Key messages

UK-Japan Symposium on Data-Driven Health:
Data strategies to predict risk, prevent and manage disease in individuals and populations.

Wednesday 26 February 2020

国立研究開発法人 日本医療研究開発機構
Japan Agency for Medical Research and Development

AMED

The **Academy** *of* **Medical Sciences**

JSPS

part of the

Japan-UK
Season of Culture
2019-20

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 November 2020.

COVID-19 notice
This symposium was held during the early stages of the subsequently global outbreak of SARS-CoV-2, the virus causing the disease COVID-19. This event was fully compliant with UK government policy, public health and travel advice at that time. Facilities to support general respiratory infection control were provided, and contract tracing information was collected from attendees. No attendees reported developing symptoms of, or being diagnosed with coronavirus to the organisers after the symposium.

UK-Japan Symposium on Data-Driven Health:

Data strategies to predict risk, prevent and manage disease in individuals and populations.

26 February 2020

Contents

Welcome from the co-Chairs

We are currently in an era of opportunity. A huge amount of investment is stimulating development in data science, promising new means to optimise health and quality of life at the person and population level.

Empowered with this opportunity, and the surge in globalisation that new technologies have fuelled, researchers, clinicians and policymakers must consider: how can we improve health not just in one country, but globally?
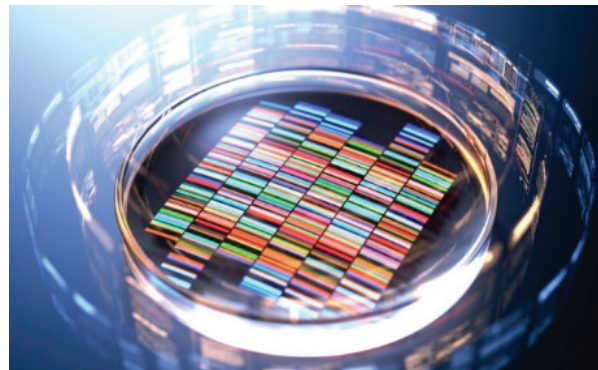
Much important groundwork has been done to explore the principles that must underpin health data research to ensure it is ethical, transparent and engages appropriately with the people and patients it affects. Amongst such activities is the public dialogue programme, policy workshop and report from the UK Academy of Medical Sciences: *'Our data-driven future in healthcare People and partnerships at the heart of health related technologies'*.[1]

Keeping these person-centred principles at the heart of our scientific approach, we are now in a unique position to rapidly expand the global offering of digital health solutions. Maximizing benefit from this exciting process will require pragmatism and close communication across the international research community to avoid duplication, and instead look for interoperable and complementary technologies that could lend to systems that can be deployed quickly, apply widely, learn, and can be refined over time – while, crucially, maintaining trust in research.

We are delighted to have co-Chaired this event on '*Data strategies to predict risk, prevent and manage disease in individuals and populations*' to address these questions, organised by the Academy of Medical Sciences, Japan Society for the Promotion of Science (JSPS) and Japan Agency for Medical Research and Development (AMED).

Through the day's presentations and discussions, speakers and participants explored health data optimisation and resources including biobanks; health data for public health and delivering care through digital platforms; and health data to support clinicians through artificial intelligence (AI) and machine learning. Moreover, the symposium was an opportunity to establish new research connections between the UK and Japan, starting conversations that could support enduring partnerships.

[1] https://acmedsci.ac.uk/policy/policy-projects/use-of-patient-data-in-healthcare-and-research

Continued progress demands an ecosystem approach to thinking about data issues, working across disciplines, industries and countries. The emerging outbreak of COVID-19 is one example of an urgent issue where we need to act collaboratively and collectively in the health data medical science innovation space.

This symposium sought common ground for collaboration between the UK and Japan, to facilitate integrated international research inclusive of different populations, experts, policy systems and the public. We are delighted participants agreed that through the sharing of experiences in the development and application of data-driven health expertise and technologies, both countries can advance medical research and the level of healthcare provided to patients and public. We hope that this event laid firm foundations for successful partnerships in the future.

Professor Yuko Harayama
Professor Emeritus
Tohoku University

Professor Jill Pell FMedSci
Professor of Public Health
University of Glasgow

Event summary

Over three sessions, this event considered data strategies to predict risk, prevent and manage disease in individuals and populations, and examine the assistive technologies that may arise from these.

