Paper A: Convenor's Update

14 October 2020

Thank you for your continued support of the UK Health Data Research Alliance. There has been outstanding engagement and leadership from colleagues across the UK, that has built momentum across several priority areas. This update highlights key updates and activities since the July Board meeting, within the context of the Digital Innovation Hub Programme developments.

Announcements

Next UK Health Data Research Alliance Board meeting, 20 January 2021

The next UK Health Data Research Alliance Board meeting is provisionally scheduled for **20 January 2021, 10:00-12:00**. Please save the date and details will be confirmed close to the time.

UK Health Data Research Alliance Symposium, 1st December

Alliance members are invited to attend the second UK Health Data Research Alliance Symposium on 1 December 2020. Please register here: https://www.hdruk.ac.uk/uk-health-data-research-alliance-symposium-december-2020/ . Given the current circumstances the event will be virtual.

The meeting will bring together people from across the UK Health Data research Alliance, Health Data Research Hubs, Innovation Gateway, industry and international players working in health data infrastructure. Attendees will have a chance to learn more about the health data infrastructure in the UK, both the public and the private sectors, and also to hear more about the international landscape and propose new ways to work together. A short update to the Board will be provided by Maria Koufali, Nottingham University Hospitals NHS Trust, and Alliance Symposium Advisory Group Member.

International COVID-19 Alliance Forum, 19 Nov 2020

On 19 November 2020 HDR UK will host a virtual event to showcase some tangible outcomes and achievements of the International COVID-19 Data Research Alliance and Workbench since the launch in June, including plans for driver projects.

Interested UK Alliance members are invited to attend the meeting. A Zoom link and more details will be sent via email in due course.

The International COVID-19 Data Research Alliance is an independent consortium of leading life science, philanthropic and research organisations that was set up to offer a co-ordinated global response to the COVID-19 pandemic. The Workbench is a new data platform that allows scientists around the world to discover, access and analyse multi-dimensional datasets whilst respecting confidentiality and privacy.

For members of the UK Alliance with high value COVID-19 related datasets, we encourage you to consider joining the International COVID-19 Alliance as well.

For more information about the International COVID-19 Alliance and Workbench contact Julia Levy Julia.Levy@hdruk.ac.uk.

New members

We are delighted to welcome two new members of the Alliance this month who signed their letter of intent to join the Alliance. We look forward to their engagement in the Alliance workstreams and to working with them making new data assets discoverable via the Gateway. A public announcement is scheduled in the coming weeks.

- **UK Brain Banks Network –** A coordinated national network of UK brain tissue resources for researchers to use. The Network supplies tissue samples to academic and industry researchers in the UK and abroad and have approval to provide tissue samples to research projects and pilot studies. The banks also work together to agree common standards of operations and harmonise protocols for consent, tissue handling and storage, quality indicators and the application process for access to tissue samples and associated data.
- **Early Detection of Disease Research Project UK –** A major Industrial Strategy Challenge Fund initiative aimed at improving early detection and prevention of major chronic diseases including cancer, metabolic diseases and neurodegenerative diseases. The study plans to recruit a cohort of up to 5 million healthy volunteers linked to routine NHS health records, representative of the whole UK population. The cohort will enable studies developing new tools for disease prevention and diagnosis such as circulating tumour DNA, biomarkers, proteomic panels, metabolomics, imaging tools and digital signatures. It is anticipated that the programme will run over at least the next 20 years as the value of the study will grow as data and samples accrue over time.

National health data research capability to tackle the COVID-19 pandemic

SAGE fortnightly reporting

Reporting to SAGE and the UKRI/DHSC rolling panel takes place on a fortnightly basis in line with SAGE briefings. The latest report can be found here (Paper B.1).

Since last board meeting, SAGE has endorsed five proposed recommendations relating to:
- Coordinated approaches to support and accelerate access to restricted national datasets
- Accelerating insights on long term impacts of health and social care changes during COVID-19 lockdown on major diseases
- Providing insights on understanding high rates of COVID-19 infections in minority groups and care homes.

National Core Studies Programme

Funding for the National Core Studies Programme commissioned by Sir Patrick Vallance to answer the immediate policy and operational needs for the UK response to the pandemic during the winter started on 5 October 2020 and will run for six months to March 2021. It will be formally announced within the next few weeks.

