New technologies that use patient data

Summary report of a workshop held on 24 April 2018

The Academy *of* Medical Sciences

The Academy of Medical Sciences

The Academy of Medical Sciences is the independent body in the UK representing the diversity of medical science. Our mission is to promote medical science and its translation into benefits for society. The Academy's elected Fellows are the United Kingdom's leading medical scientists from hospitals, academia, industry and the public service. We work with them to promote excellence, influence policy to improve health and wealth, nurture the next generation of medical researchers, link academia, industry and the NHS, seize international opportunities and encourage dialogue about the medical sciences.

Opinions expressed in this report do not necessarily represent the views of all participants at the event, the Academy of Medical Sciences, or its Fellows.

All web references were accessed in May 2018.

New technologies that use patient data

Summary report of a workshop held on 24 April 2018

Contents

Executive summary

Data-driven technologies are set to revolutionise the way healthcare systems operate. They have the potential to accelerate research and development, improve the efficiency of healthcare systems and allow patients to play a greater role in their health. However, the utility of such technologies is reliant on access to, and effective use of, patient data. In addition, ensuring that they collect, use and access these data in ways that meet the expectations of patients, the public and healthcare professionals is critical. As many of these technologies are new or emerging, principles to help guide use of patient data by these innovations will help to support their successful deployment in the health and social care system.

Therefore on 24 April 2018, the Academy convened a workshop of experts and key stakeholders drawn from across academia, digital health, data and pharma companies, the NHS, learned societies and the regulatory, funding and charity sectors to discuss the findings of a series of public dialogue workshops, carried out by Ipsos MORI. These workshops had examined public, patient and healthcare professionals views on the use of future data-driven technologies. Participants were asked to discuss the policy implications of the findings from the dialogue programme and consider how these could form principles for deployment of future data-driven technologies that use patient data in the NHS. Discussions focused on four aspects of patient data use by data-driven technologies: control; reliability; transparency; and accountability.

The key themes of discussion that emerged from the workshop were:

- The importance of **balancing access to data with ensuring appropriate safeguards for their use.** This includes establishing safe and flexible data access processes that are managed by the NHS, which would be reflected by NHS organisations acting as 'data stewards'. Ensuring appropriate access to data, and adoption of transparent data access processes may require a culture change to replace a risk-averse approach to one that builds a shared responsibility to provide access to and use of data for the benefit of patients and research.
- Ensuring that the NHS meets public and patient expectations that the **NHS will act as a responsible data steward, enabling new data-driven technologies and patient data to be developed and used for public good, whilst also protecting and respecting patient interests.** This role is particularly relevant in scenarios where specific consent is not practical and instead the NHS is trusted to act appropriately and needs to ensure that this trust is well placed.
- Establishing **appropriate governance mechanisms for making decisions about how data-driven technologies and patient data are used,** in line with any given or assumed consent or other legitimate authorisation. These should incorporate ongoing engagement with the public and patients.
- The need to **clarify the roles and responsibilities of healthcare professionals in the context of increasing availability and volume of information** for clinical decision-making, and their role in communicating information back to patients. In addition, there needs to be consideration of where accountability lies for decision-making, and errors, when using data-driven technologies. This includes the growing role of healthcare professionals in providing data to and from data-driven technologies into NHS systems.
- **Appropriate transparency is needed around how clinical decisions are made when using data-driven technologies**, and how uncertainty, accountability and clinical judgment are managed by clinicians if there is a lack of clarity around how data-driven technologies reach certain decisions ('black box systems'). A layered approach that provides differing levels of information on the flow of data across the whole pathway from collection to care delivery may help to support and reassure patients and healthcare professionals. This also **requires digital literacy to be built amongst healthcare professionals** to support their reliable use of data-driven technologies and role in contextualising the information received from data-driven technologies.
- **Transparency around how, why, where and by whom data are collected** is essential for providing context for patients, clinicians and commissioners to make decisions whether to use data-driven technologies. Transparency for patients should include what data are being collected about them, especially data which they may not be aware are being collected (such as during routine hospital monitoring).
- The **roles and responsibilities of patients in generating data and ensuring that they are fed back into the system** where possible, particularly if generated outside of the NHS.
- The need to fully understand **the reliability of patient data and the data-driven technologies** themselves (both how they use the data and the reliability of their data sources), including data integrity and provenance. This need for understanding includes looking at **ways to measure the reliability of data-driven technologies** and agreeing the levels of reliability and error that are considered acceptable, and to determine how to apply these acceptability criteria to the application of data-driven technologies.

Introduction

Professor Carol Dezateux CBE FMedSci, Professor of Clinical Epidemiology and Health Data Science, Queen Mary University of London, opened the workshop by explaining that data-driven technologies are rapidly emerging as new tools, technologies and services, built on the advances in ways that data can be collected and used. Progress in fields such as machine learning and natural language processing offers new opportunities for almost all sectors, and the medical research and healthcare sectors have been identified as likely to benefit significantly from their use. As the world's largest publicly funded health service, a particular strength of the NHS is its access to one of the largest single resources of patient data globally.[1] Therefore the UK has a unique opportunity to capitalise on the deployment of data-driven technologies in healthcare to improve patient care, optimise the healthcare system

[1] GovLab (2014). *The open data era in health and social care.* www.thegovlab.org/static/files/publications/nhs-full-report.pdf

and support important medical research. There have been numerous publications on the potential for data-driven technologies to drive innovation and improvement in the NHS.[2]

This workshop took place at a critical time with ongoing discussions on uses of patient data. Data privacy issues had been brought to public awareness by the debates on the use of data in the NHS, for example data access collaborations between NHS Trusts and private companies, and controversies sparked by a lack of transparency in these situations. These events highlighted the need for better transparency and public engagement around how data are collected, used, stored and accessed. Health data, especially, are often viewed as particularly sensitive. Professor Dezateux alluded to a number of recent studies on the topic that involved public dialogue, such as those by the Royal Society and Understanding Patient Data, noting that this dialogue is ongoing.[3,4] The development and implementation of data-driven technologies in healthcare is an important part of this broader discussion on patient data. Gaining a better understanding of the hopes and concerns of the public, patients and healthcare professionals (HCPs) on the ways in which data-driven technologies could alter the patient data landscape is required to ensure that change occurs in line with these expectations.

