# UK Health Data Research Alliance

# Principles for Participation

**March 2023**

The *UK Health Data Research Alliance* is a member-led collaboration of organisations committed to improving human health by maximising the potential of multiple forms of data at scale.

## Overview

This document provides an overview of the principles that organisations and individuals must adhere to as part of their involvement in the UK Health Data Research Alliance (the Alliance) and associated programmes, including use of the Health Data Research Innovation Gateway.

The vision is to make the UK home to data-driven research, scientific advances and innovation in healthcare to improve patient outcomes. The UK has some of the richest healthcare datasets worldwide. However, NHS and health research data are not always accessible, and their potential uses for research and innovation are not being fully realised. The aim is to increase the access and use of health data in a trustworthy and ethical way in order to develop improvements in the UK's health technology and deliver benefits to patients and the population.

The principles for participation draw on national and international best practice frameworks, policy and recommendations. They are intended for all organisations involved in the Alliance (whether as a user of data, provider of Secure Data Environments, Health Data Research Hub or data custodian in the Alliance). They will guide working practices and may be reinforced through specific agreements through which organisations will engage with each other and with the Alliance programmes.

The principles were originally developed for the Innovate UK Industrial Strategy Challenge Fund Digital Innovation Hub Programme that formed part of the Data to Early Diagnosis and Precision Medicine Challenge. As that programme concluded (March 2023), the UK Health Data Research Alliance Board agreed to adopt the majority of the principles for the next phase of its development. They are set out below followed by a summary of the existing frameworks, best practices and recommendations that underpin the principles for participation.

## Principles for Participation

Every organisation involved in the Alliance and associated programmes commits to:

1. **Demonstrate active and ongoing engagement with patients and the public** in the design, development and governance of their activities involving health data to provide assurance that these activities are in the public interest.
2. **Encourage the availability and use** of structured and unstructured health and care data, including clinical, administrative, imaging, genomic and other molecular data, for research and innovation that serves **public interest purposes in line with guidance issued by National Data Guardian** [1].
3. Promote **the protection of privacy and data security** in line with the OECD Recommendation of the Council on Health Data Governance [2] through the adoption of best practices in privacy enhancing technologies, including providing access to sensitive data through Secure Data Environments (also referred to as Trusted Research Environments) accredited by a recognised authority in line with Digital Economy Act, DHSC Secure Data Environment Policy and Charter for Safe Havens in Scotland [3, 4, 5].
4. Make data Findable, Accessible, Interoperable and Reusable by adopting the **FAIR Guiding principles for scientific data management and stewardship** [6].
5. Use a **proportionate approach to the governance** of data access based on the **five "safes"** [7] and adhere to the **Foundation Principles and Core Elements for Responsible Data Sharing** [8] set out in the Global Alliance for Genomics and Health Framework for Responsible Sharing of Genomic and Health-Related Data.

6. **Maximise the benefits of data for research and innovation through non-preferential access to data** for uses that serve the public interest, by ensuring that data remains available to, and accessible by, any organisation (concurrently or otherwise) provided the five "safes" criteria are met and the organisation meets the access requirements of the data custodian.

7. **Establish mutually beneficial ways of working in partnership** including contractual arrangements and Intellectual Property agreements in line with principles set out in the Life Sciences Sector Deal 2 [9]. (and as updated by NHS England Centre for Improving Data Collaboration and equivalents in devolved nations).

8. **Work collaboratively to increase harmonisation** and reduce the complexity of data sharing arrangements and data governance models to improve the efficiency of accessing data for trustworthy and ethical research and innovation purposes. This includes making the terms of access clear, such as expected timescales and costs, and being transparent about the type and quality of data available.

9. Contribute to a **joined-up and UK-wide offer** for researchers in all sectors by collaborating with existing, relevant health research infrastructure, embracing open ways of working, code sharing and development of reproducible analytical pipelines in line with Goldacre Review [10].

# Summary of References

1. **What do we mean by public benefit? Evaluating public benefit when health and social care data is used for purposes beyond individual care,** a National Data Guardian guidance.
https://www.gov.uk/government/publications/what-do-we-mean-by-public-benefit-evaluating-public-benefit-when-health-and-adult-social-care-data-is-used-for-purposes-beyond-individual-care

This guidance will improve public benefit evaluations by defining and standardising the concept of public benefit to enable clearer interpretation and understanding.

2. *OECD, Recommendation of the Council on Health Data Governance*, OECD/LEGAL/0433
https://www.oecd.org/health/health-systems/health-data-governance.htm

3. **Digital Economy Act 2017**.

https://www.gov.uk/government/collections/digital-economy-bill-2016

4. **Secure Data Environment for NHS health and social care data – policy guidelines**.

https://www.gov.uk/government/publications/secure-data-environment-policy-guidelines/secure-data-environment-for-nhs-health-and-social-care-data-policy-guidelines

These guidelines provide additional information about the use of secure data environments as outlined in the Data Saves Lives strategy.

https://www.gov.uk/government/publications/data-saves-lives-reshaping-health-and-social-care-with-data/data-saves-lives-reshaping-health-and-social-care-with-data

5. **Charter for Safe Havens in Scotland**. Handling Unconsented Data from National Health Service Patient Records to Support Research Statistics.

https://www.gov.scot/publications/charter-safe-havens-scotland-handling-unconsented-data-national-health-service-patient-records-support-research-statistics/pages/

6. **The FAIR Guiding Principles for scientific data management and stewardship**. Wilkinson, M. D. et al. Sci. Data3:160018 doi: 10.1038/sdata.2016.18 (2016).

