Use of health data by the life sciences industry

a UK perspective





Foreword

Ethical, consented and trusted use of health data by the life sciences industry is now a critical part of the research and development of new generation precision medicines. Through key institutions such as Health Data Research UK (HDR-UK), Clinical Practice Research Datalink and NHS Digital, a unique environment is being developed to realise the benefits of access to well-curated, easily accessible health data for the life sciences research community. Medicines Discovery Catapult and the Association of the British Pharmaceutical Industry have been working with HDR-UK to identify the health data research and development needs of industry, large and small, and determine what needs to be put in place to make the UK a more attractive place to discover and develop new medicines. In this report, we present the findings of qualitative and quantitative research revealing a snapshot of biopharmaceutical industry experience in the UK health data landscape in 2019. The findings clearly show that while there is more to do, progress is being made to develop the clear, credible health data offer that industry requires. We look endeavours to build a world-class health data infrastructure for the benefit of patients and the NHS.





Chris Molloy (CEO, MDC) & Professor Carole Longson (CSO, ABPI)

Contents



3

Acknowledgements

This report is the result of a combined effort between the Association of the British Pharmaceutical Industry and Medicines Discovery Catapult in collaboration with Health Data Research UK, NHS Digital and the Office for Life Sciences. These organisations are very grateful to the participants of the research, the members of the pharmaceutical and life sciences industry and the wider stakeholders in medical research in the UK who gave invaluable insights via workshops, structured interviews, and an online survey.



Medicines Discovery Catapult

Daily engagement with SMEs

- The sample and data access theme of Medicines Discovery Catapult operates with independence and impartiality across the entire supplier community. It has been working to map the health data landscape and provide transparency for the SME sector to access the right data provider and, where necessary, data-analysis partner.
- Medicines Discovery Catapult has in-depth knowledge of the health data needs of SMEs and can help broker access to consented data. The organisation's report, State of the Discovery Nation 2018¹, showed that SMEs regard data and the NHS as critical assets, but struggle to engage with NHS systems and access data.



Association Of The British Pharmaceutical Industry

- The ABPI represents innovative research-based biopharmaceutical companies, large, medium and small, leading an exciting new era of biosciences in the UK.
- Our industry, a major contributor to the economy of the UK, brings lifesaving and life-enhancing medicines to patients. We represent companies who supply more than 80% of all branded medicines used by the NHS and who are researching and developing the majority of the current medicines' pipeline, ensuring that the UK remains at the forefront of helping patients prevent and overcome disease.
- Globally, our industry is researching and developing more than 7,000 new medicines.
- The ABPI is recognised by government as the industry body negotiating on behalf of the branded pharmaceutical industry for statutory consultation requirements, including the pricing scheme for medicines in the UK.



Executive summary

Patients, the NHS, the healthcare industry and the life sciences economy in the UK can all benefit from the ethical and trustworthy use of health data. To understand the potential benefits in detail, it is essential to consider a wide range of views from patients, academia, industry, and a breadth of other stakeholders. In this report, two UK organisations, the Association of the British Pharmaceutical Industry and Medicines Discovery Catapult, have specifically asked a broad range of representatives from the pharmaceutical and life sciences industry about their health data needs.

Our findings, derived from a series of workshops, structured interviews, and an online survey, provide evidence of what the health data needs of industry are, and the current challenges of working with health data in the UK. It identifies that current policy being undertaken in the UK to improve access to health data are having a beneficial impact, while maintaining the highest standards of transparency and ethical oversight. It also identifies areas where future efforts need to focus if the UK is to deliver a 'clear, credible national offer' for health data research and innovation that is in the public interest.



The UK is in a strong position

The UK has some of the most detailed and diverse health data assets anywhere in the world. It is feasible, with co-ordination, good governance, and partnership working with the NHS, to enable access to health data on large populations and to make national-scale improvements to health and care. Combined with unique research expertise, outstanding talent in the NHS and universities, and a vibrant life sciences industry, the UK has an unprecedented opportunity to use data at a large scale to drive innovation, support the UK pharmaceutical and life sciences industries, and improve people's health.

