIndustry Association, State of the Discovery Nation, 2019
51. Information shared with the ABPI from Health Data Research UK
52. Forbes, The Deeply Human Core Of Roche's \$2.1 Billion Tech Acquisition -- And Why It Made It, 18 February 2018
53. Case studies provided in confidence to the ABPI
54. HDR UK, UK Health Data Research Alliance, accessed 6 December 2019
55. Health Education England, Preparing the healthcare workforce to deliver the digital future, February 2019
56. Data provided in confidence to the ABPI, unless otherwise stated
57. IMS case study, referenced in Life Sciences Industrial Strategy, 30 August 2017
58. NHS England, NHS aims to be a world leader in artificial intelligence and machine learning within 5 years, 5 June 2019
59. DHSC, Health Secretary announces £250 million investment in artificial intelligence, 8 August 2019
60. IBM, Breaking the 80/20 rule: How data catalogs transform data scientists' productivity, 23 August 2017
61. Data provided in confidence to the ABPI
62. Data provided in confidence to the ABPI DHSC, NHS challenged to go paperless by 2018, 16 January 2013
63. NHS Digital, National Information Board: Paperless 2020, accessed 2 July 2019
64. HDR UK, Sprint Exemplar Projects; accessed 5 August 2019
65. University of Leicester, Innovative healthcare project to link patient data receives government funding, 4 February 2019
66. HDR UK, Digital Innovation Hub Programme prospectus guidance notes for applicants, May 2019
67. ISD Scotland, eDRIS: FAQs; accessed 31 July 2019
68. HDR UK, Gateway; accessed 5 August 2019
69. Case studies provided in confidence to the ABPI
70. As stated in NICE's 'Statement of Intent'; accessed through NICE, Consultation on the data and analytics statement of intent, 13 June 2019
71. ABPI, Worldwide pharmaceutical company R&D expenditure by country; accessed 9 August 2019
72. PharmaTimes, NICE explores extending use of real world data to inform guidance, 27 June 2019
73. HM Government, Life Sciences Sector Deal 2, 2018
74. 2014/15 figures; cited in section 6, National Institute for Health Research, Impact and value report, 12 July 2019
75. Financial Times, Drayson floats medical AI group Sensyne Health on Aim, 14 August 2018
76. Sensyne Health, Our company; accessed 9 August 2019

The Association of the British Pharmaceutical Industry

A company limited by guarantee registered
in England & Wales number 09826787

Registered office 7th Floor, Southside,
105 Victoria Street, London SW1E 6QT

RMI-0130-0120

Design by tothepoint