June 2021
APPG AI Evidence Meeting

BIG
INNOVATION
CENTRE

AI in Public Health:

How has AI helped us cope with Covid-19?
PARLIAMENTARY BRIEF

AI in Public Health: How has AI helped us cope with Covid-19? is a Parliamentary Brief based upon the All-Party Parliamentary Group on Artificial Intelligence (APPG AI) Evidence Meeting held online on the 15th March 2021.

This APPG AI is co-Chaired by **Stephen Metcalfe MP** and **Lord Clement-Jones CBE**.

We would like to express our appreciation to the following people for their oral evidence:

- **Dr Indra Joshi,** Director of AI**, NHSX**
- **Dr Keith Grimes,** Director of Clinical Innovation**, Babylon Health**
- **Dr Eldad Kepten,** Head of Data Science, **Clalit Research Institute**
- **Dr Natalie Banner,** Understanding Patient Data (UPD) Lead**, Wellcome**
- **Raja Sharif,** Founder and CEO**, FarmaTrust**
- **Jonathan Levi,** CEO of **HACERA,** Founder of **MiPasa**

Big Innovation Centre is the appointed Secretariat for APPG AI

- CEO, **Professor Birgitte Andersen**
- Rapporteur: **Dr Désirée Remmert**

The video recording of the Evidence Meeting can be found on our websites.

www.biginnovationcentre.com | Email: info@biginnovationcentre.com | @BigInnovCentre
www.appg-ai.org | Email: appg@biginnovationcentre.com | @APPG_AI
© Big Innovation Centre 2020. All Rights Reserved

PARLIAMENTARY BRIEF

AI in Public Health:
How has AI helped us cope with
Covid-19?

All Party Parliamentary Group on
Artificial Intelligence

APPG AI Sponsors

The Group supporters – Blue Prism, British Standards Institution, Capita, CMS Cameron McKenna Nabarro Olswang, Creative England, Deloitte, Dufrain, Omni, Osborne Clarke, PwC, and Rialto – enable us to raise the ambition of what we can achieve.

Contents

1. Introduction

Artificial Intelligence (AI) technologies have played an important role in the research and development of diagnostic tools, pharmaceutics, and public health measures in the fight against Covid-19. In this meeting we discussed what we have learnt from the deployment of data-driven technologies in these different areas with a focus on safe and efficient data collection and data sharing. In particular, we addressed the question of how to enable collaboration across different organisations and the healthcare system to advance progress in medicine and public health as well as to facilitate the work of the NHS.

In this brief we will outline the main challenges that arose with the deployment of AI technologies in public health measures during the Covid-19 pandemic in the UK and abroad. Our suggestions for policymakers are based on experiences with AI technologies in clinical research, social care and healthcare, public opinion research as well as international experiences with AI technologies in public health presented by our speakers from the UK, the US, and Israel. The detailed evidence statements of our speakers can be found in the appendix of the parliamentary brief.

The APPG AI Evidence Meeting (15th March 2021) convened a group of experts in medicine, healthcare, and AI and data science from academia and businness.

- **Dr Indra Joshi,** Director of AI**, NHSX**

- **Dr Keith Grimes,** Director of Clinical Innovation**, Babylon Health**

- **Dr Eldad Kepten,** Head of Data Science**, Clalit Research Institute**

- **Dr Natalie Banner,** Understanding Patient Data (UPD) Lead**, Wellcome**

- **Raja Sharif,** Founder and CEO**, FarmaTrust**

- **Jonathan Levi,** CEO of HACERA**, Founder of MiPasa**

This meeting was chaired by co-Chair **Lord Clement-Jones CBE. C**o-Chair **Stephen Metcalfe MP** sent this apologies for this meeting.

Parliament has appointed Big Innovation Centre as the **Secretariat of the APPG AI**, led by **Professor Birgitte Andersen (CEO)**. The Project Manager and Rapporteur for the APPG AI is **Dr Désirée Remmert**.

1. Recommendations for policymakers

The expert speakers at the meeting agree that **data ownership, the reliability and transferability of data sets and AI systems, digital skills of healthcare workers, and public trust in the efficacy and safety of AI driven technologies** count to the major challenges of integrating AI technologies into the healthcare system in the UK.