Healthcare industries are data-dependent

Life sciences research and development is an information-rich, evidence-based activity. Its success relies on access to, and the interpretation of, complex data. It is therefore critical that health data can be accessed ethically by researchers and industry if we are to realise the benefits of improved patient care, new medicines, diagnostics and the emerging use of artificial intelligence (AI).

The Association of the British Pharmaceutical Industry, Health Data Research UK and Medicines Discovery Catapult undertook three studies in 2019 to understand the health data needs of this sector:

- Global pharmaceutical expert workshops in collaboration with the industry's own Pistoia Alliance
- A series of structured interviews with targeted, leading industry representatives
- A quantitative online survey of a broad group of UK representatives from the health data user community, including academic and charitable as well as commercial users of health data

Needs for health data access

Six themes emerged from engagement with industry on health data needs for research and innovation:

- Data breadth, depth and scale longitudinal, eventbased, multimodal, curated data sets on disease subgroups are needed with full UK-wide coverage
- Access and speed a single, easy-to-use route for access to data, fast response times, and streamlined contracting, governance and approvals processes are needed
- 3 Quality this is a critical feature where few objective measures are available
- Expertise the industry needs access to domain specialists, AI and applied analytics and phenotyping
- **9** Public trust Research is designed to return the benefits of improved patient outcomes, increased efficiency in healthcare and to help industry develop more effective products and services. Therefore, there needs to be mechanisms to return benefits from data insights to the NHS, patients and the public to understand what the data will be used for, and to be confident in those handling it
- 6 Access and cost the data needs to be clear and affordable, particularly for smaller organisations to access, for whom the price of access to data, including cost, time and resources, can be a barrier or reason to use data from outside the UK

The survey of future requirements indicated an increasing need for data linkage, for example between:

- · Primary and secondary care data
- · Genomics and phenotypic outcomes
- Clinical trial data and patient-reported outcome measures

The industry says it will also require support in its growing need to access health data, including support to engage the public. Experience and research show that patients and the public are willing to allow access to their health data so long as they are engaged in a transparent process and that there are benefits both to patients and the NHS. Transparency on how data is stored, accessed and used is important in building public trust and this is actively encouraged by the Association of the British Pharmaceutical Industry and Medicines Discovery Catapult.

Barriers to health data access

The research identified a number of systemic barriers that limit access to data. These include the time taken to access data, constraints to access for commercial companies, the effort to identify and assess the quality of data sets, and cost. When surveyed, members of the health data user community reported that only 25% of recent requests for data had been completely successful. While access to clinical trial data was the most straightforward – with a success rate of 45% – access to linked primary and secondary care electronic health records data was very often deemed by industry to be unsuccessful.

Progress in health data access

Respondents reported growing momentum within the UK to change, and that collaborative working among custodians of UK health data is having a positive impact; for example:

- Increasing size of web based resources such as Clinical Practice Research Datalink (CPRD) and UK Biobank
- Creating the **National Centre for Expertise** within NHSX, a joint organisation for digital, data and technology
- The establishment of the UK Health Data Research Alliance – an inclusive alliance of data custodians committed to making an increasing breadth and depth of data available for research and innovation purposes
- User-led development of the Health Data Research Innovation Gateway – providing discovery, accessibility, security, and interoperability to surface data; to support linkage; and to enable health data science projects safely and efficiently
- Developing the Health Data Research Hubs making data available, curating data, and providing expert research services

The UK has the opportunity to catalyse the innovative and ethical use of a rich base of health data to improve people's lives through increased efficiency of the NHS and the development of new, more effective medicines. This requires the combined support of policymakers, health services and professionals; data custodians; the pharmaceutical and life sciences industry; and the public and patients. Tangible steps are being taken, but concerted effort across these groups is needed to deliver this 'clear, credible offer' for the benefit of all. Use of health data by the life sciences industry : a UK perspective

Introduction

"**The data access process has to be improved** for companies to use UK data sets more"

– UK-based SME

Healthcare and life sciences are information-rich, evidence-based activities. Their success relies on access to, and the interpretation of, complex data. For the benefit of improved patient care, new medicines, diagnostics and the emerging use of artificial intelligence (AI), it is critical that data gained ethically can be made accessible to research and industry.

