



## Healthcare and Data: How do we get it right?

A summary of the meeting held at The Intellectual Forum, Jesus College, Cambridge  
13<sup>th</sup> and 14<sup>th</sup> September 2018

### Background

There is currently a huge amount of interest in the potential of data for health. The data collected as part of healthcare provision can be harnessed and used to develop improvements in services, drive advances in medical research and provide a strong basis for the UK life sciences industry to thrive. The UK has excellent resources and programmes, including UK Biobank, the 100,000 Genomes Project, the Clinical Research Practice Datalink and several world-leading longitudinal cohort studies generating rich, high-quality data. Enthusiasm for health data science is underpinned by advances in machine learning, cloud computing, digital health tools, and the development of integrated care records.

However, the promise of data-driven technologies will only be realised if there are trustworthy systems in place for managing patient data, and the public have confidence in how these technologies are developed and deployed, both for delivering care and for research. The Information Commissioner has argued that privacy and innovation can go hand in hand, but are we on track to achieve this in the health sector?

*'Healthcare and data: How do we get it right?'* was hosted at Jesus College, Cambridge on 13-14<sup>th</sup> September 2018 to try and address this question. It was held in two parts: an open session of panels and debate, and a closed discussion with a focused group of representatives to respond to the issues raised. The event's aim was to bring key stakeholders together, enable debate with a range of perspectives, and – crucially – define next steps on which to act.

The full event is summarised in this report. Rather than detail the discussions chronologically, themes have been pulled together to reflect how the issues raised cut-across the topics of the panels and debates. These issues have been summarised here to reflect the conversations of the event; they do not comprehensively capture the entire breath of health and data debates.

For this meeting report we use the term 'data-driven' technologies to refer to the broad range of tools, approaches and methodologies that rely on the use of data from patients for their development, testing and deployment. The focus was on data collected as part of clinical care and treatment, held in medical records, which could be used for purposes beyond providing individual care.

The meeting was convened by Understanding Patient Data, Julian Huppert from the Intellectual Forum at Jesus College Cambridge, and Hal Hodson, technology correspondent at The Economist. The event was held under Chatham House rules.

# Open Session

## The potential and concerns of healthcare and data

The open session included four panel discussions, two keynote presentations, and questions from the floor. (See **Annex B** for the full agenda). There was clear consensus from this session that patient data has realistic potential to beneficially transform healthcare and research, but unlocking that potential means tackling a complex range of barriers.

### 1. Potential of data-driven technologies

The potential for health data and data-driven technologies to vastly improve health research and care was widely recognised throughout the open session. Though often considered hyped – many technologies are in early stages or not-yet developed and there is a risk of over-promise – there are examples where data and artificial intelligence tools are already having positive impacts in health. These include:

- [using genomic data to improve the diagnosis and treatment of cancer](#).
- [revolutionising the way professionals carry out eye tests](#), leading to earlier detection of common eye diseases and helping address growing demand.
- improving the health of people with cystic fibrosis through research, using data which they submit to a [registry held and supported by the Cystic Fibrosis Trust](#).
- saving clinicians time by [alerting them to patients at risk of certain conditions](#).

(The [AHSN AI Network initiative](#) has also mapped current uses of AI in the NHS.)

However, to realise solutions like these at larger scale, many issues and concerns must be addressed. This is vital for the potential of data in health to be fully unlocked.

### 2. Concerns about data use in health

People at the meeting spoke from many different perspectives, but the concerns raised clustered around several themes:

- The **quality of health data and data infrastructure** in UK health and research
- **Trustworthiness of data use, data protection, and data privacy**
- The need for **public dialogues** about the use of health data
- **Effective collaborations** across healthcare systems, academia, and industry

The key points from the discussions are distilled below, centred around these themes.

#### 2.1. Data quality and infrastructure

The potential of data to improve health is dependent on high-quality, comprehensive and interoperable data, which can be readily accessed by researchers, healthcare professionals, and patients. However, in the UK the quality of existing health data and the infrastructure to collect and use it is uneven across the country and often not fit for research purposes:

**Analysis-ready data:** Current projects spend significant time cleaning or preparing data for use. This time massively delays progress across this field and limits who has the capacity and resource to make good use of data. Estimates made of the cost of 'cleaning' the bulk of NHS data ranged from £20million to £2billion.

