UK BioIndustry Association

Unifying Health Data – Sudlow Review

Executive Summary

This submission discusses the current challenges and potential solutions to unifying health data in the UK.

Challenges

- National data broken into subscale datasets and organisations: most data is held by sub-national organisations without the resources to allow secure, efficient data access
- Government initiatives supporting the continuation of subscale datasets and organisations - recent and ongoing initiatives (HDR UK hubs, Data for R&D programme) have supported local, regional and disease-specific data access
- Fragmented systems and formats - numerous Secure Data Environment (SDE)s operate independently, using different formats and standards for data storage and exchange, leading to compatibility issues.
- Lack of common infrastructure - there is no centralised health data repository, making it challenging to share and access data, particularly for SMEs
- Variable data quality - inconsistent data quality across different organisations impedes research and usage
- Public-private data disparity - health data is stored by both public and private entities, which follow different priorities and incentives, resulting in a scattered health data landscape

Proposed Solutions

- National platform - government support should be moved from subscale SDEs to one single platform
- Interconnectivity - a unified system with a single access point, standard data models, interoperability, and data search across all SDEs is recommended
- Standardised information and KPIs - clarity on data availability, uses, and limitations is essential, with standard KPIs for all health data organisations focusing on data and capabilities
- Standardised SLAs - clear service expectations, decision-making transparency, standardised application processes, and agreements are needed
- Accreditation - similar to data safe haven authentication, accrediting users and organisations would ensure appropriate data analysis training and experience. Central control or standardised data collection would help streamline multiple data access

- Uniformity of technical abilities and security - SDEs should support innovation and offer clear SLAs outlining technical services. Access to a public secure data environment and innovative proprietary analytical pipelines (PAPs) should be available to all users.

Introduction

The life science sector forms a vital part of the UK's innovation system. Small to medium-sized enterprises (SMEs) are the driving forces behind this industry's growth, contributing to the economy's vitality by transforming academic insights into therapeutic solutions for patients and internationally tradable products.

There are 6,548 active businesses in the UK life sciences sector, employing 68,900 people and generating an annual turnover of £8.1 billion. 77% of these organisations are SMEs, standing at the forefront of ground-breaking research and innovation[1].

These companies play a pivotal role in crafting novel tools and technologies that have the potential to transform healthcare provision, enhancing health outcomes. However, due to their smaller size and constrained resources relative to larger pharmaceutical corporations, SMEs should warrant special attention from government and related bodies during the formulation of policies and the design of research infrastructure.

The global life sciences sector is increasingly reliant on health data access for an array of applications, ranging from drug discovery and safety testing to patient stratification and diagnostics development. However the accessibility of health data the process involved in its acquisition both have room for improvement.

The challenges SMEs face when accessing health data are numerous, given their more restricted resources to navigate intricate governance procedures. These pioneering firms necessitate targeted attention and policies from government and other data custodians to unlock the full economic potential of research on genomic and health data.

The BIA welcome Professor Sudlow's review on *Unifying Health Data in the UK* as a timely and critically important step in ensuring the vibrancy and vitality of UK-based genomics SMEs.

Current barriers to unifying health data in the UK

Fragmented datasets and organisations

Health data in the UK is largely fragmented across multiple organisations, most of which are sub-national. SMEs may not have the necessary resources for secure, interoperable and efficient data access. For example, a local NHS trust might have a wealth of patient data but lack the IT infrastructure or personnel to facilitate broader access to this data, rendering it under-utilised.

Government initiatives and their impact

[1] Bioscience and health technology sector statistics, Official Statistics, Feb 2022

Various initiatives by government, such as the HDR UK hubs and the Data for R&D programme inadvertently sustain this fragmented health data landscape. These initiatives often prioritise local, regional, or disease-specific data access, which leads to the continuation of siloed data repositories. It discourages a more unified approach, which could allow for more comprehensive research and insights.

Narrative around SDEs

The public and stakeholder discourse around health data access has largely focussed on security and pricing, overlooking other important aspects such as feasibility and researcher priorities. This has led to the creation of plans that may not meet the needs of the research community or align with the practical capabilities of the healthcare system.

Different formats and standards

Differences in the data formats and standards employed by different healthcare organisations further complicate data unification efforts. For instance, one organisation may store genetic data in a format that is incompatible with the systems used by another organisation. As health data can take various forms - genetic data, health records, biomarker data or imaging data - the lack of uniform standards presents a significant hurdle to effective data sharing and usage.