The National Core Studies cover six key areas to enhance our understanding of the SARS-COV-2 virus and the UK's pandemic response at the national level. The purpose of these studies is to accelerate insights essential to support research vital for HMG policy developments and operational decision making.

The National Core Studies include:

- **Epidemiology and Surveillance** led by Professor Ian Diamond (ONS). Collecting and analysing data to inform safe level of restrictions and protection against imminent outbreaks.
- **Clinical Trials Infrastructure** led by Professor Patrick Chinnery (MRC) and Divya Chadha Manek (NIHR). Establishing infrastructure to run large scale trials for Covid-19 drugs and vaccines.
- **Transmission and Environment** led by Professor Andrew Curran (HSE) Understanding and mitigating transmission of the disease in workplace, transport and public places.
- **Immunity** led by Professor Paul Moss (University of Birmingham) Understanding immunity against Covid to inform back-to-work policies.
- **Longitudinal Health** led by Nish Chaturvedi (UCL) Understanding the impact of Covid-19 on long term health to inform the design of mitigating policies.
- **Data and Connectivity** led by Professor Andrew Morris (HDR UK working with ONS) Making UK-wide health and administrative data available for linkage and accessible to catalyse Covid-19 research.

The Data & Connectivity Study underpins all the other five studies and will define the infrastructure, data and services needed to create a robust and trustworthy health and cross-sectoral data capability that will support research in the key areas specified. Our aim is to build upon existing partnerships and capabilities, including the work we have been convening for SAGE. In addition to national bodies from across the four nations of the UK Health Data Research Alliance, the work will also involve the Gateway, as well as Secure Research Service (ONS), and working in partnership with ADR UK.

Paper B.2 provides a more detailed overview of the National Core Studies Programme.

Diversity and Inclusion

Following publication of HDR UK's recent Statement-on-anti-black-discrimination we committed to developing focused actions to improve diversity and inclusion of datasets for research and innovation and of people in the HDR UK community. We are keen to show tangible actions now with two focussed initiatives, the Black Internship programme and the creation of a Data diversity Special Interest Group.

Black internship programme

We propose a programme of work in conjunction with the UK Health Data Research Alliance to create a Black Internship Programme and support growth of health data science leaders. Black data scientists are currently under-represented within the health data science community. We believe we can play a role in providing them with the opportunities and expectations to lead from early in their careers, developing skills needed at key stages and providing tools.

A letter to invite Alliance members to participate in the initiative was sent on 22 September 2020 (Paper C). We received a wealth of positive responses and have now more than 15 participating Alliance members. We look forward to working with members on this exciting initiative.

For more information about the Black Internship Programme, please contact Caroline Cake Caroline.Cake@hdruk.ac.uk.

Data diversity: ethnicity coding

At the last Alliance Board meeting, the Board agreed with the collective ambition to improve the quality of ethnicity coding and use data to help increase the representativeness of research studies. The initial focus on improving understanding of current practices and first questions around collection of representative data and main challenges were discussed at the last Data Officers Group meeting on September 9th.

It was agreed that a Special Interest Group on Data Diversity would be created to review current landscape and issues around ethnicity coding towards enacting the Alliance Board commitment to

improve consistency and quality of ethnicity coding, enabling data use to increase representativeness of research.

We are now seeking volunteers to join this group, both from the data custodian and user communities.

ACTIONS FOR MEMBERS: Please contact Paola.Quattroni@hdruk.ac.uk if you want to participate in the Special Interest Group on data diversity.

Innovation Gateway

Overview of progress

The Innovation Gateway development is progressing at pace, with new content onboarded every month. In October, the Gateway reached the 500 datasets milestone, a remarkable achievement from 144 at the beginning of the year. It also has:

- 350 registered users
- 48 Tools
- 517 papers
- 340 Projects, including 190 training courses

Featured collections are also now displayed on the landing page of the Gateway, with each Health Data Research Hub having its own collection, enabling researchers to find all artifacts associated with a specific Hub quickly and easily.

In parallel, the number of searches on the Gateway increases steadily, with over 7,000 searches in the last month.

