UK Health Data Research Alliance Council Actions and meeting notes

Tuesday 10th October 2023 (10:00 – 12:00)

Immediate actions for Alliance members:

- Share information on DigiTrials recruitment services with your networks.
- Complete the expression of interest form if you would like to host an intern for the 2024 HDR UK Black Internship Programme cohort.

Actions for HDR UK secretariat:

- Share slides from the Council meeting with attendees.

Meeting notes:

Welcome, introductions and opening remarks.

Paola Quattroni, HDR UK, Head of Alliance Engagement and Strategy, welcomed the attendees and introduced seven new members to the Alliance, including:

- Understanding Patient Data
- NHS Education for Scotland
- University of Sheffield
- University Hospital Southampton NHS Foundation Trust
- The Generations Study
- use My data
- Alder Hey Children's NHS Foundation Trust

Alliance overview and programme priorities

David Seymour, Director of Infrastructure & Services, HDR UK presented slides to update on the Alliance programme priorities and governance.

Sudlow review: Preliminary Findings

Cathie Sudlow, Chief Scientist and Deputy Director, HDR UK, presented slides outlining the preliminary findings of the Sudlow Review. Q&A and discussion followed.

- Some participants expressed that the priority should be ensuring the data access and information governance infrastructure is fit for purpose. Cathie added to this by highlighting the importance of collaboration and incentivising better working across the ecosystem.
- There was discussion about the national audit data assets and the importance of linking these in the same system, to avoid duplication and redundancies. Linking at a national scale is the best way to recognise these redundancies and remove them from the system.
- Some queried whether there is desire to create a more comprehensive diagnostic element of outpatient coding and if there is an opportunity to highlight the benefit of using clinical terminology to tackle these issues. Cathie clarified that the observation came from users of data rather than those who fully understand exactly how it can be surfaced into national commissioning or other systems. Digitally mature systems allow clinicians to surface this knowledge regionally or nationally and the same applies to inpatient data, which is coded more comprehensively. It can be achieved through GP data as the information is coded. However, it's not clear that the routes are in place to link this data.
- It was inquired whether the review plans to discuss the imminent Federated Data Platform (FDP) and the potential for confusion and concern of a parallel review of data for research and insights and a data platform for operational delivery. Cathie clarified that the FDP is England-specific and the clarity of how it is distinct from the emergence of SDEs is only just becoming clear. The FDP is about building software tools to inform the day-to-day operational capabilities within the NHS that can be used regionally and when appropriate, nationally. The complete FDP service has not been fully outlined yet. There should be more information available in the coming weeks.
- Participants commented on the experiences of clinicians and their individual responsibility to input data. This might lead to unclean, incomplete data recording across the ecosystem. For example, those using that data for research find that many rare diseases have less useable data. Cathie acknowledged that the way in which data is captured in the system is not perfect and GPs have a more mature system for entering data than other areas of the sector. The more sophisticated these electronic patient records (EPR) become, the more success we'll see in rare disease research and other areas of research.
- It was noted that there is a real opportunity in outpatient diagnostic coding as there is a rollout of more modern EPR systems across NHS trusts in England. Improving this coding will better enable data flow and research. There is also an opportunity to use natural language processing and generative AI to assist with coding and dealing with unstructured data.

Horizon Scanning: updates from across the UK

Paola welcomed representatives from across the four UK nations to provide updates from their organisations. Due to conflicting priorities, a number of representatives were unable to contribute and a summary of the attending representatives is written below.

Nora Cooke O'Dowd, Chief Data Officer, Research Data Scotland provided an update.

- There has been considerable interest surrounding the new Researcher Access tool.
- There have been recent website updates and metadata catalogue integration.
- A case management system, scheduled for a February launch, has prototypes available on the Research Data Scotland website. The tool undergoes iterative updates and will feature a triage approval process, commencing with a standard pathway and an advanced pathway (currently in development).
- The standard pathway offers access to 9 Public Health Scotland datasets.
- Exploration of prioritised sorting of datasets within the new pathways is in progress.
- Paul Jackson, a recent addition to RDS, will focus on broader data access concerns.
- Collaborative efforts with HDR UK and the Innovation Gateway team encompass user research and service designs.
- Robust PPIE (Patient and Public Involvement and Engagement) initiatives are underway. This includes the joint management of the 'Scotland Talks data' PPIE panel with the Scottish Centre for Administrative Data. Cooperation with PEDRI (Public Engagement in Data Research Initiative) and Understanding Patient Data is ongoing. RDS has funded 8 projects to promote public understanding and widen participation. There will be a showcase of these projects in the new year.
- There has been active involvement in DARE UK PPIE projects, such as SACRO (Semi-Automated Checking of Research Outputs).