The symposium brought together representatives from academia, industry and the healthcare sector in the UK and Japan to explore this challenge in the context of:

- The health data landscape and resources in the UK and Japan
- Health data for public health
- Health data for clinical decision-making

This document provides an overview of the key messages to emerge from the event.

The full agenda is provided in **Annex 1**.

Professor Dame Anne Johnson DBE FMedSci, Vice President (International) of the Academy of Medical Sciences, in plenary.

The organisers are delighted that this symposium formed part of the '**Japan-UK Season of Culture 2019-20**'. This initiative aims to showcase Japan's multi-faceted appeal, from its culture and art to innovation and cutting-edge technology in areas such as medicine, science and industry. The official logo for this project, featured across the materials for this event, features interlocking waves to symbolise exchange between Japan and the UK.

Mr Satoshi Katahira, Minister for Economics at the Japan Embassy at the event reception.

The symposium was followed by a networking reception during which the organisers were honoured to receive Mr Satoshi Katahira, Minister for Economics at the Japan Embassy. The Minister delivered a toast (kanpai), during which he shared his hopes for enhanced collaboration between the UK and Japan, reiterating a theme of opportunity that had been pervasive throughout the day's discussions.

The Academy of Medical Sciences, Japan Society for the Promotion of Science (JSPS) and Japan Agency for Medical Research and Development (AMED) are most grateful to the co-Chairs for their work towards the development of this symposium.

This document reflects the views expressed by participants at the meeting but does not necessarily represent the views of all participants, the co-Chairs, the Academy of Medical Sciences, JSPS or AMED.

Key messages in brief

Both the similarities and the differences between the UK and Japan are valuable bases for scientific investigation to reveal new data-driven health solutions.

Although counterpoised on opposite sides of the world and culturally unique, both island nations face similar public health challenges such as ageing populations and a high burden of non-communicable diseases (NCDs) like heart disease. Each country has rich potential for health data insights, as both provide citizens with universal health coverage and a similar, high standard of care; this is delivered through the National Health Service in the UK and via the National Health Insurance System in Japan.

Participants at the symposium agreed that no single country has all the solutions to make data-driven health a reality. We must take an 'ecosystem approach' and find complementarities with trusted and strategic partners in order that we may advance together. In this regard, much could be achieved through bolstering scientific collaboration between the UK and Japan.

During the event, some key themes emerged as guiding principles to support rapid advancement of data-driven health solutions and, in particular, international collaboration in this area. These themes were as follows, and are explored in more detail below:

- The digital age is reshaping health promotion and healthcare for individuals and populations.
- Robust data-driven health solutions rely on strong data infrastructures for research.
- Pragmatic innovation can increase the use of existing data and health tools to accelerate research progress.
- Trust and patient and public involvement are central to meaningful health data research.
- International collaboration is invaluable to improve global health, and can be enhanced through sharing and connection.

Speakers and co-Chairs at the symposium

Key messages explored

The digital age is reshaping health promotion and healthcare for individuals and populations

Keeping people healthier with health promotion and risk prediction tools

Digital health solutions are not only changing how we treat disease, but offer the opportunity to help keep individuals healthier. Whether designed for disease risk management or promotion of healthy behaviours, symposium participants agreed that to maximise the benefit of digital or mHealth approaches they must engage the wider public, not just patients.

There is an opportunity now that the overwhelming majority of people own a smartphone, and hundreds of thousands of health and fitness smartphone applications ('apps') are available and are downloaded about 5 million times globally every day. In the UK, there are just under 7 million active users of wearable devices, such as fitness tracking wristbands. Such tools generate a huge volume of health data about individual users which is currently inaccessible to traditional health systems. Participants felt these data and tools could be harnessed for public health benefit, but such approaches must be people-centred and maintain public trust.

Liz Ashall-Payne, CEO of ORCHA,[2] explained that the biggest blockers to increased use and integration of health apps are awareness, access, trust and governance issues. To tackle these, ORCHA have built a review and accreditation engine that scrutinises health apps across areas including data privacy, security, clinical assurance and user experience in collaboration with experts. ORCHA then distributes this review information via health app libraries in the UK and abroad, which help users to find and access trustworthy apps that are tailored to their individual needs. ORCHA have worked closely with organisations, including NICE, to produce their efficacy framework, and are working with NHS Digital on projects including the NHS App, signposting how integral health applications may become to how we access healthcare in the future.