To ensure the secure and effective deployment of data-driven technologies in the healthcare sector, our expert speakers suggest that government should **invest in data infrastructure** and establish **good governance and regulation** which **balances the need for regulatory rigour** with ensuring that **the UK healthcare sector can fully reap the benefits AI technologies have to offer.**

The speakers at the APPG AI evidence meeting "AI in Public Health: How has AI helped us cope with Covid-19?" highlighted the importance of collaboration and data sharing for the development of safe and reliable public health measures and AI-driven diagnostic tools in the fight against Covid-19.[1] They further agreed that there remain certain barriers that obstruct the deployment of data-driven systems in various areas of healthcare and research. Among these the most important aspects that require clarification and further reflection are questions around **data ownership, the reliability and transferability of data sets and AI systems, digital skills of healthcare staff, and confidence in the efficacy and safety of AI driven technologies among clinicians and the public.**

It has been emphasised at the meeting that integrating AI technologies into different realms of public health and healthcare would **supplement practitioners' expertise** with important insights. Coupled with **ongoing assessment of the clinical impact** of these AI systems, according to Dr Eldad Kepten, Head of Data Science at Clalit Research Institute, data-driven technologies can offer an important contribution to the healthcare system and medical research.

---

[1] With this, our speakers corroborate the Centre for Data Ethics and Innovation (CDEI)'s finding of the high relevance of data sharing to combat the pandemic which is outlined in the Covid-19 Repository & Public Attitudes report: "[…] aside from driving forward advances in vaccine research, artificial intelligence did not play the outsized role many thought it would in relief efforts. Instead, it has been conventional data analysis, underpinned by new data sharing agreements, that appear to have made the biggest difference to the work of health services and public authorities." CDEI 2020: *Covid-19 Repository & Public Attitudes: 2020 in Review.* London, UK: Centre for Data Ethics and Innovation.

Privacy and ownership of health data

Dr Natalie Banner, Understanding Patient Data (UPD) Lead at Wellcome, points out that the digitisation of NHS services has occurred at an unforeseen pace which brought with it **new questions around patient privacy and data ownership** that have to be clarified before further steps towards digitisation and the deployment of data-intensive technology in healthcare are made. It is thus essential to strike a careful balance between individual patient privacy and the benefit for the public. Whereas patient safety must not be compromised, Dr Indra Joshi, Director of AI at NHSX affirms, **the need for regulatory rigour should be balanced with ensuring that the UK healthcare sector can fully reap the benefits which AI technologies provide in this area**. For this reason, she continues, "the NHS AI Lab is leading a programme with the health and care regulators to define an appropriate regulatory framework for AI technologies, which streamlines requirements whilst also guaranteeing the safety of patients."

Banner agrees that **access to high quality, well formatted data is essential to useful and equitable AI systems**. Accordingly, the **government should invest in data infrastructure and establish good governance and regulation**. "Otherwise", Banner warns, "there's a risk that AI tools will exacerbate health inequalities, cause harm or discriminate against some communities."

Maintaing public trust in the deployment of AI in healthcare

Maintaining public trust in the deployment of AI in healthcare will be one of the decisive factors to the successful deployment of these technologies in the future.

Joshi reports that a longitudinal study conducted by the Centre for Data Ethics and Innovation in 2020, found that the **majority of respondents were confident that "digital technology had the potential to be used in response to the outbreak**, with majorities in favour of specific use cases such as using wearable tech to support social distancing." Keith Grimes, Director of Clinical Innovation at Babylon Health, made similar observations during the course of the pandemic: "Owing to the pressing nature of the pandemic, and the absolute need to manage our patients remotely, much of the resistance to the adoption of digital health technologies - such as video consultations, symptom checkers, and remote monitoring - fell away."

However, Banner warns that **the perceived lack of data transparency by the UK government during the pandemic has eroded public confidence and undermined the response to the virus.** Banner asserts that for people to be accepting of the use of AI technologies in public health, "**trust in governance is more important than their views on how effective the technology is**. Similarly, research conducted by the Ada Lovelace Institute found that for people to have confidence in the contact tracing app, independent oversight & clarity on the boundaries for the use of data was crucial." Banner summarises that **new technologies will only be embraced if there existed clear and consistent rules, public participation, and enforcement of the rules.**

Another challenge that arose and the design and deployment of AI systems used in the fight of the Covid-19 pandemic centred on the **alignment and integration of disparate data sources.** Joshi recounts that early in the the pandemic, NHSX collaborated with NHS England's Data Analytics Team to organise a new Data Store that brought together "multiple disparate data sources from across the UK on the number of COVID-19 cases, tests, ICU bed availability and other operational metrics. Using this single source of data meant decision makers in the NHS and Government were supplied with accurate real time information."