Finally, Professor Dezateux noted that this debate was occurring in a period of important regulatory change, such as the introduction of the General Data Protection Regulation (GDPR), which will fundamentally alter the relationship between individuals and data about them.[5] In addition, the NHS 'opt-out' scheme for use of identifiable patient data is a landmark in the ongoing discussion around how the NHS can, and should, use patient data for care and research.[6] These developments illustrate the dynamic environment surrounding patient data use and the need to understand how these changes will alter the implementation of data-driven technologies in healthcare.

Therefore, the Academy commissioned Ipsos MORI to carry out a programme of public dialogue with the public, patients and HCPs. The dialogue workshops took place in February 2018, and explored the views of these stakeholders on new data-driven technologies and their implications for the use of patient data (see Box: Ipsos MORI report). This following meeting then looked to explore the policy implications of the public dialogue programme, and potential principles for different stakeholders around the use of patient data by future data-driven technologies.[7]

---

[2] House of Lords Select Committee on Artificial Intelligence (2018). *AI in the UK: ready, willing and able?* https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf
[3] Royal Society and British Academy (2017). *Data governance: public engagement review.* https://royalsociety.org/~/media/policy/projects/data-governance/data-governance-public-engagement-review.pdf
[4] Understanding Patient Data, Carnegie Trust and Involve (2018). *Data for Public Benefit.* https://d1ssu070pg2v9i.cloudfront.net/pex/carnegie_uk_trust/2018/04/Data-for-Public-Benefit-REPORT.pdf
[5] The European Parliament and Council of the European Union (2016). *General Data Protection Regulation.* http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679
[6] Department of Health (2017). *Your Data: Better Security, Better Choice, Better Care.* https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/627493/Your_data_better_security_better_choice_better_care_government_response.pdf
[7] Ipsos MORI (2018). *Future data-driven technologies and the implications for use of patient data Dialogue with public, patients and healthcare professionals.* https://acmedsci.ac.uk/file-download/6616969

Ipsos MORI report

In 2017, the Academy of Medical Sciences commissioned a programme of public dialogue by Ipsos MORI on *'Future data-driven technologies and the implications for use of patient data'*, with input from Understanding Patient Data (UPD). The dialogue involved six workshops across three locations in the UK with the public, patients and HCPs. Participants discussed case studies representing the different ways data could be gathered, analysed and applied by using modern technologies in healthcare, which were used as a platform to establish the opinions of participants and general themes of discussion. The objectives of the dialogue were to gather awareness, expectations, aspirations and concerns around future technologies, which would require patient data in the NHS to be accessed, analysed, or linked with innovative types of data and/or analysis, in order to augment clinical diagnoses, management and research.

Harnessing the 'fourth revolution' in healthcare

Professor Andrew Morris CBE FRSE FMedSci, Director of Health Data Research UK, emphasised the opportunity to harness data science for biomedical research and healthcare. He highlighted the importance of capitalising on the *'fourth industrial revolution'* in the UK – the exponential increase in available data and technologies that harness them – to drive productivity through innovation and research.[8] There are a number of drivers behind adoption of data-driven technologies in the NHS. Firstly, politicians and NHS leaders have recognised the vital role that digital technologies and greater use of patient data will play in supporting the NHS in times of increasing demand.[9,10] Parallel to the growth of the NHS is the vast expansion of clinical data arising from clinical practice and trials. Professor Morris highlighted that ensuring clinical utility of trials in rare diseases or heterogeneous conditions will require larger patient cohorts and databases, which can be enabled by linking and expanding datasets. Finally, the continued increases in computing power and data generation and storage capabilities will further drive the development of data-driven technologies.[11]

Professor Morris also highlighted the need for *'learning health systems'* where healthcare, data and informatics, and research are embedded and integrated in the system to allow ongoing improvement to healthcare through continuous learning. There are a number of case studies of

[8] Boston Consulting Group (2017). *Is UK Industry ready for the Fourth Industrial Revolution?* https://media-publications.bcg.com/Is-UK-Industry-Ready-for-the-Fourth-Industrial-Revolution.pdf

[9] Deloitte University Press (2017). *Achieving Digital Maturity.* www2.deloitte.com/content/dam/insights/us/articles/3678_achieving-digital-maturity/DUP_Achieving-digital-maturity.pdf

[10] https://www.gov.uk/government/publications/industrial-strategy-the-grand-challenges/missions

[11] Moore GE (1965). *Cramming more components onto integrated circuits.* Electronics. **38(8)**

successful application of data science to healthcare systems in this way, including assessing the effects of the smoking ban on childhood asthma and employing algorithms to identify malignant skin cancers.[12,13] He identified two key characteristics of successful learning health systems: comprehensive information infrastructures and investment in innovation and research capacity. There are a number of considerations to meet these requirements including:

- The **complex structure of the NHS**, which is a diverse mix of public and private services of varying levels of integration and with regional differences in services and infrastructure.
- Addressing a **lack of digital maturity** within the NHS.[14]
- The significant increase in the **volume of data**.
- Differing **data quality and standards**. For example, a lack of standardisation for activities such as prescriptions makes deriving learnings from healthcare difficult.
- The need for **interdisciplinary communication** across HCPs, software developers, patients, NHS trust managers etc.
- **Changes to patient recruitment** and contact through technological developments.
- The need to **maintain trust** in NHS handling of data.