The explosion of interest in data and its potential value requires engagement from data controllers and regulators to meet this demand. The UK is taking strategic steps, across its many stakeholder groups, to address this complex challenge. There is a strong desire to create a clear, credible approach that enables the pharmaceutical and life sciences industry to access health data.

The UK is uniquely positioned to lead the revolution in health data research and innovation: it potentially has some of the richest health data anywhere in the world. It is feasible to enable access to health data on a large and diverse population through the NHS and to make national-scale improvements to people's health and care. Combined with unique research expertise, outstanding talent in the NHS and universities, and a vibrant life sciences industry, the UK has an unprecedented opportunity to use data at a large scale. This will drive innovation, support the UK pharmaceutical and life sciences industry, and improve the health of the public.

The 2019 *Topol Review*², an independent report on digital health commissioned by the UK Department of Health and Social Care, found: "There is a need to complete the digitisation and integration of health and care records if the full benefits of digital medicine (earlier diagnosis, personalised care, and treatment) are going to be realised for the NHS." It also said: "The collaborative partnership will need to be stretched to include industry when potential solutions to new regulatory challenges are needed. Regulation should be an enabler, not a barrier to innovation."

Industry partnership, through skills, knowledge, expertise, and investment, is essential to realise this opportunity. This report combines primary research, interviews and analysis gained from over 150 industrial partners between January and June 2019, including large pharmaceutical, biotech, medtech, diagnostics, contract research, data intermediary, and technology companies and consulting firms.

The messages delivered by this report are clear on the views of varied stakeholders in the pharmaceutical and life sciences industry.

² 'The Topol Review: An independent report on behalf of the Secretary of State for Health and Social Care: February 2019.' Leeds: Health Education England; 2019 (https://topol.hee.nhs.uk/wp-content/uploads/HEE-Topol-Review-2019.pdf, accessed 16 September 2019).

Views of pharma and life sciences industry on health data access

Three parallel streams of structured research, with overlapping industry groups, was conducted to understand their health data access experience and future needs:

- Global pharmaceutical expert workshops supported by the industry's Pistoia Alliance
- 2 A series of structured interviews with targeted, leading representatives of the global health industry
- 3 A quantitative online survey of a broad group of UK representatives of the health data user community, including academic and charitable as well as commercial users of health data

This report focuses on the findings from the structured interviews and online survey, and analyses the steps being taken to address the health data needs of the pharmaceutical and life sciences industry in the UK.

The robustness of our findings on the industry's requirements is supported by the common themes and approaches that emerged from the structured interviews and quantitative survey data. The challenges revealed here by industry researchers are similar to those expressed in a similar survey of academic users of health data for cancer research, conducted by the National Cancer Research Institute.³

Figures 1A&B shows the demographics of the industrial input into sections of this report. Figure 1A details the participants in face-to-face interviews and Figure 1B the respondents to the online survey. The data show the effective coverage of this report and the relevance of its findings to the UK's industrial base.



Figure 1A. Expert Interview Data

Figure 1B. Online survey

³ National Cancer Research Institute (NCRI) survey on accessing health data for cancer research, conducted during June 2019 and surveying the experience of members of NCRI research groups (unpublished).

Reported health data uses



Figure 2

Why industry needs health data

This research shows that accessible, usable health data offers many opportunities for SMEs and large multinational companies. These opportunities include:

- Early-stage drug discovery data used to identify unmet medical needs, discover new targets and understand patient variation
- Improving drug translation drug discoveries moving from preclinical to clinical development
- Efficacy and safety a better understanding of the thresholds of clinical efficacy and tolerability in routine medical practice
- Clinical trial recruitment and patient stratification knowing, for example, in which patients a drug will not be effective or will present too high a risk
- Development of 'companion diagnostics' for informing the above aspects
- Medical device development such as smartphone apps and wearables
- Algorithm development and AI access to high-quality, representative data against which to calibrate and validate algorithms and AI models
- Generating commercial dossiers and supporting value-based pricing – including for health technology regulatory assessments

Different priorities across industries

- - - - - -

•••

.

. . .