In this context, Kepten illustrates a successful case of bringing together disparate data sources for an accurate warning system in Israel: "Throughout the pandemic, it was of central importance to identify which of our 4.5 million members is in the highest risk for severe COVID-19 if infected. At the beginning, when only a limited number of cases was locally available, we quickly adjusted a previous model for severe flu for COVID-19 predictions. With more cases accumulating, we transitioned to refined machine learning models that incorporate detailed EMR data and give personal and explainable risks to all members. In all stages, the personal risk scores and

explanations where passed to GPs who actively contacted nearly 100% of their high risk patients and explained their situation and recommended actions. "

Projects like these require a **sensible review of the data sources and types of data used** which can be challenging when high amounts of data have to be processed in a short time and in cases where essential information on how the original data has been collected and assessed is not available. For this reason, standards should be designed that govern **transparency about how data was collected and the methodologies used to process and analyse it to guarantee more reliable modelling and forecasting results in the health sector**.

Expert collaboration and data sharing

An overarching theme of the meeting was the **importance of collaboration with partners in the public and private sectors**. Sharing data and expertise proved to be crucial to the deployment of reliable and efficient public health measure based on data-driven systems. Joshi explains: "Through these deployments of AI, we have learned a great deal about the need to foster close collaboration between partners within the system and industry. This **collaboration between organisations that typically operate separately or in direct competition allowed the health and care sectors to move at the pace required to meet the demands of the pandemic**." Similarly, Grimes highlights the importance for developing a **strategy for the publishing of public health data in machine readable format** on which the public and private sector as well as academia can draw in the research and design of data-driven technologies. He relates that "as the pandemic took hold we [at Babylon Health] had to rapidly adapt our symptom checker tool to help people identify whether their symptoms were COVID related or not. To do this, we utilised high quality clinical evidence and guidelines from trusted sources, which were being updated in a timely fashion. While the UK does lead in this area, **the rapidly changing guidance made it challenging for us to respond as quickly as we may have wished**. […] We believe that **open standards for data sharing and secure interoperability** are fundamental to coordinated efforts to deal with the ongoing pandemic, and then to address the backlog of care that has built up. "

Skills training for healthcare workers

However, to guarantee that the UK healthcare sector can reap all benefits which AI-driven software in public health have to offer, it must be ensured that **public trust in these technologies is maintained**. Therefore, it is imperative that **healthcare staff and clinicians understand the benefits of AI technologies and are able to**

convey this confidence to their patients. Kepten explains that healthcare staff and clinicians must receive **succinct and clear information from AI systems** which they can then communicate in an understandable way to their patients. Kepten explains: "Machine-learning predictions are only important if they effectively reach the end user. For us [at Clalit] – it is in many cases the GP who is in direct contact with the patients. However, the doctor at the end does not have time nor patience to understand the musings of AI experts. Thus, **results need to be communicated seamlessly in a concise and actionable manner** ("these are your patients at risk...", "expect 50 severe cases next week"). Such communication systems cannot be put in place hastily, and must reach all relevant practitioners. In addition, a clinician that is a domain expert should work with the ML specialist to translate predictive results into clinician-relevant language. " Accordingly, Grimes argues, it is important to **enlarge the pool of AI talent among clinicians with targeted skills training**. Whereas there have been made important steps towards this goal following the Topol Review of the NHS in 2019 as well as the publication of digital competency frameworks for healthcare professionals, Grimes suggests extending this approach "to **updating medical school curricula, in order to prepare the doctors and clinical workforce of the future to become architects of these changes**."

2. Evidence statements

Dr Indra Joshi, Director of AI, NHSX

What have we learnt from the deployment of various types of AI technologies over the course of the pandemic?

AI technologies were and continue to be pivotal to the NHS' response to the COVID-19 pandemic. Early on in the pandemic, NHSX, working with NHS England's Data Analytics Team, organised a new Data Store to bring together multiple disparate data sources from across the UK on the number of COVID-19 cases, tests, ICU bed availability and other operational metrics. Using this single source of data meant decision makers in the NHS and Government were supplied with accurate real time information. This data is also the basis for the public COVID-19 dashboards shared on gov.uk. In addition to this real time information, powerful AI forecasting tools were used to model the projected spread of the disease, including localised spikes in cases weeks before they materialised. These projections are used at a national and local level to support the allocation of resources to ensure the NHS can cope with increases in demand.