<b>Fragmented data systems:</b>	Health data currently exists in disparate formats, not shared between different trusts and health providers. Not only does this hinder the delivery of healthcare, but it inhibits research and disempowers patients who are unable to access and make use of data about them.
<b>Infrastructure is not fit-for-purpose:</b>	Outdated IT equipment and software within NHS systems, combined with high-burden on healthcare professionals for data entry, prevents the collection of high-quality, consistent health data.
<b>Scalability of solutions:</b>	It is incredibly challenging to scale local solutions up to a national or global level. What works in one hospital might not in another because of a lack of joined-up standards for IT and data across health systems, or a difference in cultures – particularly in the NHS.
<b>Cultural shifts are needed:</b>	The culture of the healthcare system is often not conducive to collecting high-quality data or using it to drive better care, though there are strong counter-examples. Healthcare providers see little benefit to themselves or their patients from investing time and energy in data collection or use.
<b>Accounting for data value:</b>	Large, comprehensive data sets are of value to industries and researchers. The NHS has large data sets, but currently has no ways to financially account for them. This makes it hard to justify investing in data curation and management, or to realise its worth in commercial partnerships. The value of this data must be returned to the NHS, and the benefits of its use passed onto patients.

## 2.2. Trustworthiness in data uses

Trust is crucial for patients, healthcare professionals, and the public to have confidence in the use of data for health and research. However, trust is hard to build, and can be undermined in a single moment. Care.data, the NHS Wannacry hack, and Cambridge Analytica demonstrate how rapidly trust can be eroded. To build the trustworthiness of uses of data, many concerns must be addressed:

<b>Commercial interest:</b>	While the private sector provides much-needed innovation, many feel their profit-making interest must be 'kept in check'. There is concern that, left to their own devices, private organisations will not properly protect patient data, will seek to exploit it where possible, and will not use it primarily for social benefit.
<b>Who benefits?</b>	There are many questions about whether data collection and use is for the benefit of the patient, clinicians, researchers, governments, or private industries. This can breed reluctance to share data, or less trust that it is being used responsibly and fairly.
<b>Effective, clear regulation:</b>	Laws and regulation must be clear. GDPR has increased the focus on data protection but also generated confusion about what is or is not allowed. Those using health data must be able to say why they need the data, justify how they use it, and provide evidence of transparency, choice, and accountability. At the same time, regulation must not stifle innovation but support and guide it.

**Do no harm:** There will be times when data protection rules must be broken to protect an individual's health or when not doing so would result in harm. However, breaking rules needn't mean breaking trust – if proper regulations and clear guidelines are in place, then healthcare professionals will know how to judge what is appropriate in different circumstances.

### 2.3. Data privacy and protection

Data privacy and protection is closely intertwined with trust. Without proper respect for data privacy, not only is trust undermined, but so too are the benefits that data-use can bring to health and care.

**Getting choices right:** The use of data about a patient may require their consent, or may be justified on a different lawful basis depending on the circumstances. However, the role and limits of consent must be clear, particularly in light of GDPR and the common law duty of confidentiality. "Consenting" does not necessarily give individuals agency over what happens to data about them. It must be clear what control or choice people do or do not have over data about them.

**Data trusts:** New conceptual models may enable people to exercise control, while protecting data and facilitating its use at the same time. One example is the concept of "data trusts" – independent organisations tasked with brokering relationships between patients, their data, and researchers or organisations wishing to use it. However, there are questions about how such organisations might contend with the power of large, private corporations, and how they could be scaled.

**Health is personal:** Health data is highly personal. Many individuals are wary of sharing data in case their personal information becomes exposed from a lack of proper data privacy and protection. In cases where a health condition carries a social stigma, this wariness can prevent people seeking medical advice.

**Confidentiality:** Doctor-patient confidentiality, a longstanding pillar of healthcare delivery which builds trust and confidence, is undermined when patients feel information about their health will be shared beyond just the doctors and healthcare professionals they speak to.

**Transparency:** Fears about who else can access health data often impacts those most socially vulnerable, such as immigrants or asylum seekers. Concerns that health data may be accessed by government authorities - whether true or not - can prevent people from approaching healthcare services. Transparency about data use and the rules around it might allay these concerns.

## 2.4. Public dialogue

Those using data for healthcare and research must put public engagement at the heart of their work. This should inform individuals about the use of health data, while recognising and addressing their concerns. It must also shape decisions about how data is used in practice and what safeguards are put in place to provide public assurance.