Finally, Professor Morris explained that Health Data Research UK was established in order to improve health through data science. A need was identified for coordination between major UK funders to establish large, multidimensional datasets and so Health Data Research UK will create regional centres of excellence ('data hubs') within the NHS. They will also partner with the National Institute for Health Research (NIHR) to train health data scientists.

---

[12] Millett C, *et al.* (2013). *Hospital Admissions for Childhood Asthma After Smoke-Free Legislation in England.* Paediatrics **131(2)**, e495.
[13] Esteva A, *et al.* (2017). *Dermatologist-level classification of skin cancer with deep neural networks.* Nature **542**, 115.
[14] Academy of Medical Sciences (2017). *Digital maturity of health and social care systems.* https://acmedsci.ac.uk/file-download/77817565

Principles

New data-driven technologies will affect every aspect of the patient data pathway; from data acquisition and collection, to analysis and usage, how data are linked between organisations and how they are stored or deleted after use. Each stage of this pathway involves multiple stakeholders. Communication between these parties and a joined-up approach between different stages is needed to ensure a reliable, consistent and transparent approach to the use of patient data by data-driven technologies.

Participants considered the following areas at each stage of the data journey to identify potential principles that could help to guide decision-making for organisations and individuals involved in aspects of the patient data pipeline:

- **Reliability:** how do we ensure the reliability of organisations involved with patient data, the data-driven technologies/products, the data themselves and the evidence that they generate?
- **Transparency:** how can we make the *'who, where, why and how'* clear for the collection, access, storage and use of data by technologies in healthcare?
- **Control:** who has control over data at each step, what is the degree of control and when is control gained and lost?
- **Accountability:** who is accountable for technologies and data at each step, including their development, testing and deployment?

Who is in control?

The role of the data controller

The GDPR definition of 'data controller' is the body who '*determines the purposes and means of processing personal data*'.[15] There was consensus amongst attendees that the use of the term 'control' – used in this instance as a legal term – may be unhelpful as it could imply that the best or strongest means of management is to protect data by controlling and restricting access as far as possible. Participants felt that the remit of data controllers within the NHS is to facilitate responsible data sharing when it is of benefit for healthcare services, care or research. As such, it was suggested that 'data steward' better reflects the dual responsibilities of the NHS to ensure security of patient data and also to facilitate proactive data use in a responsible and trustworthy manner for public good. This was felt as particularly important given the high level of public trust in the NHS to use patient data appropriately. Stewardship was also thought to better reflect the increasing public empowerment over personal data under the GDPR, particularly with regard to right to access, and the legal responsibilities of data controllers to ensure the integrity of personal data.

Data use and access sharing

Participants were clear on the need for flexible and dynamic processes to manage the access of data within the NHS and between the NHS and other partners. At present, data access is often on an *ad-hoc* basis between individual companies and NHS Trusts or patient bodies. However, processes to map and log the movement and access of data in the NHS and with external partners will be critical for the NHS to fulfil its responsibilities as a data steward. One principle that arose from these discussions was the need for data to be returned to the NHS once it has been used by third parties, or having a robust mechanism for confirming their deletion or anonymisation to ensure that the NHS is the overall custodian of the patient data and responsible for long-term control. This requires oversight and a potential framework to ensure this.

Engaging the public in governance

Current data governance laws, such as the GDPR, will form the basis of future regulation of data-driven technologies and data use. Further to this, participants advocated the value of developing 'Codes of Practice' or governance panels that provide a framework specifically for control of healthcare data and application of data-driven technologies in healthcare systems. Current governance review panels such as the National Data Guardian's Panel and Scotland's Public Benefit and Privacy Panel for Health and Social Care were highlighted as examples of good practice, with proportionate governance and involving lay, legal and clinical representation. Participants highlighted that these committees are a publicly accepted way of making decisions on when data are accessed from outside of the NHS on behalf of patients, and they need to be co-designed by patients, HCPs and lay members to ensure that data-driven technology implementation aligns with public expectations. However, it was noted that public awareness of these bodies was lacking and that raising this awareness is critical to preserving trust in the NHS.

Meeting expectations on safeguarding and access

There was a general understanding that use of patient data should strike a balance between appropriate safeguarding and enabling access. In agreement with the Ipsos MORI report, participants stressed that the majority of patients and the public are happy to have data about

---

[15] Official Journal of the European Union. *General Data Protection Regulation* (2016). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:32016R0679

them used for research as long as the uses and benefits of doing so are made clear.[16] This extends beyond personal benefit. For example, it was noted that there can be very high recruitment of palliative care patients into research programmes when the reasons for the research are clearly explained, even though there is limited direct benefit for the participant. Despite this, it was suggested that controllers are often too risk-averse to actively use data; this is possibly due to cultural barriers, uncertainty of regulation and fear of failure to appropriately safeguard. This was felt to be true for both the NHS and private sector, but also between public sector organisations (e.g. NHS and wider social care). It was noted that this risk aversion does not align with the 'social contract' obligations of the NHS.