Professor Aiden Doherty is studying whether the information generated by wearable devices can help to predict and prevent cardiovascular disease (CVD) through longitudinal cohort studies in individuals registered with UK Biobank. Wrist-worn activity trackers were given to over 100,000 individuals for 7 days, providing a snapshot of their activity status. These measurements can be linked with their later health outcomes. Higher self-reported activity levels were associated with a 17% reduced risk of CVD incidence, but self-reporting can be inaccurate. Device measurements showed higher activity levels were linked with a much more substantial (52%) reduction in CVD risk, supporting the importance of exercise as part of a healthy lifestyle.[3] In the future, this research could inform a model of human movement behaviour based on 90,000 UK Biobank participants, which could be analysed for associations with different health outcomes and genetic variants and influence much broader public health approaches.

---

[2] https://www.orcha.co.uk/
[3] Under review for publication. See presentation of Adrian Doherty for data presented.

Streamlining front-line healthcare with data-driven technology and precision medicine

Ultimately, better utilisation of health data could fundamentally change health systems and care delivery. ORCHA are working with healthcare professionals to educate and train them in the world of digital health and support tailored 'prescribing' of heath apps.

At a system level, Professor David Clifton presented how machine learning and AI is being integrated into hospitals to directly improve patient care. For example, machine learning can predict which emergency department attendees might be affected by an avoidable breach of safety standards, such as being inappropriately discharged, with 85% accuracy.[4] Looking forward, his team hopes to build software to predict which ward a patient might be admitted to from the emergency department, to allow pre-emptive allocation of beds.[5] In their HAVEN project, researchers are also using electronic hospital record (EHR) data to predict in-hospital readmission to ICU.[6] Hospital systems may also be programmed to automatically order certain tests and investigations to facilitate patient transfers or discharge, reducing demands on clinicians' time. Early models of such prediction systems are currently being trialled in hospitals in Oxford, UK, and promise to be a great support for effective and streamlined clinical decision-making.

Another area where data-driven solutions can lead to more refined and effective care pathways is in precision medicine. Cancer can increasingly be regarded as a group of rare diseases requiring targeted, individualised treatment. Professor Chikashi Ishioka shared an example of personalised medicine research executed at scale through the work of the Personalized Medicine Center (P-MEC) at Tohoku University Hospital.[7] Every cancer patient in Japan now has access to a gene panel test which is reimbursed by the Japanese public medical insurance system. P-MEC are performing clinical sequencing, whole genome sequencing and other OMICS analyses on this data to identify new gene variants and to advance research into personalised medicine approaches in cancer. P-MEC work closely and reciprocally with the new Advanced Research Center for Innovations in Next-Generation Medicine (INGEM),[8] which in turn focuses on areas including Genome Drug Discovery and Clinical Genome Diagnosis. This research will deliver valuable patient benefits and reduce costs of healthcare over the longer-term.

---

[4] See slide 11 of David Clifton's presentation.
[5] See slide 12 of David Clifton's presentation.
[6] https://www.ndcn.ox.ac.uk/research/critical-care-research-group-kadoorie-centre/research-studies/hospital-alerting-via-electronic-noticeboard-haven/haven
[7] http://www.p-mec.hosp.tohoku.ac.jp/
[8] https://www.ingem.oas.tohoku.ac.jp/english/

Applying AI tools appropriately, and how good data can help

Machine learning (ML), a form of AI technology, has the potential to be revolutionary in the right clinical and public health context. It is a fantastic tool for analysing complicated, multidimensional information that is beyond our capacity as humans to fully scrutinise. ML has been shown to predict individual risk of atrial fibrillation (a heart rhythm problem) based on electrocardiograms (ECGs) that appear normal to the human eye.[9]

It has also been used not only to improve the accuracy of preoperative prediction of malignancy in ovarian cancer from 86% using traditional statistical methods to 92%, but to identify new patterns of disease among ovarian cancer patients that have not been previously recognised by clinicians.[10] Only ML tools can perform analyses such as these, which promise powerful clinical applications.