The Innovation Gateway is nearing the end of the first milestone (31 October 2020), and new functionalities have been delivered, such as:

- **New Data Access Request process** based on the ONS Five Safes. This is currently being tested by some early adopters (both hubs and national Trusted Research Environments as part of the National Core Studies Programme). Their feedback will be taken into account for future versions of the Data Access Request process.
- **Integration with the Phenomics library:** Ability to refine searches using Phenomics where datasets have been tagged.
- **Metadata quality tool:** easy visual representation of the completeness of the metadata information.
- **Technical metadata:** in depth description at variable levels of the dataset.
- **Gap analysis:** highlights searches that do not deliver any results, potentially providing information as to gaps in the market.
- **New dashboards**

In November, the Gateway enters the second part of its delivery plan, Milestone 2, until end of April 2021. During this time, the development team will focus on delivering new functionalities such as

Trusted Research Environment integration, ID federation, production of cohort builder, improvement to the data access request process, and testing all existing and new capabilities with the National Core Studies teams.

Finally, the public and patient engagement strategy is being reviewed and updated building on learnings from Milestone 1.

Access requests workflow

Work is progressing at pace to streamline, improve and accelerate current data access processes and support both researchers and data custodians in facilitating data driven research.

In the past few months we have been working closely with several Alliance members to build a Data Access Request module as part of the Gateway and provide a solution for those data custodians listing datasets in the Gateway and interested in using the Gateway interface to provide access to data for research.

The past month has also seen a speedy increase in engagement with the national Trusted Research Environments/data custodians (NHS Digital, SAIL Databank, Public Health Scotland, Northern Ireland Honest Broker Service and ONS) as part of the National Core Studies Programme. The Programme aims to provide a streamlined access to priority datasets needed to answer questions relating to the five thematic National Core Studies. We have been able to leverage HDR UK's ongoing work to support the National TREs and are now working closely with these organisations to ensure the Gateway module in development is fit for purpose.

Alliance members early adopters and national TREs are committing to adopt the Gateway Data Access module for their processes and are now working closely with us to develop the first iteration of the module.

To date the following organisations have been involved in this project:

- Healthcare Quality Improvement Partnership, HQIP (Alliance member)
- CPRD (Alliance member – part of initial harmonisation work)
- University Hospitals Birmingham NHS Foundation Trust (Alliance member, INSIGHT and PIONEER)
- Cystic Fibrosis Trust (Alliance member)
- Barts Health NHS Foundation Trusts (Alliance member)
- INSIGHT (Hub)
- Discover-NOW (Hub)
- SAIL Databank (Alliance member, BREATHE hub, National Core Studies TRE)
- NHS Digital (Alliance member, NHS Digitrials and National Core Studies TRE)
- HSCNI (Alliance member, National Core Studies TRE)
- NHS Scotland/Public Health Scotland (Alliance member, National Core Studies TRE)
- Health Research Authority (HRA) – key stakeholders

The first iteration of the data access module for the National Core Studies is expected in early November 2020.

ACTIONS FOR MEMBERS: Please contact Paola.Quattroni@hdruk.ac.uk if you want to be involved.