Patient control, choice and consent for the use of data by data-driven technologies

When consenting to the use of patient data for research or service planning, there was strong agreement that in many cases, it is impractical and unrealistic to require patient consent on a granular basis where it is repeatedly sought every time data are used outside of individual care. It was suggested that broad consent would be more effective. The new NHS data opt-out programme introduces a consent model for the use of identifiable patient data.[17] However, for this more general model of consent around patient data and data-driven technologies, participants cautioned against the *'illusion of choice'* where either terms for data-driven technologies and data use are too confusing to allow informed choice, or options are so simplified and standardised across the NHS that the choice offered to patients does not reflect their more detailed opinions on how they are used. Instead, patients should be given clear choices that enable them to distinguish between data use with data-driven technologies for different purposes or levels of 'risk' (i.e. anonymised routine clinical measurements vs. identifiable photographic data). There is potential to draw on learning from other sectors such as banking, in how to effectively communicate these choices.

It was also noted that patient opinions on data access may vary according to aspects such as their medical condition or socio-cultural environment. For example, patients may have preferences for the types of research that are conducted using data about them. Therefore, participants suggested that a 'one-size fits all' model is unfeasible, instead favouring a layered approach to giving patients control over the level of access to data about them, and whether these data are anonymised or identifiable, for different research or planning purposes.

However, it was noted that increasing the requirements for patient consent through such a model could prevent useful applications of these data for research and planning, and may be complex to implement. In particular, participants recognised that that a poorly implemented consent model could severely undermine research and the NHS's wider capabilities as a learning healthcare system by increasing gaps in data sets. As such, it was felt that using a generalised consent model may be a necessary trade-off to ensure maximum research and public health benefit from patient data, which would assist the improvement of care over the longer-term. The risks of using such a model could be appropriately managed in a similar manner to how other risks, such as human error, are managed and communicated to patients.

---

[16] Ipsos MORI (2018). *Future data-driven technologies and the implications for use of patient data Dialogue with public, patients and healthcare professionals.* https://acmedsci.ac.uk/file-download/6616969
[17] https://digital.nhs.uk/services/national-data-opt-out-programme

Accountability

Delivering the 'social contract' of the NHS

There was a strong consensus that the NHS is the steward of a 'social contract' between the public and the healthcare system.[18] Part of this social contract is an implicit agreement that the NHS has a responsibility to use patient data and new technologies for public good whilst also ensuring privacy and security. This encompasses uses of patient data beyond individual care to applications in research with a broader public health focus or longer-term aims. Anecdotal accounts from participants and perspectives from the Ipsos MORI report and previous research suggest that patients' views are well-aligned with this.[19] As the use of data-driven technologies in the NHS evolves, participants agreed that these new technologies should be used in a way that fulfils this role of the NHS.

Participants also noted that a push for digital maturity in healthcare systems could prioritise innovation over utility for patients. Therefore, participants reiterated the need for a framework for decision-making in the NHS around the acquisition and use of patient data and the implementation of associated data-driven technologies. While new governance may be difficult to establish and too slow to cope with the dynamic pace of change in this field, generalised guidelines could be of use.

The evolving role of the healthcare professional

Participants agreed that the increasingly vast amount of data available to HCPs raises important questions about their shifting responsibilities. There was debate as to whether HCPs should be responsible for making patients aware of all incidental findings; some thought this was a patient 'right' and others felt it impractical. Concern was raised that the rise in information available to HCPs may lead to increased 'negative action' where pressure to respond to previously unavailable data could lead to relatively unnecessary interventions that may do more harm than good. It was noted that the Hippocratic Oath to *'first do no harm'* will become only more relevant in an age of increasing potential for medical interventions.

There was general agreement that there remains a need for HCPs to put data in context and make clinical judgements. The increase in both the quantity and sources of data would not alter this responsibility but may make it more challenging. Indeed, participants anticipated the value of software in summarising information and there is a level of accountability with the software developers to ensure that HCPs are provided with relevant information. It was not clear where the accountability lies for clinical errors in this situation (whether with the HCP, the developer or both) and so this is a key area for continued discussion. This could fundamentally change the accountability of HCPs. However, it was noted that for existing technologies, such as medical imaging, the risks associated with clinical errors are accounted for and managed, and that this may not be fundamentally different for future technologies.

Does more patient control mean more accountability?

It was broadly agreed that the introduction of certain data-driven technologies, alongside the new regulatory landscape, could lead to patients having more control over data about them. For example, patients will need to decide whether data collected via data-driven technologies

---

[18] The Academy's 2015 report on *'Exploring a new social contract for medical innovation'* explored the role of the NHS in carrying out its social contract to meet the expectations of different stakeholders on innovations by considering their perspectives on the value/benefits and risks of innovation. https://acmedsci.ac.uk/file-download/38377-5673dcc2f036b.pdf

[19] Ipsos MORI (2016). *The One-Way Mirror: Public attitudes to commercial access to health data.* https://wellcome.ac.uk/sites/default/files/public-attitudes-to-commercial-access-to-health-data-wellcome-mar16.pdf

such as wearables are provided to their HCP. Equally, patients should have the control and choice to use advice given by HCPs as much as they wish, and maintain autonomy over their lifestyle. For example, patients should have the right to choose to make unhealthy lifestyle decisions without disclosing this to HCPs. Whilst these scenarios are new, the issues around disclosure are not. It was felt that patients would still be responsible for providing their HCP with accurate information as best they could, with the end decision as to what information and how much is provided resting with the patient.