To support the effort to combat the pandemic, NHSX co-established the National COVID-19 Chest Imaging Database (NCCID), which brings together medical images of patients suffering with COVID-19. The NCCID is designed to enable the

development of software that helps doctors and researchers to understand the progression of the disease, assess severity of COVID-19 in patients, and identify factors that may complicate recovery. Researchers at Addenbrooke's Hospital, Cambridge, are developing an algorithm to inform more accurate diagnosis of patients presenting to hospital with COVID-19 symptoms.

Outside of NHSX, there were many other uses of AI including supporting drug discovery and vaccine development.

Through these deployments of AI, we have learned a great deal about the need to foster close collaboration between partners within the system and industry. This collaboration between organisations that typically operate separately or in direct competition allowed the health and care sectors to move at the pace required to meet the demands of the pandemic. NHSX has capitalised on this momentum with the continued operation of the Data Store and NCCID which will be invaluable in uncovering new opportunities to improve the efficiency of the health and care system and learning more about COVID-19 to support better patient outcomes respectively. Protection of patient information is a central component of this work.

The deployment of AI during the pandemic also reinforced the need to ensure the safety and efficacy of AI technologies before they are used in real world settings. Under no circumstance do we want to compromise on patient safety, but we need to balance the need for regulatory rigour with ensuring the benefits of AI technologies can be felt on the frontline. The NHS AI Lab is leading a programme with the health and care regulators to define an appropriate regulatory framework for AI technologies, which streamlines requirements whilst also guaranteeing the safety of patients.

How could data sharing and processing be improved to facilitate research in medicine and public health?

Following the positive impact data sharing had on the health and care sectors response to the pandemic (under the Health Service Control of Patient Information Regulations 2002), the government launched a review in February 2021, to focus on the more efficient and safe use of health data for research and analysis for the benefit of patients and the healthcare sector.

The review will complement the forthcoming Data Strategy for Health and Social Care that will set the direction for the use of data in a post-pandemic healthcare system.

The Secretary of State for Health and Social Care has asked Dr Ben Goldacre, Director of the Data Lab, Nuffield Department of Primary Care Health Sciences,

University of Oxford, to undertake a rapid review of data usage and sharing across the health system (report pending April 2021).

In addition to this review and to ensure the capitalisation on the opportunity to make the most of the large quantities of data generated by the NHS, the NHSX Centre for Improving Data Collaboration is exploring how the value of health data can be realised. This includes improving guidance and support for health and care providers when entering agreements with industry.

AI in the NHS – How to achieve better collaboration for intelligent problem solving?

There are a number of barriers to accelerating the use of AI-driven technologies in health and social care settings such as: clarifying ownership within the health and care system; ensuring there is reliable data sets and effective digital systems; ensuring staff understand the benefits of AI technologies, and clinicians and the public have confidence in the use of AI technologies.

The NHS AI Lab is working across the health and social care system to bring together system partners, local organisations and technology developers. By convening these groups to collaborate, the NHS AI Lab is spurring and enabling safe adoption of AI technologies into the health and care system at scale and pace.

For example, one of the Lab's programmes, the £140 million Artificial Intelligence (AI) in Health and Care Award will benefit patients by combining the power of artificial intelligence with the expertise of the NHS providers and clinicians to improve health and care outcomes. The Award will increase the impact of AI systems to help solve both clinical and operational challenges across the NHS by accelerating the testing and evaluation of technologies within NHS and social care settings.

Has citizens' attitude towards data sharing changed during the pandemic?

There is significant government activity and investment to advance the life sciences sector by improving access to health and care data and encouraging data-driven innovation. It is essential that these activities are aligned to patients' attitudes towards sharing data.

The Government's Centre for Data Ethics and Innovation longitudinal study of over 12,000 people between June and December 2020, testing public attitudes in the light of the pandemic found that 72% felt digital technology had the potential to be used in response to the outbreak, with majorities in favour of specific use cases such as using wearable tech to support social distancing.

Additionally, The National Data Guardian, the body responsible for representing patient views, is currently undertaking public dialogue to understand public attitudes towards the sharing of health and social care data for data-driven research and innovation in England. Full findings from the public dialogue will be reported by the National Data Guardian in 2021.

Dr Keith Grimes, Director of Clinical Innovation, Babylon Health

Over the course of the past 16 months, healthcare providers and medical device manufacturers have learned a great deal about the need to innovate at pace, whilst maintaining the same commitment to safety and quality. Owing to the pressing nature of the pandemic, and the absolute need to manage our patients remotely, much of the resistance to the adoption of digital health technologies - such as video consultations, symptom checkers, and remote monitoring - fell away. As a result, the global healthcare community was able to safely support and manage patients in their own home, identify those at highest risk and help them shield, and use high quality large datasets to undertake critical research on therapeutic interventions at pace.