But while ML is producing exciting science, it should not be regarded as a 'silver bullet' for improving health care. It must be correctly applied to the right medical need and the right target population, and relies heavily on the availability of good quality data. It is typically not as useful for heterogeneous, routinely collected clinical information including binary variables, which are core data underpinning healthcare systems. ML can also be easily confounded and so may be vulnerable to error or fraud. For example, a small amount of visual 'noise' introduced into a digital image at a level undetectable to the human eye can completely change an AI prediction of whether a skin lesion is benign or malignant.[11]

Professor Iwagami noted that ML it not always better than traditional methods, noting a systematic review found no performance benefit of ML over logistic regression for clinical prediction models to date.[12] However, Professor Eiryo Kawakami explored how population data-based solutions may help us close this gap. In a study on diabetes risk, machine learning found nine health states in people without diabetes ranging from healthy groups, to pathological states with several risk factors for diabetes. These states were derived from an analysis of 2.5 million medical examination records from 300,000 people over 5 years in Japan, using a landscape model based on 10 risk factors associated with diabetes onset. This kind of risk stratification based on high-quality, large-scale longitudinal data could change how we define and predict disease, and open up new, tailored disease prevention approaches over the next decade.

Moving forward, we need to maintain a clear vision of what AI and ML can offer in the immediate term, which may be concentrated in particular fields including imaging. Realising the full potential of these technologies to improve public health will require rich, diverse, linked and accessible longitudinal population health data. Building such data resources is therefore the first step to reap the full benefit of AI and ML, in addition to continuing to develop, share and standardise clinical phenotype data – which may fundamentally change how we view health and disease.

---

[9] See slide 17 of David Clifton's presentation.
[10] Kawakami E, *et al*. (2019). *Application of Artificial Intelligence for Preoperative Diagnostic and Prognostic Prediction in Epithelial Ovarian Cancer Based on Blood Biomarkers*. Clin Cancer Res. **25(10)**. 3006-3015.
[11] Finlayson S *et al.* (2019). *Adversarial attacks on medical machine learning.* Science **363**, 1287-1289
[12] Christodoulou E, *et al*. (2019). *A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models*. J Clin Epidemiol. **110**.12-22.

Robust data-driven health solutions rely on strong data infrastructures for research

Facilitating health data research relies upon accessible, flexible and secure health data systems. Biobanks are an excellent example of this kind of resource. In Japan, the Tohoku Medical Megabank Organization (ToMMo) was established in 2012 to support the whole genome analysis of the Japanese population.[13] Now, ToMMo offers its over 150,000 participants disease-susceptibility information, personalised prevention approaches and medical care based on their personal characteristics including genome information. ToMMo has generated valuable insights into genomic features specific to Japanese ethnic groups, and has recently disclosed 4.7K Japanese single nucleotide polymorphisms derived from 4,700 Japanese volunteers. ToMMo is now one of 12 biobanks in Japan which are networked for cross-searching of specimens and data.

In the UK, the vision in the UKRI industrial strategy to build a robust national health data infrastructure, led to the establishment of Health Data Research UK (HDRUK).[14] HDRUK is an independent organisation that aims to provide the tools, policies, technologies and processes to enable data access at scale so that it can be harnessed for health and care impact. Their 20-year vision is that medicine will change such that almost every interaction will be influenced by big data. To facilitate this, HDRUK hopes to achieve unification of diverse data at scale, ultimately allowing studies on up to 66 million people, including hospital admissions and biobanking data.

Pragmatic innovation can increase the use of existing data and health tools to accelerate research progress

Ensuring high-quality data: use it to improve it

High data quality is fundamental for robust research and the development of accurate, reliable new AI technologies. However, 'data quality' means different things to different groups. For example, biobanks use measures like International Organization for Standardization (ISO) standards to evaluate quality, but researchers' preferred indicator is often the published research arising from a data set. This problem around how data quality is perceived, communicated and therefore delegated is a key issue for the research community to address.

The related process of quality assurance through data curation and harmonisation is highly involved and resource-intensive. In addition to HDRUK's extensive work on data linkage, they have established 7 data 'hubs' in a £50 million programme to improve data at scale. The hubs are centres of excellence that make data available for trustworthy research, and have developed high-profile global partnerships with industry partners, including Microsoft and Roche, and charities. This example of cross-sector collaboration shows how by working together, multiple partners can achieve more than single institutions.

---

[13] https://www.megabank.tohoku.ac.jp/english/
[14] https://www.hdruk.ac.uk/

However, in some cases fully harmonising a data set before releasing it for use may mean that it is poorly discoverable for extended periods, and it may even become obsolete. For example, just 15% of tissue samples collected for research in the UK and their associated data are used. Arguably, this is the least-reported type of sample misuse and represents an enormous amount of wasted research potential. Non-use of health data also compounds quality concerns, because if it is unclear if, how or when data might be applied in research, the incentive to robustly steward its quality is low.