.

.

.

.

. . . .

Source: online survey

.

.

Although many of the identified needs for health data are industry-wide, different subsectors have different priorities at different times. SMEs, for example, are more focused on early research, discovery, and development (e.g. target identification, safety studies), with a necessity for rapid engagement and delivery. Figure 2*, shows the results of the online survey for reported health data use.

Larger firms, meanwhile, have use for data throughout the product life cycle, all the way through to regulation, market access and product monitoring.

We observed a high demand for patient data to inform clinical trial recruitment. Other main uses for health data were understanding current treatment pathways, unmet medical needs, and comorbidities.

The most commonly cited data sources were web-based databases, the Clinical Practice Research Datalink (CPRD), NHS Digital, NHS trust clinical records and the UK Biobank, along with free-to-use data sets like that of the International Cancer Genome Consortium (ICGC), plus commercial data sets from a range of data intermediaries and/or from U.S. health insurers.

*Participants represent UK and global small to medium-sized enterprises (SMEs) and large companies, including; drug discovery and biopharmaceuticals, contract and clinical research organisations, medical technology companies, charities, information systems specialists, medical device developers, specialist consultancy companies, academia, funding bodies, NHS research, and practitioners

Types of health data that industry uses

Figure 3*, shows the types of data respondents had requested access to, as revealed by the quantitative online survey run over one month:

- · 92 separate data access request attempts were recorded
- The data type with the highest attempt rate was clinical trial data (22% of responses)
- The data type with the lowest attempt rate was administrative secondary-care data (5% of responses)
- There was a high demand for clinical real-world data from electronic health records, disease-specific data (registries) and genomics data



Figure 3.

In interviews, companies mentioned several disease areas they considered to be priorities for data access. Cancer, central nervous system and inflammation/immunity were important for over half of those interviewed (Figure 4**). Many companies prioritised access to linked clinical and 'omics' data to support, for example, the linking of genomic data to phenotypic outcomes, needed mainly for early research but also for trials of personalised medicines. Companies reported a need for primary care and secondary care data sets to be linked to be able to build an improved understanding of the whole patient and build a more complete picture of the effectiveness and safety of treatments, and patient outcomes across entire patient journeys.

Reported therapy area coverage



Figure 4.

Source: structured interviews

Source: online survey

*Participants represent UK and global small to medium-sized enterprises (SMEs) and large companies, including; drug discovery and biopharmaceuticals, contract and clinical research organisations, medical technology companies, charities, information systems specialists, medical device developers, specialist consultancy companies, academia, funding bodies, NHS research, and practitioners

Reported health data access attempt

Industry data access experience

While the pharmaceutical and life sciences industry has clearly identified the potential benefits of appropriate access to health data, it reports difficulty in realising them. Most companies have experienced delays and uncertainties in accessing data. "The data access process has to be improved for companies to use UK data sets more," was the appeal of one UK SME. Figure 5* reveals the most common barriers to access identified in the online survey. The time taken to access data, access constraints for commercial users, the effort to identify and assess the quality of data sets, and cost were common barriers. The interviews also reinforced that UK data had a good reputation for quality, but a poor one for delivery.



Source: online survey

Successful

Figure 5.



Of the 92 separate data access requests recorded, only 25% were deemed completely successful. As seen in Figure 6*, the data type with the highest access success rate (45%) was clinical trial data. It is also the most demanded type of data, as seen in (Figure 3). The responses also show:

Health data access success by data type

- The data type with the highest percentage of unsuccessful access attempts was secondary -care electronic health records data – 43% of access attempts were unsuccessful
- Access to primary-care data was of limited success, with no fully successful attempts, and most (88%) being only partially successful.

Unsuccessful

Partially successful



Clinical trial Secondary Care – EHR Disease registries Genomics Primary care Volunteered health data Phenotypic Research cohorts Secondary Care – administration

Figure 6

**Interviews with UK representatives of international and national companies. Of these, a third were pharmaceutical companies, another third were UK SMEs, and the remaining third were split between data, medical technology, contract research, diagnostic, and advisory firms.