Metadata Onboarding

The focus of this workstream is making datasets discoverable through the Gateway and improving the quality of the metadata available. MetadataWorks are continuously engaging with Alliance members to onboard new datasets and improve the quality of metadata per each dataset listed in the Gateway. To date, 19/33 Alliance members have listed at least one dataset in the Innovation Gateway (410 datasets) and 7/7 Hubs have listed at least one dataset (83 datasets). Of these, all have summary level metadata and 185 have technical metadata. The graph below shows the average of metadata quality per Alliance member. Alliance Board members are encouraged to review how the metadata for their organisation's datasets are represented on the Gateway and provide feedback.
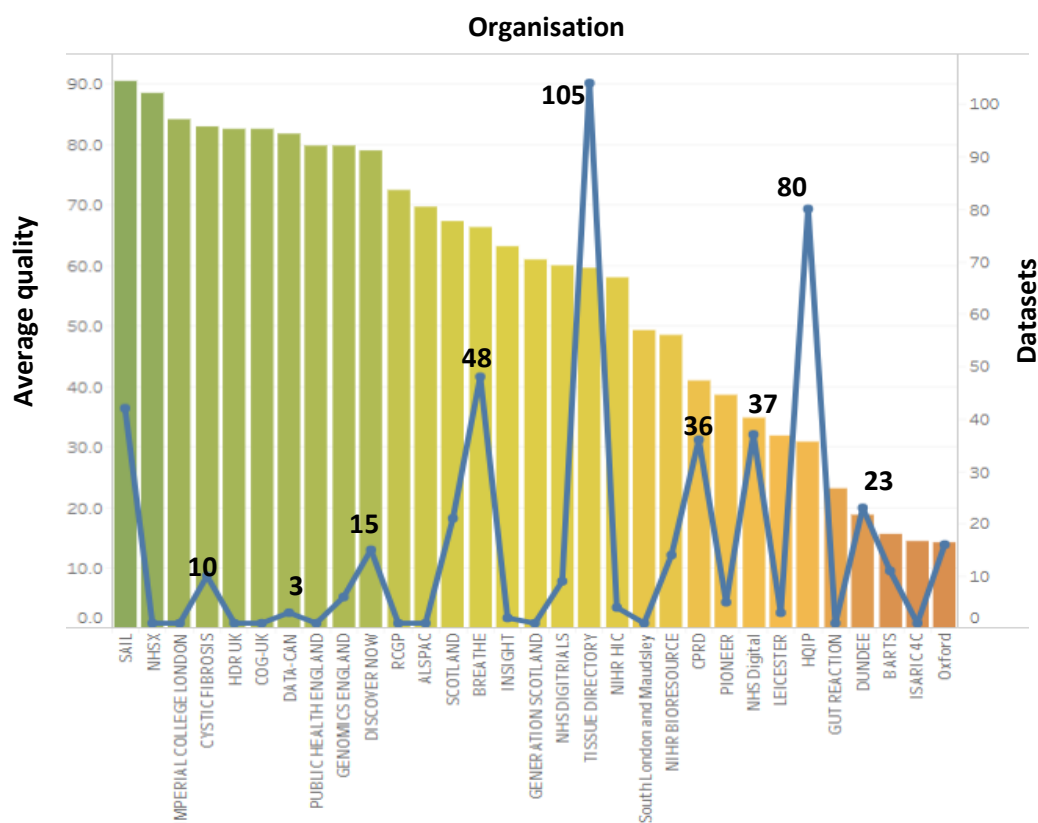

Figure 1. Metadata quality per Alliance member (as of 10.10.20). 'Metadata quality' is the average quality for all datasets listed by each Alliance member.

ACTIONS FOR MEMBERS: A number of Alliance members have not yet had capacity to engage with the process and are encouraged to do so.

Data Utility

The Data Utility project aims to provide a common means of articulating the usefulness of a dataset for a given purpose. This is intended to support user activity through the Gateway, describe the impact of any data improvement activity by Hubs, and see how learnings can be applied across the UK Alliance. It includes the development of Data Utility Evaluation framework and testing open and open-source tools for data profiling. We thank all Alliance members who have been contributing to this work in the past months. The Data Utility Evaluation framework has now been published on the HDR UK website.

Activities relating to Data Utility and two case studies will be discussed during the meeting.

In addition, a paper on the 'Principles of Data Standards' has been developed in consultation with data officers across the HDR UK community (the Data Officers Group). We received feedback from 50 individuals across more than 30 organisations including patients and the public. The vision of the Principles of Data Standards is to encourage behaviours to improve data usefulness and usability through the provision of clear guidance and recommendations. This will greatly increase the possible impact and scale of research across the Hubs and Alliance.

We are now conducting a survey to understand the data standards landscape. The aim is to get a sense of what data models and supported standards are in use across organisations. We invite all Alliance members to contribute to this short survey here.

Practitioner, Public & Patient Involvement and Engagement (PPPIE)

To support the growing and wide range of involvement and engagement activities across the Institute, we launched our new network, HDR UK Voices. With an initial focus on recruiting members of the public to join the network, it offers anyone who is interested an opportunity to influence and shape our work. By joining the network, HDR UK Voices members will be informed of new opportunities as and when they arise, and they will have a variety of options to choose from to suit their needs. For example, they could help us write engaging information about our work, share their views on challenging and complex topics related to health data or help shape research projects.