OPEN DATA OPEN STANDARDS & INTEROPERABILITY

At Babylon, we have been pioneers of a 'digital first' approach to delivering accessible, affordable, and high quality healthcare in the UK and around the world. As the pandemic took hold we had to rapidly adapt our symptom checker tool to help people identify whether their symptoms were COVID related or not. To do this, we utilised high quality clinical evidence and guidelines from trusted sources, which were being updated in a timely fashion. While the UK does lead in this area, the rapidly changing guidance made it challenging for us to respond as quickly as we may have wished. We would recommend that an expanded strategy and plan be developed for publishing public health data using open standards, in a machine readable format. We would also request that healthcare providers be able to register for notifications of these changes whether they are from the public, private, or academic sector.

We believe that open standards for data sharing and secure interoperability are fundamental to coordinated efforts to deal with the ongoing pandemic, and then to address the backlog of care that has built up.

FRAMEWORKS AND GUIDANCE

We are pleased to see the continued global leadership in the healthcare AI space shown by NHS X and NICE, with the publication of the Digital Technology Assessment Criteria (DTAC) and Evidence standards framework for digital health technologies. For those in the industry, clarity and consistency are key to research, development and delivery. We also welcome the continued engagement with our sector, and look forward to ongoing collaboration with regulators and the NHS

TRAINING

I lead the Digital Health guild at Babylon, a team of product & technology proficient clinical subject matter experts that have a critical role in developing, building, and testing our digital health technologies. We use these to help patients self-care, support clinicians in caring for their patients, and allowing the providers to intelligently and effectively manage their population, where appropriate. These clinicians have required specific additional training to allow them to do this work, as the pool of talent remains small. The Topol Review of the NHS, published in 2019, made specific recommendations on preparing our current and future workforce for a digital future. Progress has been made on the implementation of those recommendations, such as the Topol fellowship at Health Education England, and the publication of digital competency frameworks for healthcare professionals, but we would recommend that this approach also be extended to updating medical school curricula, in order to prepare the doctors and clinical workforce of the future to become architects of these changes.

Dr Eldad Kepten, Head of Data Science, Clalit Research Institute

In the previous year, Israel has been no exception in its having to deal with the dangers and effects of the Sars-Cov-2 pandemic. From the early stages of protecting at-risk populations, through the continuous attempts to mitigate spread and prevent hospital overload, and up to the mass vaccination campaign. Although each country has its unique setting, opportunities and challenges, we believe that lessons can be learned from Israel's experience, and in the current context – the use of AI in tackling some of the challenges listed.

Israel has a universal healthcare system composed of four competing payer-provider public organizations. These organizations (HMOs) are responsible for the long-term wellbeing of their members, who can switch freely between providers. Clalit is the largest of the HMOs, caring for more than half of the Israeli population, from all socio-demographic segments of society with some over-representation of the old, the minorities, and the underprivileged. In the early 2000's, Clalit digitized its health records, including all community, hospital and specialist services. This has created a rich database, covering tens of millions of patient years, with thousands of features, fully recorded usually from birth to old age. The Clalit Research Institute (CRI) is responsible for turning these vast amounts of data into actionable medical and operational knowledge.

CRI includes professionals from diverse backgrounds, which work together to solve high gain challenges. Medical experts bring domain knowledge, epidemiologists devise unique approaches to detangle population level confounding, and data scientists and statisticians extract numerical insights based on these inputs. Our

products are clustered into several domains – basic medical research (predictive and causal), personally tailored medical recommendations for each member, and population level analytics.

During the pandemic, CRI produced several noteworthy models and solutions in these domains. For example:

- Warning system for members at risk – throughout the pandemic, it was of central importance to identify which of our 4.5 million members is in the highest risk for severe COVID-19 if infected. At the beginning, when only a limited number of cases was locally available, we quickly adjusted a previous model for severe flu for COVID-19 predictions. With more cases accumulating, we transitioned to refined machine learning models that incorporate detailed EMR data and give personal and explainable risks to all members. In all stages, the personal risk scores and explanations where passed to GPs who actively contacted nearly 100% of their high risk patients and explained their situation and recommended actions.
- Prediction of hospital loads – predicting temporal trends is a difficult task in an ever-changing pandemic. However, we utilized our personalized data to predict for each infected individual when they will be hospitalized and at what probable severity. Thus, we could give an estimation for each hospital as to the loads they should expect in the near future based on current infections, and prepare accordingly.
- Vaccine efficacy studies – In parallel to the vaccination campaign it was of important to measure the real world efficacy of the vaccine. Such studies demand the live connection of an epidemiological design, analysis code that is subject to changes, and a complex database. Eventually, this retrospective study was published in the medical literature, and further raised the confidence in the vaccine. While this is not the typical AI imagined by the public, this type of study relies deeply on coupling data with coding and mathematical capabilities.