Michael Chapmam, Director of Research and Clinical Trials, NHS England provided an update on behalf of NHS England.

- Momentum is building around the Data for R&D programme, both on secure data environments (national and sub-national). There is a partnership working with HDR UK in many areas, including developing the Innovation Gateway as the front door to the system. NHS DigiTrials, initially developed as part of the Digital Innovation Hub programme, are seeking expressions of interest for four further studies to help develop their recruitment service pilot. This is an opportunity to use the service behind recruitment to NHS Galleri and Our Future Health. Deadline for expression of interest is 31 October 2023. More information is available here.
- Work is still ongoing following the merger of NHS Digital and England. An Advisory Group for Data is in development. This group will advise on internal and external access arrangements for confidential information and will tackle data issues in a transparent way. Terms of reference are yet to be finalised and recruitment is ongoing to supplement the former Independent Group Advising on Release of Data (IGARD). Minutes from this advisory group can be found here.
- Finally, Garry Coleman is developing how NHS England provide transparent assurance on the use of data (building out from the Audit function) and is keen to work with Alliance on how to achieve this.

Panel presentations: Tackling health inequalities through data science (chaired by Paola Quattroni, HDR UK Head of Alliance)

Dona Reddiar & Phinehas Arkorful, HDR UK presented slides on Tackling the Underrepresentation of Black Scientists through the Black Internship Programme.

- Paola Quattroni invited questions and comments from the attendees.
- Dona shared an expression of interest form for Alliance member organisation to complete if they would like to be a host organisation for the 2024 Black Internship Programme cohort.
- Sarah Cadman shared information on the Amazon Web Services mentorship programme that is available to all interns and alumni. The alumni network also provides networking and mentoring opportunities for the alumni and existing interns.
- Monica Jones commented on the opportunities for apprenticeship and conversations with Cambridge Spark around improving existing opportunities across the ecosystem. There is an appetite to align existing opportunities and to expand them further across other fields.

Mavis Machirori, Senior Researcher, Ada Lovelace Institute provided a verbal presentation on the work at the Ada Lovelace Institute regarding health inequalities impacted by system designs.

- The institute has been engaged in a collaborative partnership program with the Health Foundation since the onset of the pandemic. This program, which recently concluded, primarily focused on understanding the impact of data-driven systems on health inequalities. During the pandemic, many technologies were deployed to enhance health services with an assumption that these innovations would bring positive outcomes, often overlooking their unintended consequences on health inequalities.
- A survey of over 2,000 UK adults was conducted, specifically addressing the use of deployed systems, including digital technologies like chatbots and access to GP services. The findings revealed a significant disparity in access to these services, stemming from barriers such as limited access to certain hardware and a lack of knowledge in utilising these resources. Individuals earning less than £20,000 or belonging to ethnic minority groups were notably less likely to benefit from these services. Many survey participants were aware of these disadvantages, and there were specific concerns for individuals where English was not the first language.
- The work also delved into the systemic aspects contributing to these disparities. The Ada Lovelace Institute collaborated with commissioners and service providers, as well as organisations with access to extensive databases. A lack of communication between national and local systems was evident. Main areas of concern included the quality of data input, data standardisation, data cleaning, and data sources.
- One significant finding was that organisations committed to addressing inequalities lacked a clear understanding and metric of the health inequalities they aimed to tackle.
- Local authorities and local services have more nuanced understanding of their populations and their specific needs. This local insight is often not effectively communicated to national organisations, leading to difficulties in identifying and addressing these needs. Additionally, there

are redundant effort at the national level to understand local needs that were already known to local organisations.
- The Ada Lovelace Institute conducted peer research involving the training of six researchers to engage with their communities, conduct interviews, and gain insights into the needs of their local areas. This revealed that many system designs failed to consider the unique requirements of certain communities and made incorrect assumptions about their access to resources.
- Without nuanced context, systems can often be designed with the best of interest but still enable further health inequalities. Further information on 'Access denied?' is available on the Ada Lovelace Institute website.