Participants suggested that the best way to drive high data quality is to encourage its active use in research, attracting a shared interest and responsibility around good data management. Moreover, there was considered to be a duty to use data that have been consented for use in research. This points to an important role for data access tools that incentivise increased use of existing health data resources and, both as a by-product and intentional process, help improve the data over time. An example of such a tool is ATLAS,[15] developed by the UKCRC Tissue Directory and Coordination Centre (TDCC).

With the aim of rapidly increasing the research use of tissue samples by making existing data more visible, ATLAS placed pragmatic innovation at the centre of its approach. The developers released data with a baseline level of harmonisation for immediate use, while also implementing a longer term harmonisation and data standards strategy. This was achieved by:
- Using existing harmonisation standards such as the Observational Medical Outcomes Partnership (OMOP) Common Data Model, rather than defining new ones;
- Looking for 'obvious' harmonisation opportunities such as word matches;
- Enabling users to filter database search results flexibly and perform 'live' harmonization;
- Tracking user searches to inform targeted follow-up harmonisation activities.

The launch of ATLAS: an advanced data search tool for researchers

Professor Phil Quinlan announced the exciting launch of the data discovery tool ATLAS.

ATLAS is a platform that provides the international research community with access to detailed, live information about the data associated with tissue samples held by some of the UK's largest bioresources, cohort studies and human sample resources. The project has been led by the University of Nottingham (lead of the UKCRC TDCC) and HDRUK, supported by Nic Timpson from the University of Bristol.

The platform can be accessed at: https://biobankinguk.org/atlas/

This 'minimal viable product' model, popularised by start-up business entrepreneurs, was also core to the development of HDRUK's Health Data Research Innovation Gateway. The Gateway provides a common entry point to discover and enquire about access to UK health datasets held by members of the UK Health Data Research Alliance.[16] Tools such as these adopt a philosophy that can guide and protect our efforts in the rapid-moving landscape of health data science: start now and get better, but don't wait to get started.

[15] https://biobankinguk.org/atlas/
[16] https://ukhealthdata.org/

Future-proofing data and the role for real-world evidence

This adaptive approach also places an emphasis on the agility of data resources. The pace of improvements in technology means that past innovations are not always efficiently used. If a new, preferred technology becomes available, for which existing data resources are insufficient or incompatible, data collection will often be started afresh. So begins another lengthy cycle of data recruitment, curation and research, racing against the clock to generate insights before the technology is superseded and becomes redundant.

Future-proofing the health data that we collect for research (and the systems we use to store, discover and access that data) would help ensure the benefit this data generates is continuous, cumulative and enduring. This means investing in interoperability, internationally recognised standards, and data systems that are open rather than a 'black box' - even when these are procured from third parties. Successful efforts at this scale require political support. An example is The Global Alliance for Genomics and Health (GA4GH), a policy-framing and technical standards-setting organization, seeking to enable responsible genomic data sharing within a human rights framework. Even more ambitious endeavours could be to embark on the same path for disease phenotype data, or harmonise the format of routinely collected hospital data extracted from electronic medical records across countries.

The consistent, speedy development of technology also raises questions about how we can best evidence the safety and efficacy of digital tools designed to support health and wellbeing. This is particularly challenging for concepts such as personalized AI, where a system might dynamically learn and change for each patient. However, it is also a current issue for technologies in clinical use today. For example, smartphone apps must be updated within 18 months or are considered redundant by ORCHA.

This is because if they do not update along with our phones, our data security could be compromised. Randomised controlled trials (RCTs) also take around 18 months – meaning by the time a product could be proven clinically effective by an RCT, it must have updated and changed.

Several popular apps have been through multiple slow and expensive trial cycles to evidence successive versions of their product. Blinded studies are also a challenge to achieve in the field of new technologies - for example, it is difficult to blind for use of a wearable device. Although randomised trials must remain the gold standard for evaluating efficacy, this has fed discussion about what the next generation of digital regulation might look like. Could we integrate the use of real-world evidence and patient reported outcome measures (PROMs) into regulation, to better support safe and progressive innovation?