So what have we learned from projects such as these, and what can we recommend?

First, we saw that models are as good as the data that feeds them. By connecting the whole health cycle of each Clalit member to a central database, we can create highly complex models that look into thousands of potential features. However, such multi-year history data can only be acquired over time, as some outcomes are best predicted by data from distant past records. Since the turnover of members in HMO's is very low, it was beneficial to plan ahead and collect data long before its application in ML. Today we are reaping the fruits of this continuous effort.

Second, machine-learning predictions are only important if they effectively reach the end user. For us – it is in many cases the GP who is in direct contact with the patients.

However, the doctor at the end does not have time nor patience to understand the musings of AI experts. Thus, results need to be communicated seamlessly in a concise and actionable manner ("these are your patients at risk...", "expect 50 severe cases next week"). Such communication systems cannot be put in place hastily, and must reach all relevant practitioners. In addition, a clinician that is a domain expert should work with the ML specialist to translate predictive results into clinician-relevant language.

Finally, medical experts are much less likely to believe the recommendations of a model without establishing a track record of excellence through medical publications and past clinical trials. During the past years, we have worked jointly with practitioners from different specialties on research, development, and clinical intervention projects. It may prove useful to foster a longstanding collaborative relationship between the AI researchers and the medical end users.

We believe that that we are only at the beginning of the AI revolution in healthcare. Increasing amounts of data, coupled with ever improving algorithms, will enable precise personalized predictions in more and more medical domains. There is no desire to replace practitioners, but rather supplement their expertise with real-time insights, delivered just in time, in clear actionable language in accordance with their preference and coupled with ongoing assessment to the clinical impact of these interventions.

Dr Natalie Banner, Understanding Patient Data (UPD) Lead, Wellcome

During Covid-19, the way data is collected and used has rapidly changed. Digitisation of NHS services has occurred at a pace not previously believed possible. And the rules around access to health data have been relaxed by the Control of Patient Information notices, allowing more data to be used for vital research like the RECOVERY trial or the work of OpenSafely.

However, lots of these changes happened with little oversight or scrutiny. Speed has come at the cost of transparency and good governance. For example, it took the threat of legal action for Government to publish the datastore contracts and admit a Data Protection Impact Assesment wasn't completed for Test & Trace. A recent report published by the Public Administration and Constitutional Affairs Committee argues that Government's lack of data transparency during the pandemic has damaged public confidence and undermined the response to Covid-19.

These issues must be addressed before the UK starts to routinely develop and deploy AI in healthcare. Access to good quality, complete, well formatted data is fundamental to AI, and critical to ensuring AI technologies are useful, useable and equitable. That means investing in data infrastructure and establishing good governance and regulation now. Otherwise, there's a risk that AI tools will exacerbate health inequalities, cause harm or discriminate against some communities.

Why now is the time to establish strong data governance

Data is more visible and tangible than ever before

- Covid-19 has made the importance of the effective and responsible use of

data clearer than ever.

- Statistics on infection rates, hospital capacity and mortality are discussed daily - all of which come from data collected by the health and social care system.

- According to National Data Guardian (NDG) polling, 56% of people have learned more about how health and care data is used to monitor public health and for research during the pandemic.

- But this new profile also brings more scrutiny. Understanding Patient Data's latest research found that media reporting on health data has increased during the pandemic, with stories about risks or issues getting the most attention.

- The public will only embrace new technologies if the rules & governance are right

- Recent research by the Centre for Data Ethics and Innovation (CDEI) found that for people to believe technology could help with the Covid-19 response, trust in governance was more important than their views on how effective the technology is. Similarly, public engagement run by the Ada Lovelace Institute found that for people to have confidence in the contact tracing app, independent oversight & clarity on the boundaries for the use of data was crucial.

- The CDEI research also found only 43% of people think the existing rules and regulations are enough to ensure that technology is used responsibility. There is still work to do to setup the right governance mechanisms and explain them to the public.