Patient choice about how data are shared can also apply to sharing across different parts of the NHS. It was questioned whether services require access to a patient's entire clinical record, for example whether a physiotherapy clinic requires sexual health information. There was agreement that different services may only require access to specific parts of a patient's electronic health record (EHR) that is relevant to the service, but it can be difficult to determine the relevance of information in an EHR, especially when considering multimorbidities or polypharmacy. Patients may assume that all of the data about them are available across the NHS and this could compromise care or lead to errors, such as a patient neglecting to tell HCPs about a medicine allergy because they assume it is known. There is also a risk that this could exacerbate existing issues of data silos within the NHS, with access to patient data between NHS centres varying on an individual basis.

Despite the challenges, it was noted that increased patient control also represents a significant opportunity to engage people in their healthcare. East London NHS Trusts were highlighted as an example of where improvement to EHRs alongside a cultural shift in primary care have helped to promote healthcare as a shared two-way dialogue between patients and HCPs.

Reliability of data-driven technologies and the data journey

Reliability of patient data

Participants were optimistic about the potential for data-driven technologies to improve the quality of data and generate new types of data for integration into EHRs. It was highlighted that healthcare data degrades quickly, with information becoming unreliable over time due to its relevance to the context in which it is captured at the time. Thus, technologies that enable up-to-date information would improve assessments and increase the potential for longitudinal analyses of health. This idea of data integrity is key to understanding its reliability. Equally, broader data such as lifestyle data was recognised as extremely useful for determining clinical outcomes, and future data-driven technologies have the potential to incorporate this information into EHRs.

However, participants also recognised that new sources of patient data from data-driven technologies raise questions around the reliability of data from these new technologies, for example from wearable devices, and how this can be assessed. The need to retain information about the provenance of patient data collected outside of primary or secondary healthcare was stressed, and to put mechanisms in place for assessing – or auditing – their reliability. It was recognised that varying quality of data is not a new challenge in the NHS with varying quality based on the accuracy of the technology/test, ease of use, skill of the HCP etc. However, knowing where the data came from is a prerequisite to assessing its reliability.

Reliability of analyses and data-driven technologies

There was a consensus that the introduction of data-driven technologies in healthcare will bring both opportunities and risks in terms of reliability. While there was broad agreement that current governance may be sufficient to cover general data use, it was suggested that there are

important areas around ensuring reliability of data-driven technologies that demand further attention. For example, the importance of finding ways to assess the reliability of machine learning algorithms, particularly in assuring that there is no bias in the system. There was also discussion on whether reliability could *itself* be reliably measured, and how reliability was defined.

A distinction was drawn in public dialogues around privacy and use of personalised and depersonalised data. Participants emphasised the importance of this distinction, and noted that it is critical to understanding when and how patient data can be reliably depersonalised as this will affect the acceptability of using these data.

It was also posed that adoption of data-driven technologies may lead to the creation of new *types* of unreliability, but may or may not increase the overall *amount.* For example, the current system has a level of managed unreliability (such as patients omitting information or not showing up for appointments) that data-driven technologies could help to reduce, whilst simultaneously introducing other forms of unreliability as outlined above. Balanced risk-benefit analyses may be useful in deciding the trade-offs and overall value of introducing a technology, as there is a risk of limited adoption if the system is too averse to new risk.

Reliability of individuals and organisations using data-driven technologies

Participants noted that the reliability of data-driven technologies is also dependent on the individuals using them. Ensuring that patients and HCPs are using data-driven technologies correctly (whether in roles as data providers or users for healthcare delivery) is important for building trust in the quality of data generated. This point was also made in relation to emerging machine learning and AI-based clinical tools; how can a clinician assess the reliability of a technology when they do not, and perhaps cannot, understand the decision or the process? Here there was agreement on the need for appropriate training for HCPs to allow them to effectively use data-driven technologies, and that the data-driven technologies themselves should be designed to operate in an easily explainable manner. It was suggested that HCPs should be sufficiently trained to have a general understanding of new data-driven technologies and to play an active role in co-designing these systems for use. It was also noted that developers themselves should understand the reliability of the data sources used to develop data-driven technologies and that mechanisms for assessing whether developers are accounting for the quality of data should factor into decisions to use their data-driven technologies.

Building trust through transparency

What data are being collected?

Participants stressed the need for full clarity around any integration of patient data derived from data-driven technologies into patient records, such as those obtained from remote monitoring. There was concern over 'context collapse', where patients may consent to the collection or use of data without fully understanding all the contexts in which it will be collected or used. Additionally, transparency around the types of data that data-driven technologies are collecting is important, especially as many technologies will be collecting information (e.g. location, activity) that is surplus to their specific remit. Despite these concerns, it was highlighted that this does not preclude the widespread use of technologies but that transparency around the context in which they are used is key, and patient views should be accounted for when considering the appropriateness of using such technologies. An example was given where patients with chronic obstructive pulmonary disease were able to leave hospital early if they

agreed to be monitored with a remote device.[20] This allowed the patient to return home whilst also saving NHS time and resources and represents a compromise between the patient and health service that benefits both parties, but patients retain the choice to refuse monitoring and remain in hospital. It was agreed that in other circumstances, such compromises may be less favourable to the patient where they involve aspects such as remote 'surveillance', either in terms of the benefits or the privacy rescinded. As a result, the use of technologies that have these risk-benefits is often context-specific, and therefore public engagement and patient choice are critical to ensuring that these technologies are implemented appropriately.