Barriers encountered when accessing data

Industry has six key requirements for health data

Six themes emerged from the engagement with the pharmaceutical and life sciences industry on what they say it needs from health data:

- Data breadth, depth and scale longitudinal, event-based, multimodal, curated data sets on disease subgroups are needed with full UK-wide coverage
- Access and speed a single, easy-to-use route for access to data, fast response times, and streamlined contracting, governance and approvals processes are needed
- Quality this is a critical feature where few objective measures are available
- Expertise the industry needs access to domain specialists, AI and applied analytics and phenotyping
- Public trust Research is designed to return the benefits of improved patient outcomes, increased efficiency in healthcare and to help industry develop more effective products and services. Therefore, there needs to be mechanisms to return benefits from data insights to the NHS, patients and the public to understand what the data will be used for, and to be confident in those handling it
- Access and cost the data needs to be clear and affordable to access, particularly for smaller organisations for whom the price of access to data, including cost, time and resources, can be a barrier or reason to use data from outside the UK

Data breadth, depth and scale

The respondents asked for clinical and 'omics' data to be linked. This would support early research to understand and stratify disease. They also wanted data on accessible, relevant patient populations to assess the feasibility of trial design and ways to recruit specific, uncommon patient types to clinical trials. While some of these uses could be enabled by the extension or improvement of current systems, the following key changes are required:

- · Direct linkage to secondary-care data
- · Rapid assessment of patient numbers across sites
- · Ability to recruit patients identified from national data sets

Secondary-care data sets, linked to primary-care data if possible and available longitudinally, are very valuable. Companies noted that the involvement of NHS hospital trusts was vital to enabling timely access to detailed data on specific diseases and treatments, which is not available from national data sets. The lack of pre-approved contracts and national guidance made it difficult, however, for hospitals to efficiently grant data access for research.



"The involvement of NHS hospital trusts is vital to enable timely access to detailed data" - Global pharmaceutical company

Other types of data linkage desired by the respondents were between genomic and phenotypic outcomes, and data and patient-reported outcome measures (PROMs). There was a marked increase in the desire to use linked genomic and phenotypic data, particularly from the diagnostic companies taking part in the structured interviews. Other areas of expansion for data use are epidemiology, comorbidity, pharmacovigilance, polypharmacy, and cohort recruitment.

Access and speed

The most consistent message for the UK was to prioritise transparent, predictable, quick access to data, for example by creating pre-approved contracts for data access.

While there is a strategic drive to enable access to health data, the wide range of data custodians across the country makes it challenging for the pharmaceutical and life sciences industry to gain this access.

Most companies in this research had experienced delays, uncertainties and a lack of clarity. These factors had prevented data from being accessed, and companies felt that the process must be improved. Creating pre-approved contracts for data access, for example, would help. This issue is a particularly stark one for SMEs – only 25% of data access attempts were successful according to our survey.

"Slow turnaround times mean that we can't work with the UK"

Global contract research organisation

"The data access process has to be improved for companies to use UK data sets more"

– UK-based SME

Quality

Companies recognise that significant work is needed to prepare data sets for a specific research question. They want transparency, though: they need to know early on if the data cannot answer the question or cannot be accessed, rather than run the risk of this happening later. Some companies preferred research cohort data because of its quality, and others were keen to understand the availability of associated bio-samples for translational research.

Expertise

78% of survey respondents wanted support in identifying data sets and assessing their quality. Having clarity about where data is and how to access it is a key need, and this was amplified in the interviews of health data experts. The SME sector needs particular support in data access, and in understanding the access process and the commercial constraints for commercial companies. It also needs support in data analysis.

Public trust

The interviewees from the pharmaceutical and life sciences industry saw improved patient engagement as a priority. Sharing the examples of good practice and championing positive outcomes for patients would benefit the sector in breaking down the barriers to wider uptake.