Fatemeh Torabi, Data Scientist, Swansea University presented slides about harnessing routinely collected ethnicity data to advance diversity in data and analytics, on behalf of the Alliance Diversity in Data – Ethnicity Coding working group.

- Paola Quattroni invited questions and comments from the attendees.
- David Seymour inquired about the accessibility of the code featured in the presentation slides for use by other researchers. Fatemeh explained that they have made available a published guide detailing the categorisation of code in Wales, which required the adoption of various codes to facilitate this process.
- Paola concluded by reminding attendees that the Ethnicity Coding working group are actively working on a community-led paper on ethnicity coding. The paper is being drafted to capture emerging recommendations for recording ethnicity data across the settings in the UK to better enable research. The recommendations will be published shortly. There are currently high-level outlines on the Alliance website, which will be further refined in the paper.

Appendix

Attendees

Name	Organisation
Adam Marsh	Optimum Patient Care
Alison Elderfield	HDR UK
Alex Knight	HDR UK
Andrew Morris	HDR UK
Andy Boyd	UK Longitudinal Linkage Collaboration
Anna Dominiczak	Scottish Government
Ariane Alamdari	Leeds Teaching Hospital NHS Foundation Trust
Ben Crosby	HDR UK
Cathie Sudlow	HDR UK
Chris Carrigan	use MY data
Chris Gush	Healthcare Quality Improvement Partnership
Claire Argent	HDR UK
Claire MacDonald	Manchester University NHS Foundation Trust
Claire Richardson	Gloucestershire Hospitals NHS Foundation Trust
Clara Fennessy	HDR UK
Colin Brown	NHS Education for Scotland
Dal Bansal	NPIC (National Pathology Imaging Co-operative
Damon Chow	HDR UK
David Harrison	iCAIRD (Industrial Centre for Artificial Intelligence Research in Digital Diagnostics)
David Seymour	HDR UK
Dona Reddiar	HDR UK
Elliot Bridges	Human Fertilisation and Embryology Authority
Em Loveridge	Optimum Patient Care
Fatemeh Torabi	University of Swansea
Fergus McDonald	DARE UK
Garry Coleman	NHS England
Geoff Hall	University of Leeds
Graham Prestwich	National Pathology Imaging Co-operative
James Peach	UK BioIndustry Association
Janet Valentine	Association of British Pharmaceutical Industry
Jo Knight	University of Lancaster
Kathy Harrison	DataLoch
Kay Snowley	HDR UK
Laura Cocking	NIHR BioResource

Liz Perraudin	Association of Medical Research Charities
Louis MacGregor	The Brain Tumour Charity
Lucy Lennon	University College London
Luke Readman	NHS England
Mavis Machirori	Ada Lovelace Institute
Melissa Lewis-Brown	Cancer Research UK
Michael Chapman	NHS England
Michele Afif	The Brain Tumour Charity
Monica Jones	University of Leeds
Nada Karrar	HDR UK
Nathalie Kinsgton	NIHR BioResource
Nick Kennedy	Royal Devon University Healthcare NHS Foundation Trust
Nicola Perrin	Association of Medical Research Charities
Nicola Armstrong	Health and Social Care NI
Nicola Hamilton	Understanding Patient Data
Nora Cooke O'Dowd	Research Data Scotland
Olivier Roth	BioIndustry Association
Paola Quattroni	HDR UK
Pete Stokes	Bennett Institute for Applied Data Science
Phinehas Arkorful	HDR UK
Rachel Brophy	HDR UK
Rob French	Cardiff University
Sarah Cadman	HDR UK
Stephen Burrows	HDR UK
Sue Mason	The University of Sheffield
Susan Hodgson	CPRD
Tim Hubbard	HDR UK / Genomics England
Tom Barlow	Scottish Government
Uwaye Ideh	HDR UK
Yvonne Silove	Healthcare Quality Improvement Partnership