Lack of a common data infrastructure

At present, the UK lacks a common data infrastructure for health data, meaning there is no central repository where all health data is stored, or that data from various sources is incompatible. Health data is spread across various systems, leading to accessibility challenges, especially for SMEs which may not have the resources to navigate these disparate systems.

Data quality

The quality of health data is not consistent across different organisations, which can undermine its usefulness for research. While some organisations might have meticulously recorded and maintained data others may not have the same level of diligence. There is also a lack of clarity around what data is available and where.

Public and private health data discrepancies

Health data is stored by both public and private organisations, with differing priorities and incentives. . his dynamic creates a fractured landscape where different health data is held by different organisations, each operating under their own terms and with different access points, making a unified approach challenging.

UK BioIndustry Association

Potential solutions to unify health data in the UK

Government support should be moved from subscale SDEs to a single platform

The assumed value of UK data relies on it being unified and accessible. Initiatives are helpful if they support subnational data to be accessible in one secure and efficient platform, but not if they are building multiple datasets with multiple access procedures on multiple platforms.

SMEs cannot take the time and risk to try to rebuild national-scale datasets by applying to many SDEs. It is unlikely that all the individual regional SDEs will have the resources to be secure, efficient and enable cutting-edge research.

Interconnectivity - - a unified system with a single access point, standard data models, interoperability, and data search across all SDEs is recommended.

The value of data is enhanced considerably when it is supplemented with data from outside the original data source. Third parties should be able to add data and value to the SDE from data sources that are not traditionally captured by the NHS or social care environments.

This interconnectivity should include:

- A single point of access for all SDEs
- Standardisation of data, through using internationally recognised standard data models (e.g., OMOP, openEHR and Mauro)
- Standardisation of functionality to support a trusted framework of interoperability. For example, through the introduction of standard APIs like FHIR (Fast Healthcare Interoperability Resources)
- The ability to search or link data across SDEs though interoperability and the use of identifiers
- Ensuring that analyses can be performed across the whole of the UK population
- Ensuring that data can interconnect with international data sets

Standardised information and KPIs - clarity on data availability, uses, and limitations is essential, with standard KPIs for all health data organisations focusing on data and capabilities.

Organisations that hold data should make it easier for SMEs to know exactly what data is available. This includes:

- Being clear about data availability
- Being clear about what the data has been used for

- Being clear about what the data can be used for

There should be standard KPIs that all health data organisation use, which should focus on:

- Data – including volume, velocity, variety, variability, veracity, visualisation, and value
- Capabilities – including what features and functions are available
- Certifications – including what compliance and standards have been achieved
- Services – descriptions of what services are available including any platform downtime. For example, Genomics England displays this publicly online
- Standardised Service-Level Agreements (SLA)s in all SDEs - clear service expectations, decision-making transparency, standardised application processes, and agreements are needed
- Using service level agreements so that users are clear about what service to expect
- Communicating the decision-making process, ensuring applicants can see clear details of the basis for the decision and examples of why applications would be or have been refused
- Using a flow diagram to show the application process with timelines, decision points and decision-makers
- Using standardised agreements where data is not centralised. This would mean that where an applicant is establishing access with several organisations, e.g., NHS trusts, one application process can be used

Accreditation for all

Accreditation of users and organisations similarly to the authentication of data safe haven. This would show that an individual has had the appropriate training and experience to perform the analysis in any given environment. This would also facilitate streamlined access to multiple data collections.

Uniformity of technical ability of data environments and security

SDEs should be optimal for innovators in the life science sector, including SMEs. If an SDE limits how data can be used or analysed by design, it becomes a barrier to, rather than an enabler of innovation.

Access to SDEs should be under a clear SLA, which outlines the technical service provided.

The ability to support the deployment of innovative proprietary analytical pipelines (PAP) on flexible computational platforms should be available to all users.

Securing access to a public secure data environment or the ability to use a PAP should apply to all public secure data environments.

The BioIndustry Association (BIA) is the voice of the innovative life sciences and biotech industry, enabling and connecting the UK ecosystem so that businesses can start, grow and deliver world-changing innovation.

Our members include start-ups, biotechnology and innovative life science companies, large pharmaceutical companies, universities, research centres, tech transfer offices, incubators and accelerators, and a wide range of life science service providers: investors, lawyers, IP consultants, and IR agencies. We promote an ecosystem that enables innovative life science companies to start and grow successfully and sustainably.

For any further information on the contents of this submission please contact the BIA policy team at oroth@bioindustry.org