Engagement exercises have shown that citizens are willing for their data to be used in research if data security can be reliably promised and if there are clear, explained benefits to patients and the NHS. People also want a clear process when giving their informed consent for the ethical use of data to improve healthcare. The process of giving informed consent needs to be balanced between comprehensive but also easyto-understand information. A public engagement exercise commissioned by the Health Research Authority has found, for example, that patients need "a straightforward and accessible consent process." And patients who have donated tissue to biobanks would be happy to share personal data if a clear explanation was given of how their data would be used.⁴

Public engagement is vital to improve trust in the industrial uses of health data. A series of workshops with patients and members of the public was recently conducted by the Health Research Authority and the University of Sheffield.⁵ Participants were initially very reluctant about sharing anonymised patient data with commercial organisations for reasons other than direct care. Before the workshops, only 18% felt this would be acceptable. But after the engagement, 45% were supportive.

The following examples show that support from the public for health data use can be gained through engagement:

- Volunteer panels benefit from a willingness of participants to take part without direct personal gain – the UK Biobank, for example, gains broad consent for the use of anonymised data and keeps its participants involved through progress updates, newsletters and a dedicated 'participant resource centre'.⁶
- UseMyData describes its members as patients, relatives, and carers who support and promote the sharing of health data to improve patient treatments and outcomes, "knowing that effective safeguards to maintain confidentiality and anonymity ... are applied consistently, transparently and rigorously."
- A survey of over 3,000 asthma patients found that 88% would be comfortable with their confidential health data being used for research to develop new asthma treatments. This altruism rose to 94% being willing to share their anonymised health data with an analytics company if a tool could be developed to target people at particular risk of an asthma attack.⁷
- Following the recent publication of the UK's revised principles on the use of health data, the UK government is undertaking a further round of detailed engagement with stakeholders. This includes commissioned public engagement work being linked through the group Understanding Patient Data, which will inform the development of a full, final policy framework. The full policy framework will be published later in 2019.

Access and cost

Data budgets were seen by participants in our research to be larger for supporting regulatory submissions for health technology assessments, for value-based pricing and for supporting clinical trials. Funding for data access and analysis is thus biased towards later-stage drug development – where budgets are bigger. For SMEs, where activities are typically at the earlier stages of drug discovery, these costs can be a barrier to using health data. Data repurposing activities within larger organisations could allow for greater and more diverse insights from the same data sets.

.

.

.

.

• • • • • • • • • • •

•

⁴ 'Public support for greater data sharing with biobanks.' London: Health Research Authority; 2018 (https://www.hra.nhs.uk/about-us/news-updates/public-support-greater-data-sharingbiobanks, accessed 16 September 2019)

⁵ 'Sharing anonymised patient-level data where there is a mixed public and private benefit – a new report.' London: Health Research Authority; 2019 (https://www.hra.nhs.uk/about-us/newsupdates/sharing-anonymised-patient-level-data-where-there-mixed-public-and-private-benefit-new-report, accessed 16 September 2019).

⁶ Sudlow C, Gallacher J, Allen N et al. 'UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age.' PLoS Med. 2015;12(3): e1001779. doi: 10.1371/journal.pmed.1001779.

⁷ West B and Cumella A, 'Data sharing and technology: exploring the attitudes of people with asthma.' London: Asthma UK; 2018 (https://www.asthma.org.uk/datareport, accessed 16 September 2019).

Progress

The custodians of UK data are already acting as a group on what we have learnt from engagement with industry, the NHS, academia, charities, patients and the public. The following tangible actions are now happening to realise the ideal of a clear, credible industry offer for UK health data:

Health Data Research UK

Health Data Research UK (HDR-UK) is the national institute for health data science. Its mission is to unite the UK's health data to enable discoveries that improve people's lives. Its work includes the following:

Public trust

HDR-UK has a commitment to earning, building and sustaining public trust through being clear and transparent about how health data is being used, and by involving patients and the public at every stage of the innovation process. HDR-UK established a Public Advisory Board in January 2019 to give advice on its strategic priorities and approach across its research sites and hubs.

UK Health Data Research Alliance

Launched by HDR-UK in February 2019 to facilitate partnership across NHS organisations and other health data custodians, leading to an ethical, consistent approach to the use of data and to public engagement. The alliance has 18 members, including national custodians across the four UK nations, NHS trusts, charities, and custodians of genomics data.