Clarity around decision-making

There was agreement that algorithms and AI have the potential to blur or limit transparency in clinical decision-making. Firstly, there may be difficulty in understanding how these data-driven technologies reach a decision if this is not clearly available for the user – described as a 'black box' way of working. Therefore, there is a need to understand what qualifies as sufficient transparency for a user such as a clinician to be able to comfortably use a data-driven technology in decision-making. New data-driven technologies also need to be understandable for the user to reduce the grey areas in decision-making, whether the users are HCPs, patients or both. There is a growing field around audit and 'understandability' in data science and this has important implications for healthcare technologies. In addition to ensuring that design of data-driven technologies facilitates ease of interpretation, it was also suggested that patients need to be appropriately informed of the use of data-driven technologies in decision-making in their care. Therefore, HCPs need to be trained to effectively communicate the use of data-driven technologies to patients. This should extend to data-driven technologies that collect data outside of a clinical setting, so that the patient can take responsibility for the generation and communication of these data, and the HCP can take responsibility for interpreting the data taking into account their limitations or expected reliability.

Generally, it was recognised that whilst full transparency of how these systems operate will be unfeasible in many cases, a 'layered information' approach that provides information around the flow of data would be important. Developers could be transparent about these data flows that underpin data-driven technologies and provide this information to service commissioners, HCPs and patients in an understandable manner. Through these approaches, trust would be nurtured between developers, commissioners and HCPs and patients.

The chain of transparency

In order for patient data to be used in a trustworthy manner, it was agreed that the origins and distribution of data need to be made clear. Once patient data are in the healthcare system, patients need to be aware of who is holding data about patients, how these data are being stored and how they are used, both for direct care and other purposes. As noted in the Ipsos MORI report, there is a lack of understanding in the public sphere about the structural complexity of the NHS and how its component parts interact, and therefore lack of clarity around where control lies at different stages of the patient data journey.[21] Participants discussed this in terms of patient and public understanding of how data are linked and shared within the NHS, and it was felt that public understanding of these areas was often poor.

Transparency on the use of patient data for research was also noted as a key area. Examples were given of systems that provide clarity and transparency on the use of patient data for research, such as the Health and Care Research initiatives in Wales, which provides information on research projects that are using NHS data.[22] However, it was stressed that similar UK-wide

---

[20] www.digitalhealth.net/2014/09/scottish-copd-patients-united4health/

[21] Ipsos MORI (2016). *The One-Way Mirror: Public attitudes to commercial access to health data.* https://wellcome.ac.uk/sites/default/files/public-attitudes-to-commercial-access-to-health-data-wellcome-mar16.pdf

[22] https://www.healthandcareresearch.gov.wales

initiatives do not exist. Additionally, systems like the NIHR Maudsley Biomedical Research Centre's Clinical Record Interactive Search (CRIS) were used as exemplars to illustrate the successful use of de-personalised patient data for NHS research. CRIS links de-personalised datasets from within the Trust for use in research.[23]

Participants discussed the impact that good governance, transparency and trustworthiness can have in reducing the requirement for individualised consent for uses of health data for different purposes. Part of this process is the need for improvements to how patients are informed on the outcomes of research involving data about them. Engaging patients in conversations about the data about them that is collected or held by the NHS, to enable them to make informed decisions about the uses of these data, was suggested as beneficial.

[23] https://crisnetwork.co/

Identifying responsibilities

The patient data journey involves many different stakeholders at varying points across the acquisition, analysis, delivery and linkage of data. Identifying where responsibilities lie at each of these different points is a key step to ensuring informed, consistent and trustworthy practice. It will also enable organisations and individuals to play a greater role in managing accountabilities to the benefit of all.

Throughout the workshop, participants were asked to consider which stakeholders they thought were responsible for taking ownership of the principles raised in discussions. It was recognised that any broad principles will have ramifications across the patient data pipeline, and therefore coordination between stakeholders will be crucial. However, it is important that these stakeholders are engaged relative to their impact in different areas, and made aware of particular responsibilities.

Responsibilities across the data journey

The role of the empowered patient

Patients are playing an increasing role in their healthcare and new data-driven technologies such as wearables will further encourage their active engagement in how data are collected and used. For example, patients will be responsible for deciding whether to provide clinicians with

data obtained from such devices. This requires the patient to have a level of understanding and digital literacy to allow them to collect the data appropriately and then share or communicate these data effectively with a clinician. Such responsibilities already exist to some extent, with patients able to offer information to a clinician that may not appear precisely relevant to their direct care.

Responsibilities for decision-making support by data-driven technologies

It was agreed that in order for patients to exercise these new responsibilities in a meaningful way, systems should be put in place to prevent harm. The accuracy and reliability of data-driven technologies need to be well-established and ensured as far as possible, so as to be useful for patients and HCPs. Both the NHS and regulators are responsible for deciding on the standards for these technologies (including when authorisations are required during development and testing), and the developers themselves for adhering to these. In the context of data-driven technologies used by HCPs, developers have an additional responsibility to make these technologies sufficiently understandable for HCPs – particularly if they may direct clinical decision-making. Participants were less clear about where the responsibility of HCPs lies in ensuring the reliability of new data sources used in clinician decision-making. However, assuming that reliability and accuracy of this data is assured, it was felt that HCPs have a responsibility to make full use of information provided to them. Participants also highlighted the need for developers to be transparent about what types of data are recorded by their data-driven technology.

The NHS has a responsibility for the ethical deployment of data-driven technologies that use patient data; a formal mechanism could be established which builds upon current systems that evaluate the utility of data for research. Participants stressed that these need to give careful consideration to avoiding a 'context collapse' and the 'illusion of choice' in how patient data are collected both in and outside of hospital. It was also suggested that HCPs may have a role in informing patients of the advantages and limitations of data-driven technologies to enable them to make informed decisions about data sharing and whether to use these technologies when a choice is available to them.