ORCHA are taking steps in this direction by collecting PROMs on prescribed digital products in a crowd-sourcing initiative built like a clinical trial, to support more personalised prescribing. However, the viability of such approaches is likely to vary depending on the digital product, its target population and the research or regulatory questions at hand. For example, Japanese attendees reported that data from apps recording symptoms experienced in mild conditions such as dry eyes or hayfever were often unreliable: high incidences of missing or nonsense data were unsuitable for longitudinal research. Robust patient and public involvement (PPI) and user-developer collaboration may reveal new solutions to improve real-world evidence generation.

Trust and PPI are central to meaningful health data research

Discussions also highlighted the importance of ensuring public trust in data-driven technologies, for which PPI is key. The UK Academy has previously explored this area in '*Our data-driven future in healthcare People and partnerships at the heart of health related technologies*' report,[17] and a joint event with HDRUK and the Collaboration for the Advancement of Sustainable Medical Innovation (CASMI): '*Realising patient and NHS benefits from health and care data: from policy to practice*'.[18]

Just by living we are generating data that is acquired, integrated and often monetised. For example, there are around 7 million active users of wearable health devices in the UK,[19] which can generate up to around 180 million data points per person per week[20] – more information than it may ever be feasible to feed into health and care systems in real time. With the rise in popularity of health applications, there are also now around 5 million downloads of health and fitness apps every day.[21] These are not only used by the younger generation, although this cohort may be the best informed around data privacy issues – the biggest increase in users of smartphones is in those aged over 65.

Since it is clear that we are on the path to a digital future, it is key that user populations are brought into technology development and health research processes. An example of this kind of PPI is the RUDY JAPAN initiative.[22] Using an online portal to collect information on the symptoms and daily life of patients with intractable diseases and rare diseases, patients and researchers can advance medical research together. As part of this project, hereditary angioedema patients have worked with researchers to define what information the RUDY JAPAN portal will record about their condition, and developed an app to help them input and self-monitor this data.

This is a growing success story for PPI in Japan, which is at an earlier stage of integration into research culture than in the UK. Indeed, Japanese cultural norms around turn-taking and listening are very effective in supporting the sharing of ideas and opinions between differently positioned stakeholders.

Organisations such as Understanding Patient Data[23] (UPD) in the UK are now supporting broader conversations on the use of patient data, particularly data collected routinely within the health service that might be used for purposes beyond individual care without explicit consent. UPD works with patient groups, charities, NHS organisations and policymakers to bring transparency, accountability and public involvement to the way patient data is used in, for example, research and planning. In Japan, the Evidence Generating Commons Project[24] is similarly seeking to create a shared space for policymakers, patients and researchers to work together to progress medical research. This project extends past data issues, to question whether patients are getting an adequate voice in what research is funded, and how PPI could be embedded in priority setting.

---

[17] https://acmedsci.ac.uk/policy/policy-projects/use-of-patient-data-in-healthcare-and-research
[18] https://acmedsci.ac.uk/more/events/live-stream-event-realising-patient-and-nhs-benefits-from-health-and-care-data-from-policy-to-practice-
[19] See slide 2 of Aidan Doherty's presentation
[20] Doherty A *et al*. (2017). *Large Scale Population Assessment of Physical Activity Using Wrist Worn Accelerometers: The UK Biobank Study.* PLoS One **12(2)**, 738-744.
[21] See slide 2 of Liz Ashall-Payne's presentation
[22] https://rudy.hosp.med.osaka-u.ac.jp/
[23] https://understandingpatientdata.org.uk/
[24] https://www.jst.go.jp/ristex/stipolicy/en/project/project31.html

Effective PPI around new technologies is a prime area for global collaboration, as evidenced by a newly launched project to investigate the best practice of Artificial Intelligence (AI) in healthcare, to ensure the benefits for all in the UK and Japan.[25] Led by Professor Jane Kaye, Director of The Centre for Health, Law, and Emerging Technologies (HeLEX) in Oxford, UK, and Professor Beverley Yamamoto of Osaka University, this inter-disciplinary project will consider how to develop a platform for engagement that can address issues of trust, responsibility, accountability and transparency, and influence the way AI technologies are implemented in healthcare.