Health Data Research Innovation Gateway

The purpose of the Health Data Research Innovation Gateway is to provide a common portal to enable the UK's health data to be discovered and accessed safely and responsibly. The gateway is being developed in an agile way, with user co-design through a 'sandbox', involving testing by researchers and innovators. It will be available from the start of 2020 to enable data discovery and access.

Health Data Research Hubs

Health Data Research Hubs are centres of excellence with expertise, tools, knowledge, and ways of working to maximise the insights and innovations developed from health data. They are formal collaborations between the NHS, academic organisations and industry, working in partnership with patients and the public. Each hub will demonstrate and drive the utility of expert health data science based on a particular disease, condition, clinical trial or real-world evidence.

UK plc

Major national stakeholders and policymakers recognise the importance of this clear, credible offer and are working in concert to deliver elements of the offer. Relevant areas of tangible progress include the following:

CPRD

CPRD collects de-identified patient data from a network of GP practices across the UK. Primary care data are linked to a range of other health related data to provide a longitudinal, representative UK population health dataset. The data encompass 42 million patient lives, including 13 million currently registered patients.

New linkage data is now available for patients from an additional 1211 practices. It has increased the coverage periods for many linked data sources and includes the National Cancer Registration and Analysis Service and Quality of Life of Cancer Survivors data. It also contains data linked to patient 'postcode'.

Linkage of primary and secondary healthcare data (NHS Digital)

NHS Digital, NHSX and partners are seeking to establish a new approach for the use of general practice data for planning and research, and enabling secure linkage of this to other key data sets such as hospital data in a way that is supported by the public and medical professionals. In time, this will enable the strategic replacement of the current General Practice Extraction Service (GPES) solution as part of the new GP IT Futures framework.

Improved governance principals, further patient engagement and commercial support for industry access transactions (Office for Life Sciences)

To ensure that NHS organisations have the necessary guidance to be able to enter into appropriate arrangements for the use of their data, the UK government is developing a policy framework. This was underpinned by the five guiding principles published in draft form in the second Life Sciences Sector Deal in December 2018 and revised in July 2019.⁸

In addition to the policy framework, the UK government is in the process of establishing a national Centre of Expertise to provide expert commercial and legal expertise for NHS organisations when negotiating data partnerships. The centre will sit within NHSX to ensure it aligns with the wider policy landscape for data, digital and technology.

A new service could be incrementally available from summer 2020. Further detail on the specifics of the approach will be available in due course.

Ongoing and future needs

Despite the challenges outlined in the research reported here, the pharmaceutical and life sciences industry sees high value in the UK's health data – and a growing need for access to it. There is a clear role to support simpler industry access to this data, and capture value for patients and the nation.

We heard strongly that there was a need for a central health data discovery gateway that would act as a broker for distributed data sets. This would show quickly what data was available – and would be particularly valuable if it included the terms and timelines of access plus measures of quality. We believe this clarity about where data is, and how it can be accessed, is a key need, and it was one echoed by the healthcare experts in our research interviews.

A consistent message from UK industry was to prioritise transparent, predictable, quick access to data, for example by creating pre-approved contracts for data access.

The quantitative survey suggests the need for health data will rise (103 responses indicated a future need, compared with 77 already making use of health data). Support will need to be given for this growth across the full range of data types and uses. Some 78% of respondents need support in the identification of data sets and the assessment of their quality. The survey of future requirements indicated a rising need for data linkage, for example between:

- Primary and secondary care data
- · Genomics and phenotypic outcomes
- · Clinical trial data and patient-reported outcome measures.

The SME sector in particular needs support for accessing health data. More than two-thirds of companies with fewer than 10 employees called for advice and support to understand the access process and access constraints for commercial companies.

The survey also suggested a requirement for increased active support in data analysis, adoption of new analytics capabilities, skills and training.

.

.

• • • • •

78%

of respondents need support in the identification of data sets and the assessment of their quality

Conclusions and next steps

This engagement with the pharmaceutical and life sciences industry shows that the UK health data landscape is evolving fast. The industry has been clear about what it needs from the availability of ethically sourced health data to be able to make meaningful, innovative use of it. The UK now has the opportunity through 'system' design' to provide a clear, credible offer on responsible access to health data. All sizes and shapes of pharmaceutical and life sciences organisations need to be able to access data in a single, timely, cost-effective manner. Many, especially SMEs, need support to access data - and all organisations need to build partnerships with patients and the public to earn trust.