Data governance and stewardship

The NHS as a steward

As an agreed steward for patient data, the NHS has a clear responsibility to use the data it controls for public good; be that through quality monitoring, research, public health initiatives or individual care. Given the complex structure of the NHS and the multiple organisations involved, participants recognised the need to explicitly identify its responsibilities and those of third party data users. It was agreed that NHS data should continue to be used in line with public values, particularly in cases where explicit consent is not sought from patients, and this will require ongoing dialogue with patients, HCPs and the wider public. Additionally, examples of research initiatives within the NHS, such as CRIS, were highlighted as successful ways to represent the views of patients and the public in decision-making through lay membership.[24] Participants agreed that care for the security and sanctity of data is the responsibility of third parties when in their hands, although the responsibility for due diligence around this was debated. In line with the NHS acting as the long-term steward of data, there was some

---

[24] https://crisnetwork.co/

agreement that third parties should be obliged to offer something in return to the NHS for the use of NHS-held data.

Data management

Participants recognised the role of providers of data management systems. The use of personalised data, and mechanisms for effectively depersonalising data, were identified as worthy of scrutiny, as public dialogue has identified different perspectives on how these two data types are used. Providers of software systems for data de-personalisation are responsible for ensuring that identifiable information about patients is kept private, and this is critical to nurturing trust around data use between the public, NHS and third parties.

To encourage a two-way conversation around data use, participants suggested that patients needed to be better informed about the benefits of sharing data about them (e.g. for research and planning), especially in circumstances where the data will not benefit their direct care but instead benefit wider society. This would allow patients to make informed decisions about the use of data about them held in the NHS. There was discussion around where the responsibility for this lies, with suggestions that HCPs have a responsibility to 'make the case' for use of patient data for research and third parties to convey the outputs of data use.

Conclusions

The digitisation of the NHS and the associated technologies that will accompany it are poised to drive transformational changes in the way that healthcare services are delivered, patients interact with the health system and research is conducted. Once a digital, linked-up NHS system is achieved, these technologies are likely to emerge rapidly, and it is imperative that the system is prepared to allow these technologies to be robustly critiqued for their effectiveness and reliability as well as their impact on the NHS workforce and patients. In addition, their adoption and use should be guided by principles that embody the views of patients and the public, and acknowledge the social contract that the NHS is held to as a body trusted by the public to make decisions for the benefit of society. The discussions in this workshop explored these areas and built upon the findings of the Ipsos MORI report to determine a number of areas where guiding principles will be useful in steering the development and use of data-driven technologies by stakeholders across the health space. These principles include:

- The public are highly supportive of data use for research, as shown in the Ipsos MORI report and previous public dialogues. They are open to the opportunities new technologies bring to further this agenda.
- The stewardship of patient data by the NHS includes its obligation to use data as effectively as possible for social good, alongside the need to keep data safe and secure and to protect patients' privacy.
- The NHS should seek to use data and data-driven technologies proactively in areas where there is clear benefit either for individual patients or society as a whole.
- As the use of patient data ultimately is stewarded by the NHS, any advances in products or technologies derived from the data should be fed back into the NHS and the products available equitably across the nation.
- Data management technologies are an important part of ensuring that data are used effectively. In particular, methods to track the location and travel of data will help ensure transparency.
- There is a need to ensure that data-driven technologies are implemented in a way that retains or strengthens the HCP-patient relationship where possible.
- Increased patient engagement in the acquisition and use of data about them could be used to encourage a two-way dialogue between HCPs and patients.

- The increase in patient data available to HCPs will require technologies that can effectively summarise and visualise data, but HCPs should still be responsible for putting this in a clinical context.
- The provenance of patient data, particularly when influencing clinical decisions, needs to be made transparent, both to HCPs and patients. Alongside this, methods for assessing the quality of this data should be developed.

Annex 1: Agenda

Tuesday 24 April, 08.30-14.00
Academy of Medical Sciences, 41 Portland Place, London, W1B 1QH

08.30-09.00	**Registration and refreshments**
09.00-09.10	**Welcome and introduction** Professor Carol Dezateux CBE FMedSci, Professor of Clinical Epidemiology and Health Data Science, Queen Mary University of London
09.10-09.30	**Health data science, clinical care and new technologies: where do we want to be in 10-20 years' time?** Professor Andrew Morris CBE FRSE FMedSci, Director, Health Data Research UK
09.30-09.50	**Overview of report and findings from public dialogue** Sarah Castell, Head of Qualitative Methods and AMS project lead, Ipsos MORI
Panel discussion of IPSOS Public Dialogue report *Chaired by Professor Jonathan Montgomery, Professor of Healthcare Law, University College London*	
09.50-11.00	**Panel discussion** Discussants: • Ms Hilary Newiss, Chair, National Voices • Professor Corri Black, Co-Director, Aberdeen Centre for Health Data, University of Aberdeen • Dr Charles Gutteridge, Chief Clinical Information Officer, Bart's Health NHS Trust • Professor Mandy Chessell CBE FREng, Distinguished Engineer, IBM
11.00-11.20	**Tea and coffee**
11.20-12.30	**Break out groups: What principles should guide data access and use, for and by new health technologies?** Participants split into four groups in a world café format to discuss four areas in turn: **reliability; transparency, control; and accountability.** Each of these were considered in the context of data collection and acquisition, analysis, delivery and linkage and legacy.
12.30-13.00	**Lunch**
Discussion: Developing a framework for principles to guide data access, acquisition, and generation and/or use for and by new health technologies *Chaired by Professor Jonathan Montgomery, Professor of Healthcare Law, University College London*	
13.00-13.50	**First principles: what are the core aspects to consider, who is accountable and how can we 'operationalise' these principles?** Break-out group chairs presented the consolidated outputs from each theme and reviewed with audience. In particular, looking at who is responsible for the different principles; how can we be doing these; and what and who are the enablers?
13.50-14.00	**Summary and next steps** Professor Carol Dezateux CBE FMedSci, Professor of Clinical Epidemiology and Health Data Science, Queen Mary University of London
14.00	**Close**