- Local engagement:** Public engagement on health data must exist at a local level, embedded in and spread across communities. A larger, national public campaign would not be able to address the complexity of issues or be able to respond to public concerns.
- Empowering patients:** The use of data must not disempower patients. They must not feel ‘trapped’ by their data, unable to contest it when it’s wrong, or beholden to data which is incomplete or inaccurate.
- Building confidence for health professionals:** Healthcare professionals lack confidence and capability to use data and have open conversations with patients about data. This comes from a lack of clear guidance and training, and further deepens public concern. Existing tools for data protection are perceived by health professionals as a hindering ‘check-box’ exercise, rather than a useful resource.
- Multiple publics:** There isn’t a single ‘public’ responding to the use of data in health, but rather a range of publics with differing experiences. For example, patients are often more comfortable with data use when it will directly improve their quality of life. Public dialogue must be able to engage across a wide range of interests and perspectives, including those who are seldom heard by policymakers and researchers.

## 2.5. Collaboration between sectors

Unlocking and developing the potential of health data requires collaboration from a range of stakeholders: health systems, patients, researchers, governments, charities, industry and more. Currently, however, there are several barriers to this collaboration.

- Lacking leadership:** There is no clear, visible leadership across organisations and sectors, with a common vision and language. Without this, organisations are left to forge ahead separately, meaning solutions and developments are fragmented, disconnected, and imbalanced.
- Business models:** A lack of guidelines or models for fostering collaboration between the health system and the private sector leads to disconnected approaches and a lot of wasted effort seeking to form partnerships on different terms. This makes it more challenging for smaller organisations and researchers with innovative ideas to bring them to fruition.
- Data across health, research, and innovation:** Health data is used across a spectrum of organisations. At one end are those directly providing healthcare; at the other are those undertaking data-driven research that might lead to medical advances. Frameworks, guidelines, and regulation must be able to account for these different uses, and for collaboration between them.

## Closed Session

### Defining barriers and identifying actions

The closed session aimed to more clearly define the issues raised by the open discussions, before identifying and mapping gaps in existing work to overcome them. The output of this session was a set of specific actions that participants could potentially undertake, collectively or with identified others, to address the systemic issues and drive better uses of data.

### 3. How to get healthcare and data “right”

The broad range of existing issues were narrowed into specific barriers which must be tackled to get healthcare and data “right”. Overcoming these barriers requires:

- Developing the UK’s current **health data infrastructure** and improving access to **analysis-ready data**.
- Making sense of over-abundant, inconsistent **guidelines and regulation**.
- Building **trustworthiness and confidence in data uses** and addressing a need for **effective engagement** with public communities, patient groups, and healthcare professionals.
- Providing health professionals with **resources and training** to make best use of data in their work. This includes **research** being considered alongside direct care as an important rationale for data collection and use.
- Establishing the right **business models** and frameworks to support collaboration between healthcare systems, academia, and industry.

The steps necessary for each of these points were then detailed further:

#### 3.1. Improving data quality and infrastructure sustainably

An **efficiently interoperable IT infrastructure** is needed within and across health systems, to ensure analysis-ready data can be collected, shared, and used easily, with high trust and low cost.

Research is needed to understand how to **maximise the utility of existing data**, through curation and algorithmic means. This should explore the costs, viability, and benefits of doing so.

**Standards for collecting, handling, augmenting, and sharing data** should be developed and adopted across healthcare systems for maximum patient and societal benefit.

**Standards for interoperable data-processing algorithms** should be adopted, so different technologies and software can communicate with one another across IT systems.

**Models for using data across direct care and research** should be developed, by building on existing examples, such as the NHS Local Health & Care Record Exemplars (LHCRES) if they prove to be successful.

#### 3.2. Developing guidance and best practice

The plethora of **existing guidance needs scoping, mapping, and synthesising**, to make it more consistent and allow healthcare professionals to action it locally and nationally.

A **single advisory body** could be established to provide clear advice on best practice in data uses, while maximising innovation. It could provide a unified view across relevant regulators (ICO, MHRA, HRA, CAC) and those with oversight of health data use to help develop common

understandings of practical, technical, legal and regulatory issues for better use of health data. It should synthesise guidance, not add to it, and should engage researchers, industry, healthcare professionals, patients, and the wider public. It could be linked to a regulatory 'sandbox' to allow potential innovations in health data use to be tested.

Organisations in the health data landscape must **collectively engage on a political level**. This must communicate the barriers to, and requirements for, UK excellence in digital health to the Secretary of State for Health and Social Care, and across local councils, MPs, and government.