The benefit to having an umbrella network of HDR UK Voices is that it can be used to enable other groups and populations such as researchers and practitioners to join and also ensure their voice is heard and incorporated as we develop our work. We will shortly be starting work to draw in practitioners that cover a range of professions and specialisms to sit within our network. We see this group acting as a critical friend to the Institute and would ask Alliance members to consider areas of their work in which the practitioner voice is needed and would have a high level of impact.

Current priorities are focused on running involvement and engagement activities around data access and how best to implement the outputs. Working closely with the HDR UK Public Advisory Board we

are developing a set of recommendations on how patients and the public should be involved in the data access decision making process (we aim to publish Nov 2020). In collaboration with useMYdata members, we are working on ways in which to better communicate work around data access to the public with a first priority being developing a lay explanation of Trusted Research Environments. Moving forward, we will build on this dialogue with patients and the public to consider work around data access/release registers and activities related to the Data and Connectivity programme of work. This includes, but is not limited to, public perceptions around private sector data, moving data across the four nations, how we best communicate research to the public from data access -> insight -> outputs and moving beyond the COPI notices. We will work with other organisations who may already be planning work in these areas to not duplicate work.

To gain a better understanding of current PPIE practices and explore areas for further collaboration across the Alliance members, a survey that will act as a mapping exercise has been sent to your PPIE leads. Where we do not have contacts for your PPIE leads we will email you directly for your guidance in finding the right person to connect with in your organisation. We will hold an Alliance Communications and PPIE Leads meeting in early-mid Nov 2020 and invites will be sent shortly.

ACTIONS FOR MEMBER:

Please encourage your PPIE leads to complete the mapping exercise survey by **Friday 30 October 2020**.

Please share the HDR UK Voices Network with your contacts and encourage them to sign up.

Health Data Research Hubs

The development of the Health Data Research Hubs continues at pace while the impact on the Hubs from the pandemic has been somewhat normalised. The opportunities and initiatives developed over this period are being absorbed into Hub operations, Hub services and sustainability planning.

The Hubs have been instrumental in providing data services to key decision makers during the past three months including:

- Working with Hub consortium members HUMA's to use their remote monitoring service. The Discover-Now/Huma team (in collaboration with NHSx) designed 'Community Virtual Wards' where clinicians could rapidly identify deteriorating patients, react quickly to avoid further complications, and better manage system capacity and improve overall survival rates. The team deployed in two health systems in both primary and secondary care and by using patient reported data from the App were able to connect into linked identifiable data which was then deidentified and made available to Discover-NOW researchers. Within two weeks of launch and at the peak of COVID-19 in London, there were over 150+ patients on the Community Virtual Ward with 96% of patients using the app daily as prescribed. The team estimate the service increased secondary care efficiency by 500% based on HCPs making over 150 calls to patients per day which reduced to 30 calls per day using this remote monitoring service. The Discover-NOW and NHSx team led what is believed to be one of the fastest ever implementations of a digital service in the NHS – going live in just four weeks.

- PIONEER facilitated data collection in four regional NHS hospitals within Birmingham by designing and implementing a fully electronic COVID-19 screening and management system and real-time COVID dashboard. This system guided clinical management while facilitating collection of clinical and biomarker data, by capturing details as to where a particular patient lived, presentation severity and their hospital journey from admission to outcome. The dashboards allowed public health, infection control officers and senior clinicians to visualise the distribution of positive, negative, and suspected cases around hospitals in the city.

The Milestone 2 assessment of Hubs is due in March 2021 and work is well underway to sign-off the approach, timescales, deadlines, activities, panel, assessment criteria and dissemination of the results of this crucial milestone to which Hubs must:

- Evidence that the quality of the datasets has been improved (curated) and that the curated data is discoverable through the Gateway
- Provide publishable enhanced service case studies (e.g. from industry, academia, and NHS) that demonstrate impact and value to researchers and innovators from a range of sectors, and to patients and populations
- Provide evidence that the Hub is continuing to engage and involve patients and the public in a meaningful manner.