The FAIR Guiding Principles

To be Findable:

- F1. (meta)data are assigned a globally unique and persistent identifier
- F2. data are described with rich metadata (defined by R1 below)
- F3. metadata clearly and explicitly include the identifier of the data it describes
- F4. (meta)data are registered or indexed in a searchable resource

To be Accessible:

- A1. (meta)data are retrievable by their identifier using a standardized communications protocol
    - A1.1 the protocol is open, free, and universally implementable
    - A1.2 the protocol allows for an authentication and authorization procedure, where necessary
- A2. metadata are accessible, even when the data are no longer available

To be Interoperable:

- I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (meta)data use vocabularies that follow FAIR principles
- I3. (meta)data include qualified references to other (meta)data

To be Reusable:

- R1. meta(data) are richly described with a plurality of accurate and relevant attributes
    - R1.1. (meta)data are released with a clear and accessible data usage license
    - R1.2. (meta)data are associated with detailed provenance
    - R1.3. (meta)data meet domain-relevant community standards

7. **The Five Safes**. *Desai, Tanvi; Ritchie, Felix; Welpton, Richard (2016). "Five Safes: designing data access for research"*. Bristol Business School Working Papers in Economics

| | |
|---|---|
| Safe projects | Is this use of the data appropriate? |
| Safe people | Can the users be trusted to use it in an appropriate manner? |
| Safe settings | Does the access facility limit unauthorised use? |
| Safe data | Is there a disclosure risk in the data itself? |
| Safe outputs | Are the statistical results non-disclosive? |

8. **Foundation Principles and Core Elements for Responsible Data Sharing** set out in the Global Alliance for Genomics and Health Framework for Responsible Sharing of Genomic and Health-Related Data. https://www.ga4gh.org/genomic-data-toolkit/regulatory-ethics-toolkit/framework-for-responsible-sharing-of-genomic-and-health-related-data/#fp

Foundational Principles

- Respect Individuals, Families and Communities
- Advance Research and Scientific Knowledge
- Promote Health, Wellbeing and the Fair Distribution of Benefits
- Foster Trust, Integrity and Reciprocity

**Core Elements for Responsible Data Sharing**

It is good practice for those involved in genomic and health-related data sharing to have core elements of responsible data sharing in place. The following Core Elements of the Framework aid in the interpretation of the Foundational Principles to individuals and organizations involved in the sharing of genomic and health-related data. The Core Elements should be interpreted in a proportionate manner that acknowledges different levels of risk and community cultural practices. This Framework applies to use of data that have been consented to by donors (or their legal representatives) and/or approved for use by competent bodies or institutions in compliance with national and international laws, general ethical principles, and best practice standards that respect restrictions on downstream uses.

- Transparency
- Accountability
- Data Quality and Security
- Privacy, Data Protection and Confidentiality
- Risk-Benefit Analysis
- Recognition and Attribution
- Sustainability
- Education and Training
- Accessibility and Dissemination

9. **Creating the right framework to realise benefits for patients and the NHS where data underpins innovation – Guiding principles1 set out in the Life Sciences Sector Deal 2**

https://www.gov.uk/government/publications/creating-the-right-framework-to-realise-the-benefits-of-health-data/creating-the-right-framework-to-realise-the-benefits-for-patients-and-the-nhs-where-data-underpins-innovation

- Principle 1: Any use of NHS data, including operational data, not available in the public domain must have an explicit aim to improve the health, welfare and/or care of patients in the NHS, or the operation of the NHS. This may include the discovery of new treatments, diagnostics, and other scientific breakthroughs, as well as additional wider benefits. Where possible, the terms of any arrangements should include quantifiable and explicit benefits for patients which will be realised as part of the arrangement.
- Principle 2: NHS data is an important resource and NHS organisations entering into arrangements involving their data, individually or as a consortium, should ensure they agree fair terms for their organisation and for the NHS as a whole. In particular, the boards of NHS organisations should consider themselves ultimately responsible for ensuring that any arrangements entered into by their organisation

are fair, including recognising and safeguarding the value of the data that is shared and the resources which are generated as a result of the arrangement.

- Principle 3: Any arrangements agreed by NHS organisations should not undermine, inhibit or impact the ability of the NHS, at national level, to maximise the value or use of NHS data. NHS organisations should not enter into exclusive arrangements for raw data held by the NHS, nor include conditions limiting any benefits from being applied at a national level, nor undermine the wider NHS digital architecture, including the free flow of data within health and care, open standards and interoperability.
- Principle 4: Any arrangements agreed by NHS organisations should be transparent and clearly communicated in order to support public trust and confidence in the NHS and wider government data policies.
- Principle 5: Any arrangements agreed by NHS organisations should fully adhere to all applicable national level legal, regulatory, privacy and security obligations, including in respect of the National Data Guardian's Data Security Standards, the General Data Protection Regulation (GDPR) and the Common Law Duty of Confidentiality.

10. **Better, broader, safer: using health data for research and analysis**

https://www.gov.uk/government/publications/better-broader-safer-using-health-data-for-research-and-analysis

https://www.goldacrereview.org/