World Health Summit 2022: How to Achieve a Global Health Data Space

IN A ROUNDTABLE discussion at the World Health Summit (WHS) 2022, located in Berlin, Germany, and also available digitally, experts in the field of health data debated barriers, opportunities, and next steps for achieving a global health data space. Key concepts such as anonymisation, infrastructure, data quality, and data standardisation were analysed by the expert representatives from pioneering international data projects, including the World Health Organization (WHO), EBRAINS, and ERA-NET NEURON.

INTRODUCTION

The expert debate was chaired by Petra Ritter, Berlin Institute of Health, and Bernstein Centre for Computational Neuroscience Berlin, Germany, who opened the session by introducing the key aspects of human health data and explaining what a global health space might mean.

Health data are sensitive, personal data, and can present a risk to the patient if they become public. Brain scans, for instance, can provide insight into an individual's demographic, personality, intelligence level, attention span, and even their degree of satisfaction with their own life. This all represents information that can easily be abused or monetised and, therefore, needs to be protected. Protection can come in the form of anonymisation, when all unique individual features are removed from patient data, or pseudonymisation, where names and birthdates are removed or replaced. However, neither system is perfect, Ritter highlighted, as anonymisation removes information relevant for precision medicine and personalised therapy research. The biometric information remaining in pseudonymisation can also allow for potential re-identification.

WHAT VALUE DOES A GLOBAL HEALTH DATA SPACE PROVIDE?

Paweł Świeboda, Chief Executive Officer of EBRAINS AISBL and Director General of the Human Brain Project, broke down the 'tip of the iceberg' problem to explain to the audience why a global health data space is needed and the value it could provide. Currently, only a fraction of usable health data is being put to proper use. For the primary data user, the proper use of data means enhancing the quality of processes, health service outcomes, and continuity of care for patients. For the secondary data user, this means building a knowledge base and improving research effectiveness in academic institutions and industry.

The value of a global health data space has also been emphasised by the recent shift in ideology towards value and outcome-based health care, which Świeboda described as "the single most powerful driver" behind efforts to push a global health data space. Increased representativeness of data is an additional horizontal benefit that would be provided by a global health data space. Global cohorts ensure that diverse populations can benefit from

"Currently, only a fraction of usable health data is being put to proper use."

targeted treatments. Świeboda took the time to emphasise the need for equal representation, highlighting current disproportionately high levels of Western patient cohorts in Alzheimer's disease research.

A further key issue is the current reproducibility crisis in medical research, fragmentation of data sets with uneven quality, different governance models and policies for reuse, and sharing all impact to enable researchers to repeat and reproduce findings. Świeboda emphasised that a global space for health data sharing could aid in addressing this crisis.

CURRENT BARRIERS TO A GLOBAL HEALTH DATA SPACE

However, before reaching this goal of a global health data space, multiple barriers currently stand in the way. During the roundtable discussion, each of the experts shared and explained challenges that their organisations had encountered. Świeboda introduced some key current challenges which were explored in the Towards European Health Data Space (TEHDAS) report, published in February 2022. The report

listed key barriers to data sharing within Europe, such as different interpretations of key terms, including anonymisation and secondary use, and different understandings of General Data Protection Regulation obligations.

Mehdi Snène, Chief Executive Officer of the International Digital Health and AI Research Collaborative (I-DAIR), works on building research networks and growing health research in the Global South. Snène highlighted that infrastructure still represents a significant barrier for global health data sharing. Sharing high-quality data requires that researchers have access to equipment for data collection, software for processing statistics and analysing data, and hardware for securely storing data. Snène highlighted that this is an ongoing challenge in the Global South, which needs addressing to allow the development of South-South and South-North collaborative consortia for data sharing, education, and training.

Ritter emphasised that an overarching challenge for protecting health data arises from varying legal definitions between regions. Definitions of anonymisation and pseudonymisation vary across different jurisdictions. Ethical and mobile frameworks that provide the basis for legislation surrounding health data protection and sharing also vary between regions. This creates challenges with data interoperability, and can reduce the security of data sharing.

NEXT STEPS TOWARDS GLOBAL DATA SHARING

The expert representatives from each organisation provided a detailed overview of how they were working towards tackling these challenges to increase the ease of global data sharing. Marlies Dorlöchter, Head of Division International Health Researcher, DLR-PT, Germany, and Coordinator for ERA-NET NEURON on behalf of the German Federal Ministry of Education and Research, shared the two central pillars the ERA-NET NEURONS project uses, with a view to creating a joint global data space. Firstly, the funding programme of the organisation has stringent review steps, examining the soundness of statistical analysis and the quality of data, with data management plans mandatory in the projects it endorses. Secondly, capacity building and training is also required, with workshops, webinars, and data training for researchers participating in projects. In this way, the project can ensure the creation of high-quality data that is interoperable and shareable. Dorlöchter explained how they felt the depth of these requirements created excellence in the outcomes, with 1,133 peer reviewed publications arising from the 80 research projects funded by ERA-NET NEURONS.

Świeboda discussed the European Health Data Space, highlighting how this project has focused on innovation to address legal and infrastructure barriers. The European Commission (EC) project aims to clear the way to greater accessibility to health data, and to achieve this it has created new legislation that, instead of replacing the General Data Protection Regulation, will act on top of it to make data more freely available. Świeboda emphasised that this legislation is based on a core nexus to health data sharing, of patient trust versus accessibility. The bulk of the legislation created as part of the European Health Data Space has focused on creating greater trust in citizens to encourage them to share data.

EBRAINS is a research infrastructure, and part of the European Health Data Space. It is a collection of tools and services underpinned by computing and storage capacity. The infrastructure allows researchers access to fair, high-quality data, with both machine- and human-readable metadata. The high quality of the data is ensured by curation from expert neuroscientists who monitor it. Access to data is restricted, and individuals requesting access must define the purpose of data use, ensuring high levels of protection and security.

Both Snène and Świeboda highlighted the importance of building human capacity and communities for their respective data sharing projects. Snène highlighted that for Global South nations where I-DAIRs is active, there is a need to tailor capacity development programmes to build knowledge in individuals and institutions. Świeboda similarly emphasised the importance of building collaborative communities in the EBRAINS project to enable trust, accessibility, and interoperability of shared data.

CONCLUSION

The roundtable discussion, which enabled key figures from international health data sharing organisations to share their experiences, revealed key themes. Challenges presented by the current lack of global health data sharing such as the 'tip of the iceberg' challenge and reproducibility crisis were highlighted, in addition to the current barriers including infrastructure, legislation, and human capacity that stand in the way of creating this space. However, the experts agreed on key immediate and efficacious steps, such as increasing data interoperability, developing global data quality frameworks, and writing legislation to build citizen trust, that could pave the way to a global health data space. ●