Six themes have emerged in the requirements for health data expressed by the pharmaceutical and life sciences industry:

- 1 Breadth, depth and scale of health data
- A single, easy-to-use route for access to the data, fast response times, and streamlined contracting, governance and approvals processes
- High-quality data
- Expertise in specialist and emerging areas such as Al and applied analytics
- 9 Public trust including by returning benefit and insights from the data to the NHS and public
- 6 Cost-effectiveness for all sizes of organisation

The UK is making progress in engaging the custodians and users of health data. Each of the organisations behind this report, and others across the UK, are working in concert to create a clear, credible, ethical, consistent approach to the use of health data. This work includes the creation of the UK Health Data Research Alliance and Centre for Expertise to improve access, contractual agreements and data standards, and to engage the public; the Health Data Research Innovation Gateway, a user-led development to provide a secure and efficient route to discover and access data; and Health Data Research Hubs.

This work will continue to be co-developed with industry so that the data from UK patients benefits them, the NHS and the industrial wealth of the nation. The UK has the opportunity to catalyse the innovative and ethical use of a rich base of health data. This demands the combined support of politicians and policymakers, health services and professionals, data custodians and users, the pharmaceutical and life sciences industry, and the public and patients.

The UK now has the opportunity through 'system design' to provide a clear, credible offer on responsible access to health data

Appendix: research methods

Three parallel streams of structured research with overlapping industry groups was conducted to understand their health data experience and desires:

- Global pharmaceutical expert workshops supported by the industry's Pistoia Alliance
- Health data industry leaders in a series of structured interviews
- A quantitative survey of a broad group of UK representatives of the data-user community, including academic and charitable as well as commercial users of health data

Consortium workshops

A structured series of workshops supported by the industry's own informatics consortium, the Pistoia Alliance, were undertaken. This alliance has a senior-level membership across over 100 international firms. The present report contains feedback from an initial meeting held in April 2019. These workshops in 2019 and 2020 will continue to refine the global industry need.

In-depth interviews

Following the first Pistoia workshop, 29 interviews were undertaken with UK representatives of international and national companies. Of these, a third were pharmaceutical companies, another third were UK SMEs, and the remaining third were split between data, medical technology, contract research, diagnostic, and advisory firms. The large global companies had representatives of more than one department (e.g. early research, clinical trials, real-world evidence). The interviews were conducted around a structured questionnaire with open questions on the company's needs for access to UK health data.

Online survey

An online survey ran for one month. There were 52 responses from 41 unique, self-identified organisations from across a spread of organisation size, geography and sector coverage. All regions of the UK except Northern Ireland were represented, although most respondents were from three UK hubs of the North-West of England, London and the East of England.

 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0

The respondents to the online survey were reached through Medicines Discovery Catapult's communications network. They represent UK and global small to medium-sized enterprises (SMEs) and large companies, including drug discovery and biopharmaceuticals, contract and clinical research organisations, medical technology companies, charities, information systems specialists, medical-device developers, specialist consultancy companies, academia, funding bodies, NHS researchers and practitioners.

Data collected

The interviews with health data industry leaders collected information on:

- Health data priorities
- · What data is required and what for
- · Where/how it is currently obtained
- · Limiting factors

The quantitative online survey of UK stakeholders collected data on:

- · Size, location, and type of organisation
- Type of health data accessed
- · Uses of health data
- The success of previous data access attempts and what barriers there were (if any)
- Support needed for future access attempts



info@md.catapult.org.uk $(\simeq$

@meddisccat

(in linkedin.com/company/meddisccat

01625 238 734

y

Medicines Discovery Catapult Mereside, Alderley Park Alderley Edge Cheshire **SK10 4TG**



@ABPI_UK

www.linkedin.com/company/abpi

 (\mathbf{C}) 0207 930 3477

Association of the British Pharmaceutical Industry 7th Floor Southside 105 Victoria Street London SW1E 6QT 0 0 0 1