In the meantime, questions remain around how to bring together PPI approaches and industry, and how to guide people and patients to trustworthy health technologies on the mass market. Trusted review services such as ORCHA have a role in this, by feeding back to application developers about how they could improve against frameworks that have been agreed with national and international partners for safety, efficacy and other areas. App 'libraries' of reviewed and approved products can help users and healthcare professionals to find appropriate products to fit individual needs, like a formulary for digital health. Systems like this that train healthcare professionals to confidently engage with digital health solutions and incorporate strong governance can help boost access to high quality digital health solutions and protect the public. Ultimately, inclusive development and prescribing approaches that bring together industry, healthcare professionals and members of the public to find person-centred solutions can help promote trust over the long term.

International collaboration is invaluable to improve global health, and can be enhanced through sharing and connection

Sharing data, methods and open science approaches for global benefit

Open science is a concept widely embraced by the research community as a highly progressive and productive approach to scientific collaboration and innovation. The concept of open science includes open data; indeed, data science is a discipline where the sharing of research materials and approaches is particularly well-practiced and embedded. Participants suggested that a challenge now exists to translate this sharing culture into health research systems, which are often more traditionally structured.

Dr Tamami Fukushi from AMED outlined how open science principles were being assimilated in Japan, following recommendations from the Science Council of Japan in 2016, to establish a research data infrastructure that ensures the openness of interdisciplinary research data.[26] In the same year, the 'FAIR Guiding Principles for scientific data management and stewardship' were published in *Scientific Data*.[27] These guidelines promote that data research systems prioritise the findability, accessibility, interoperability, and reuse of digital assets, and have been supported by the G20 international forum.

In 2018, Japan passed a law that established a system for the anonymisation of medical data by certified agencies.[28] This system aims to promote advanced R&D relating to health and medical care through secure and proper use of anonymised medical data. The research community are now mobilising to adapt data systems to the provisions of the Act, and by December 2019 two agencies were approved for anonymizing health data.

---

[25] https://www.law.ox.ac.uk/news/2020-01-28-ensuring-benefits-ai-all-designing-sustainable-platform-public-and-professional
[26] http://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-23-t230-en.pdf
[27] https://www.go-fair.org/fair-principles/
[28] http://www.japaneselawtranslation.go.jp/law/detail/?id=3441&vm=04&re=02

Concurrently, biomedical science-related working groups in the Science Council of Japan have developed infrastructure recommendations for sustainable life science data management.[29] Such high-level advocacy for, and guidance around, data connectivity is crucial to join up the multi-agency, disconnected health data systems that have existed in the UK and Japan.

This work can help to facilitate the activities of data networking and harmonisation groups such as HDRUK, and optimise the life-course research possible through medical biobanks. These groups should also carry forward open science principles. As part of its innovation gateway, HDRUK is embedding open standards and publishing research code and algorithms so that others can reuse them. Symposium participants considered this particularly beneficial for phenomics and medical ontologies, as this is such a vast, complex area in which researchers commonly duplicate effort. Discussions queried what research funders and data scientists can do to further incentivise sharing in the future, and how such systems can be scaled-up to share data safely and securely at the international level.

Mobilising young researchers to forge international connections

Symposium participants discussed that while improving international sharing of health data and research methods may be a protracted process, younger researchers are typically highly mobile and well positioned to share and enhance learnings between countries. JSPS offers a number of such opportunities though programmes across all fields from postgraduate level and above at institutional, group and individual levels.[30]

These include fellowships for UK researchers to conduct joint research activities with a host in Japan, and equally for early career Japanese researchers to work at an institution in the UK. They also support bilateral and multilateral projects, all with the overarching aim of fostering international scientific collaboration.

Expanding collaboration across the UK and Japan

Participants reflected on the respective strengths of the UK and Japan in respect of health data. The UK has delivered large-scale implementation of genomic initiatives including the 100K Genome Project and National Genomic Medicine Service, and is making advances in comprehensive data linkage. However, its data focus has traditionally been on health outcomes and episode data, rather than life-course.

Japan shows significant strengths in the organisation of complex systems and robust processes. For example, coding practices in medical insurance systems are a great research resource, and have potential clinical utility in other country settings.

These diverse strengths and areas of expertise demonstrate how much could be learned through UK-Japan collaboration to better tackle problems shared by both countries, such as ageing populations and a high burden of non-communicable diseases (NCDs).