- **Good governance is crucial to the equitable use of new AI technologies**

- Covid-19 has demonstrated the power of data, and more data is available than ever. At the same time, AI technologies are improving and beginning to show potential in the health sector (although yet to really deliver on the promise).

- It's well known that data and AI have the potential to be discriminatory, and cause harm to people. In health, these risks could have a high impact both on individuals and groups, potentially in ways that are not immediately obvious.

- AI technologies are not unique in posing these risks – many algorithms and decision-making processes can embed bias or discrimination, but these may be more hidden in black-box AI systems.

- It's important to acknowledge such risks and specify steps to mitigate them. The public cares about health inequalities and believes the benefits from using data should help improve health for everyone. Research by Understanding Patient Data found that 81% of people believe the benefits from health data partnerships should be distributed across the NHS.

- **What does good governance mean?**

- Clear and consistent rules: this makes it easier for people to understand what's happening to data, and for industry or academic partners to play by

those rules. There's currently a lot of inconsistency in the system - a patchwork of rules and processes does not inspire public confidence.

- Public participation: people should be involved in setting the rules for data use. It's not just the right thing to do, people expect it: Understanding Patient Data's research found that 74% of people believe the public should be involved in decisions about how NHS data is used.

- Enforcement of the rules: AI ethics principles or good-practice guidance are not enough on their own. As well as defining and incentivising best practice, Government also needs to establish, explain and enforce sanctions for bad practice or data misuse. This requires legislation with clear red lines to mitigate the risk of harm. The upcoming National Data Strategy for Health and Care, along with the 'Integration and innovation' whitepaper, offer a good opportunity to establish this.

Risks and opportunities

This is a unique moment. The UK is emerging from a pandemic which has made the importance of effective and responsible use of data clearer than ever. The policy landscape is shifting, with the National Data Strategy, the National Data Strategy for Health and Social Care, and the Goldacre Review all in development, alongside the AI Council's Roadmap.

The potential is exciting, but like with any innovation, this technology needs to be carefully stewarded to ensure it develops in safe and effective ways. We have seen from the enormous success of the Covid-19 vaccination programme just how valuable a robust, clear regulatory process is for ensuring new, innovative vaccines are safe and to inspire public confidence in them. We should take a similar approach to AI technologies, especially when they're deployed in a high-stakes sector like health. There is a risk that data and AI technologies are discriminatory and cause great harm. But there's also a clear opportunity, for the UK to become world-leading in trustworthy data governance, making more data available for the development of new technologies and lifesaving research.

Note: If you have questions about this response, please get in touch at hello@understandingpatientdata.org.uk.

Raja Sharif, Founder and CEO, FarmaTrust

Raja Sharif
Founder & CEO, Farma Trust

FarmaTrust has been working on blockchain track and trace technologies since 2016 although we formally incorporated in 2017. Understanding HOW the data we collect is central to our strategy.

We specialize in the pharmaceutical and healthcare sector. So we are familiar with and have solved issues around sovereignty of data, keeping data "on-prem" for NHS hospitals, GDPR and other sensitive data issues.

Our solutions track assets from the point of production to the point of consumption with full chain of custody, identity and condition data.

We have TWO divisions tracking:

i.   pharmaceutical products eg medicines or medical devices, vaccines, test kits, PPE
ii. tissue samples, organs and treatments etc stem cell or cell and gene therapies – for conditions such as blood cancers and spinal muscular atrophy

We have expanded our service offerings, so we also have the additional capabilities which eases healthcare administration and improve the patient experience:

— Health passes/Vaccination Apps to whom medicines, vaccines and test kits we track are administered
— Scheduling Solutions for cell and gene therapies (and cancer treatments) – funded by Innovate UK provided funding to get around COVID challenges

such as removing the "waiting room", "tagging vulnerable or infectious patients" and patient tracking to ensure attendance at pre-arranged appointments

— Data Analytics including predictive analytics to prevent waste and shortages.

Some of the benefits that these data driven technologies combined with Artificial Intelligence play in public health measures include:

— eliminating fake medicines, vaccines, COVID tests
— reduce administration burden through automation of purchasing processes, payments, and stock allocation
— elimination of fraud in the supply chain
— prevent waste through over-ordering
— prevent stock-outs or shortages
— automate regulatory compliance

One of the key advantages of systems like ours is:

i. Data can be easily anonymised and blockchain is built on cryptography
ii. All stakeholders can have access to the same data in real time – allowing true data driven decisions or immediate recommendations from AI engines
iii. Our systems are specifically created to technology neutral and blockchain platform neutral – so interoperability is assured and freedom to switch is easily done.