### 3.3. Building workforce capability, capacity, and confidence

**Clear guidance and best practice advice** is needed for health and data professionals, to raise their confidence to use data and engage with patients on health data issues.

There is a need **to invest in training** for professionals across healthcare, academia, and industry. This must happen at all levels: from specialised health informatics and data science for health leadership and research, to general data literacy for healthcare delivery. Greater data literacy for healthcare professionals could significantly improve the quality and completeness of data collection.

There must be **clear and effective regulation** in this space, to allow those who use data and digital technologies to understand, with certainty, what the rules and their responsibilities are.

Researchers and industry professionals need **trusted spaces to engage with health systems, regulators, and the public** on tentative and exploratory ideas, so that they can innovate and push boundaries responsibly and with confidence.

### 3.4. Enabling public conversation and public confidence

**Clear, consistent messaging** on the use of data in healthcare and research should be developed and shared across healthcare professionals, researchers, and public communities. This includes providing clarity around patients' rights and whether data is used on a consent basis or an opt-out basis.

We must **capitalise on the current climate** which brings data and health into the public interest. Examples driving this climate include: the forthcoming NHS App; the Secretary of State's drive for digital innovation in the NHS; innovation in technologies; increasing potential for patients to access their own data; and high-profile data issues like GDPR and Cambridge Analytica.

Mechanisms for **engagement at a local level** should be adopted across different communities, based on telling powerful stories about why data is important and who benefits. These should allow for **open, transparent conversations** about the use of data, while enabling individuals to provide feedback to those using it.

The power of **relationships within communities** to create a strong trusted dynamic should not be underestimated. Conversations should allow space to talk about failure or things going wrong, to enable others to learn from mistakes.

The right balance needs to be struck between **narratives at a local level and scripts or strategies** at a broader national level, to effectively amplify core messaging about the potential and benefits of data use.

Patients must be **empowered by data**, not imprisoned by it. They should be able to access their own health records, use data to manage their care, and be able to correct inaccurate data.

**Workforce conversation and confidence** are a critical part of this. Healthcare professionals are important gatekeepers, directly interacting with patients and holding their trust. If they lack information or confidence, this will prevent constructive conversations and possibly lead to greater public concern about data use.

### 3.5. Establishing an environment for innovation at pace and scale

There must be **effective business models** to allow researchers and industries to work with health systems at pace and scale. These models should reduce barriers to implementing innovations rapidly, without exacerbating health inequalities between different regions.

Small to medium sized enterprises must be **encouraged and supported** to engage with health systems. They are often highly innovative, but lack the resources to bring solutions to complex health systems.

The health sector should **learn from other sectors** which have incorporated data solutions – such as banking, retail, and travel.

The **value of data** created and held by the NHS and patients must remain with the healthcare system. The NHS and patients must see the benefits of solutions which use data, be able to account for data assets, and not be exploited.

A principle of **regional digital equity** should be adopted, to ensure fair value and benefit is spread across the entirety of the NHS, patients, and charities, regardless of where data is collected and technologies are developed.

## 5. Taking action

Addressing these steps requires specific action, and there is a range of existing work being done by the organisations represented at the closed session. This includes:

- HDR-UK developing a model for data use in research and invest in bioinformatics skills and training for the entire system.
- Building skills and digital literacy across NHS and with universities.
- Developing resources for computer science training.
- Making the case for continuing and expanding the work of *Understanding Patient Data*.
- Consultation on and public engagement with NHS England's recent [Code of Conduct for data-driven health and care technology](#).

These existing actions were discussed with a view to develop further collaboration on them. In addition, a set of further actions were established, which members present at the closed session agreed to pursue:

- |   |   |
|---|---|
| <b>Actions for guidance and best practice</b> | <ul style="list-style-type: none"><li>- Map out and synthesise the existing guidance on data use.</li><li>- Engage at political level to raise the need for joined-up approach to regulation and guidelines.</li><li>- Facilitate meetings to develop terms of reference for an advisory group around guidance and best practice.</li></ul> |
|---|---|



- Actions for public Dialogue**

  - Consider an op-ed with a clear message that there is collective effort to get these issues right.
  - Follow up with NHS App team on information provision about data use beyond care and the national data opt-out.
  
- Actions for data infrastructure**

  - Convene a group to push adoption of open standards, funding open development of critical infrastructure software for basic transfer of care data.
  - Discuss creation of a safe space so innovators can talk to regulators at an early stage, including on public engagement.
  