---

[29] http://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-24-t279-1.pdf
[30] https://www.jsps.go.jp/english/programs/index.html

To realise the full potential of such joint working will require mechanisms for international data sharing, and thorough reciprocal testing of clinical and research tools in each population. For example, coronary risk scores work in Caucasian European ancestry, but need to be validated in Asian populations. Collaborations between Japan and the UK could allow for data sets to be combined or expanded, introducing greater genetic diversity in the data and making research outcomes more generalizable. Participants reflected on the scope to further explore what particular regulatory, ethical and other challenges exist to enhancing the sharing of health data resources internationally.

Overall, symposium participants noted that while differences exist between the UK and Japan, these are often subtle in the context of health research and a wealth of opportunities for collaboration exist. Strength also lies in the countries' differences for the purposes of study, and shared learning. Ultimately, the UK and Japan have similar public health philosophies, and the same aim of improving health in society and health care. Their advanced research and healthcare systems form grounds for strong research partnerships in the future.

Annex 1: Agenda

09.00 – 09.30	Registration

Opening

09.30 – 09.45	**Welcome messages from the organisers** Professor Dame Anne Johnson, Academy of Medical Sciences (AMS) Prof Nobuo Ueno, Japan Society for the Promotion of Science (JSPS) Dr Ryo Takagi, Japan Agency for Medical Research and Development (AMED)
09.45 – 10.00	**Opening comments by the Co-Chairs** Professor Jill Pell, University of Glasgow Professor Yuko Harayama, Tohoku University

Session 1: The health data landscape and resources in the UK and Japan
Chair: Professor Jill Pell

Aims: To explore the range, availability and accessibility of data to support health research, and how such data structures might be strengthened and networked globally.

10.00 – 10.15	**Improving the use and utility of UK routine data** Professor Andrew Morris, HDR UK
10.15 – 10.30	**Understanding and predicting disease through AI and mathematical analysis of health and medical data** Professor Eiryo Kawakami, RIKEN & Chiba University
10.30 – 10.45	**ATLAS: Making the UK's Cohorts Discoverable** Phil Quinlan, UKCRC Tissue Directory and Coordination Centre
10.45 – 11.00	**The Advanced Research Center for Innovations in Next-Generation Medicine (INGEM) and personalised medicine** Professor Chikashi Ishioka, Tohoku University
11.00 – 11.30	**Panel discussion**
11.30 – 12.00	Coffee & refreshments

Session 2: Health data for public health
Chair: Professor Yuko Harayama

Aims: To consider approaches for disease prevention and management through digital health platforms and how health data strategies can inform public health policy, while maintaining public trust and robust patient data principles through appropriate patient engagement.

12.00 – 12.15	**Determining & evaluating the quality of health apps** Liz Ashall-Payne, ORCHA
12.15 – 12.30	**Using wearables and machine learning to predict cardiovascular disease risk** Dr Aiden Doherty, Oxford Big Data Institute

12.30 – 12.45	**Involving patients in research** Professor Beverley Anne Yamamoto, Osaka University
12.45 – 13.00	**We need to talk about health data: some insights from Understanding Patient Data (unable to attend)** Dr Natalie Banner, Wellcome Trust [This presentation was not given as Dr Banner was unable to attend on the day].
13.00 – 13.30	**Panel discussion**
13.30 – 14.30	Lunch

Session 3: Health data for clinical decision-making
Chair: Professor Toru Suzuki

Aims: To examine how health data might inform machine learning and artificial intelligence solutions to improve clinical care, taking account of the ethical and regulatory considerations, challenges, opportunities, and pathways to implementation for this technology.

14.30 – 14.45	**Health data applications in clinical care** Professor Masao Iwagami, Tsukuba University
14.45 – 15.00	**Using machine learning and clinical AI to improve patient care** Professor David Clifton
15.00 – 15.15	**Ethics and regulatory policy for utilising health data toward open science in Japan** Dr Tamami Fukushi, AMED
15.15 – 15.45	**Panel discussion**
15.45 – 16.15	Coffee & refreshments

Audience discussion
Chairs: Professor Jill Pell, Professor Yuko Harayama, Professor Toru Suzuki

Aims: To summarise shared learning, challenges and opportunities, and identify collaborative ways forward for health data innovation.

16.15 – 16.45	**Audience discussion**

Closing

16.45 – 17.00	**Wrap-up comments by the Co-Chairs** Professor Jill Pell, University of Glasgow Professor Yuko Harayama, Tohoku University
17.00 – 19.00	**Networking reception** With a toast (kanpai) from Mr Satoshi Katahira, Minister for Economics at the Japan Embassy.