So we believe a closer look at:

— technological infrastructure is required in the health sector
— regulations need to be clearer and consolidated
— data sharing education needs to be better (eg a packet of aspirins is NOT protected under data protection)

Citizens' attitude to data sharing is as follows:

i. patients generally do not mind their data being shared if it helps them or society generally, but they do not want to give it to commercial organisations that may make profits without some kind of return.
ii. The public still want the ability to control access to their identifiable data or "personal data"

We are initiating a project with a professor at Oxford University – its aim being that if a subject withdraws their consent to use their clinical data, then how do we withdraw that data from not just the companies that got that data directly, but second/third layer data use by companies that got the data indirectly.

Our Vaccination and COVID testing App has a blockchain base and therefore the User has full control over who gets access to their data and for how long. This can be used to open up the economy and travel. More interesting the data can be anonymised so that AI engines can be used for predictive analytics.

What we see is a paper Vaccination card but what we should be using is a Vaccination App – its available for use NOW with full data controls for the user, GDPR compliance, and can even operate access controls like office entrance turnstiles.

We believe that healthcare really needs to move to updating its IT infrastructure, and focusing on new technologies, such as blockchain, IoT and AI.

Jonathan Levi, CEO of HACERA, Founder of MiPasa

At MiPasa, we work with multiple parties that connect to a network where data is shared. Before that, we worked on clouds and infrastructure products where we needed to share information, however not with everybody. In our system there is always an auditor.

Everything has a separation between the private sharing of data by individuals or companies and groups versus the public version. We took the same technology that people use for supply chain and go to the entire procurement process and we do the same thing with data. MiPasa is a browser-based cloud native platform that allowed us to integrate data from Johns Hopkins University, from Oxford University, from the New York Times, from every country that initially shared public data sets. What we wanted to do is to bring everybody to the table. I noticed early on by mentoring a few Hackathons that the main problems arose around responsible AI.

We suggested building what we called reproducible AI. We built a system that can work with multiple languages so we can get data from Estonia, France, and the UK to name but a few. Here we looked at different metrics. MiPasa allows people to see a collection of publications on the topic of their specific interest. One can either look at the open data, we also allow people to publish their research and their code. If anybody is worried about how a certain graph was produced, they can build a supply chain with all different versions of a model. One can hence track the data set which was used to produce a certain model.

We support machine learning, AI, and statistical analysis. Our main objective is to version all the copy data as it arrives, to version the code and the models on which each respective insight is based, and thus to create a reproducible end-to-end

result. People make mistakes, this can happen easily when working with large data sets under a lot of pressure. Often, care providers need to work with many partners, some of them can be infrastructure partners, but others are caregivers with whom they must share information.

Our system allows you to define what you are sharing and under what condition, to work with unstructured data, and to code against this data was very minimal effort. The idea is that everything is cloud-based. You can work on premises, so basically people should not worry about exporting data if they do not need to. There is full visibility as to who can access, who can read, who an update, and who can model against the data. Actually, one can check later on whether people have been using the data that we have published.

Many publications are not published in a machine-readable format or with logging. This means, it is very difficult to ascertain whether this data is useful. We, by contrast, aggregate data, create a time capsule of global events, show a live map of all changes, and if there is any mistake, we can report it. The basic concern is around people who are afraid of publishing data, - it is natural in times of such a global crisis that not all systems are up to date. Sometimes the creators of flawed modelling receive substantial criticism. We wanted to allow them to publish their models and to get feedback about them. This started a conversation around modelling and data – this process produced results that were easily comparable. Our goal was to create visibility, transparency, and a "repeatability" across the entire chain of data processing. We eventually created a fully functional software service platform based on the need that emerged in the Covid-19 crisis.

We should be more worried about what we share with partners – it is a big organisation – for how long, and under which conditions. It is true that it is difficult to make a request every time under crisis, but we have a set of policies that can enforce the whole data economy - from the collection of data to tracking the publications based on the data. It is a full end-to-end solution that guarantees that we meet the strictest rules in terms of data privacy, compliance, sharing, and exporting.

Contact

APPG AI Secretariat

Big Innovation Centre
14-16 Dowgate Hill
London EC4R 2SU
United Kingdom

info@biginnovationcentre.com
www.biginnovationcentre.com

appg@biginnovationcentre.com
www.appg-ai.org