- Actions for workforce confidence and capacity**

  - Factor workforce engagement and digital literacy into engagement efforts, alongside public and patient engagement.
  - Map and synthesise identifiability terms defined in statute onto [Understanding Patient Data's identifiability spectrum](#).

The organisers of the meeting will consider convening a follow-up set of discussions with the closed group and others as needed, to further develop the actions required, and check progress against them.

## What's next?

This is a highly complex area, which will require significant sector and organisational siloes to be broken down and barriers to be removed, and collaboration between a huge range of organisations and bodies. This takes leadership, time, commitment, and resource to do well.

The strength of these discussions, as noted by those present, was that it brought together key health and data organisations who may not otherwise have had a space to converse. Not every participant agreed with all the points expressed above, but there was a clear commitment to ongoing dialogue and collaboration to ensure that the NHS and patients fully benefit from the potential of data for healthcare and research.

When healthcare and data is done "right", enormous benefits for patients and society can be realised. This is already evident through the examples raised at this event, and many more beyond them. The discussions of this event demonstrate a collective recognition of the barriers to getting healthcare and data right, and a commitment from a diverse range of people and organisations to overcome them.

---

Meeting notes compiled by Aidan Peppin

For any further queries about the meeting or to comment on this report, please get in touch at [hello@understandingpatientdata.org.uk](mailto:hello@understandingpatientdata.org.uk)

## Annex A – Representatives at the closed session

- Andrew Morris, Health Data Research UK
- Dame Fiona Caldicott, National Data Guardian
- Antony Walker, TechUK
- Mustafa Suleyman, DeepMind
- Andrew Elland, DeepMind
- Lorraine Jackson, Department of Health and Social Care
- Louise Wood, National Institute of Health Research
- Aisling Burnand, Association of Medical Research Charities (for evening of 13<sup>th</sup> only)
- Steve Wood, Information Commissioner's Office
- Andrew Davies, Association of British Healthcare Industries
- Tom Fowler, Genomics England
- Graeme Tunbridge, Medicines and Healthcare Products Regulatory Agency
- Indra Joshi, NHS England
- Iain Buchan, University of Manchester
- Hal Hodson, The Economist
- Dawn Monaghan, NHS England / Information Governance Alliance
- Lord Paul Drayson, Sensyne Health (for evening of 13<sup>th</sup> only)
- Lord Parry Mitchell, House of Lords
- Janet Allen, Cystic Fibrosis Trust
- Julian Huppert, Jesus College Cambridge
- Natalie Banner, Understanding Patient Data
- Nicola Perrin, Wellcome
- Aidan Peppin, Understanding Patient Data / Wellcome – taking notes

## Annex B – Open session agenda, Thursday 13<sup>th</sup> September

<b>Welcome and opening remarks</b> - Julian Huppert & Natalie Banner
<p><b>Panel 1: If data is the answer, what is the question?</b> Examining the realistic potential of data to drive improvements in health, care and research</p> <p><i>Pearse Keane, Moorfields Eye Hospital</i>                      <i>Andy Smith, useMYdata</i> <i>Sylvie Delacroix, University of Birmingham</i>                      <i>Janet Allen, Cystic Fibrosis Trust</i></p>
<b>Keynote speaker</b> - Dame Fiona Caldicott
<p><b>Panel 2: Barriers to better use of data</b> Identifying the key bottlenecks, conflicts and challenges that are undermining current efforts</p> <p><i>Aisling Burnand, AMRC</i>    <i>Sam Smith, medConfidential</i> <i>Justin Whatling, Cerner</i></p>
<b>Keynote speaker</b> - Norman Lamb, MP
<p><b>Panel 3: Where are we now?</b> Exploring exciting innovations that are making a difference to healthcare and research now</p> <p><i>Lord Paul Drayson, Sensyne Health</i>    <i>Carol Dezateux, UCL/OneLondon LHCRE</i> <i>James Peach, Medicines Discovery Catapult</i>    <i>Mustafa Suleyman, DeepMind Health</i></p>
<p><b>Panel 4: How should we get there?</b> Building a vision not just of how maximising the value of health data would all work, but also how it could be fair, right and well-governed.</p> <p><i>Iain Buchan, University of Manchester</i>    <i>Dawn Monaghan, NHS England</i> <i>Robin Hopkins, 11KBW law practice</i>    <i>Julia Powles, New York University</i></p>
Wrap-up remarks and close – Hal Hodson