Annex 2: Attendee list

Co-chairs
Professor Carol Dezateux CBE FMedSci, Professor of Clinical Epidemiology and Health Data Science, Queen Mary University of London*
Professor Jonathan Montgomery, Professor of Healthcare Law, University College London and Chair of the Health Research Authority

Speakers and panellists
Professor Corri Black, Co-Director, Aberdeen Centre for Health Data Science, University of Aberdeen
Ms Sarah Castell, Head of Qualitative Methods, Ipsos MORI
Professor Mandy Chessell CBE FREng, IBM Distinguished Engineer, IBM Master Inventor, IBM*
Dr Charles Gutteridge, Chief Clinical Information Officer, Barts Health NHS Trust
Professor Andrew Morris CBE FRSE FMedSci, Director, Health Data Research UK
Ms Hilary Newiss, Chair, National Voices

Participants
Dr Adrian Baker, Policy Manager, British Heart Foundation
Dr Kambiz Boomla, Clinical Senior Lecturer, Queen Mary University of London*
Ms Jen Boon, Senior Policy Adviser, Office for Life Sciences
Dr Helen Bulbeck, Patient advocate and Coordinating Group member, use MY data
Dr Tony Calland MBE, Joint Vice Chair, HRA Confidentiality Advisory Group
Mr Matt Case, Policy Adviser, Cancer Research UK
Ms Victoria Cetinkaya, Senior Policy Officer, Information Commissioners Office
Ms Vicky Chico, Data Policy Advisor, Health Research Authority
Dr Robert Chipperfield, Medical Affairs Director for Primary Care, Merck Sharp and Dohme Limited
Mr Stephen Critchlow, Founder and CEO, Evergreen Life
Professor Chris Holmes, Statutory Professorship in Biostatistics, University of Oxford/Nuffield Department of Medicine
Ms Marie Kane, Chief Operating Officer, NorthWest EHealth
Mr Brendan Krause, Vice President for Europe, Optum Labs
Dr Natasha McCarthy, Head of Policy, Royal Society
Dr Amara Nwosu, Medical doctor and Clinical Lecturer in Palliative Care, University of Liverpool*
Ms Reema Patel, Programme Manager – Data Ethics & AI, Nuffield Foundation
Ms Nicola Perrin, Head, Understanding Patient Data
Professor Ronan Lyons, Clinical Professor of Public Health, Swansea University
Mr Jonathan Sellors, Legal Counsel and Company Secretary, UK Biobank*
Professor Liam Smeeth FMedSci, Professor of Clinical Epidemiology and Head of Department, London School of Hygiene and Tropical Medicine
Ms Juliet Tizzard, Director of Policy, Health Research Authority
Dr Darren Treanor, Diagnostic Digital Pathology Lead, Royal College of Pathology
Dr Mary Tully, Reader in Pharmacy Practice; Director of Public Engagement for Connected Health Cities, University of Manchester*
Dr Rhoswyn Walker, Head of Informatics Research, Medical Research Council
Mr Gary Warner, Managing Partner, Pharmoutcomes
Dr Philippa Westbury, Senior Policy Advisor, Royal Academy of Engineering
Dr Frank Wiegand, Vice President Global Medical Affairs & Market Access, Neurosciences, Janssen
Professor John Williams CBE, Professor of Health Services Research, Swansea University
Dr Eva Woelbert, Data Science Portfolio Manager, MQ: Transforming mental health
Professor Dame Til Wykes DBE, Professor of Clinical Psychology and Rehabilitation, King's College London*

Academy staff
Ms Liberty Dixon, FORUM Policy Manager, Academy of Medical Sciences
Mr James Drew, Policy Intern, Academy of Medical Sciences
Mr Nick Hillier, Director of Communications, Academy of Medical Sciences
Mr Alberto Lazari, Policy Intern, Academy of Medical Sciences
Dr Rachel Quinn, Director of Medical Sciences Policy, Academy of Medical Sciences
Ms Holly Rogers, Communications and Engagement Manager, Academy of Medical Sciences
Dr James Squires, Policy Officer, Academy of Medical Sciences
Dr Naho Yamazaki, Head of Policy, Academy of Medical Sciences

* indicates a member of the project Steering Group

Annex 3: Glossary

- **AI:** Artificial intelligence
- **CRIS:** Clinical Record Interactive Search – this is a system that was developed by the NIHR Maudsley Biomedical Research Centre to enable secure access to anonymised information extracted from the South London and Maudsley NHS Foundation Trust EHR system.
- **Data-driven technology:** Technologies that acquire, use or analyse patient data.
- **DPA:** Data Protection Act
- **EHR:** Electronic health records
- **GDPR:** General Data Protection Regulation
- **HCP:** Healthcare professional
- **Machine learning:** an application of artificial intelligence that allows systems to automatically learn and improve from experience without being explicitly programmed.
- **Natural language processing**: a branch of artificial intelligence that helps computers understand, interpret and manipulate human language.
- **NIHR:** National Institute for Health Research
- **UPD:** Understanding Patient Data

The Academy of Medical Sciences