UK BioIndustry Association

BIA's vision for data 2035: 5 pillars to support research and innovation

BIA's vision for data 2035: 5 pillars to support research and innovation

A supplementary paper to accompany BIA's health data vision 2035: harnessing a national asset for growth

Contents

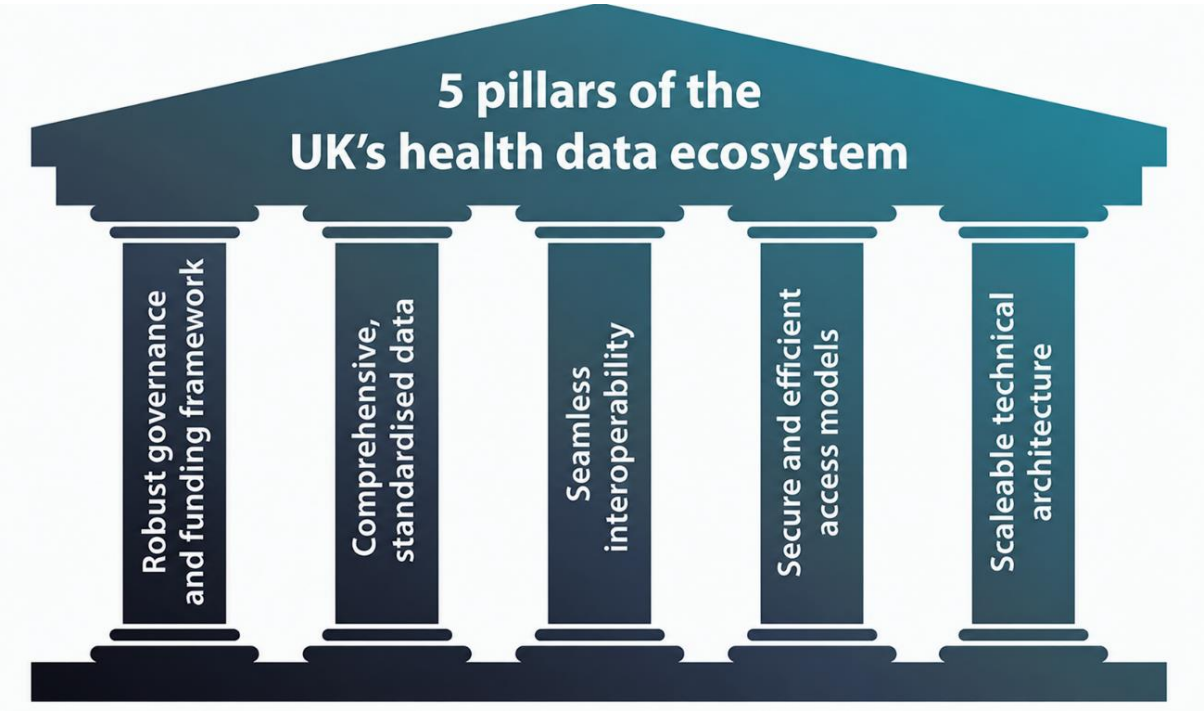
BIA's vision for data 2035: 5 pillars to support research and innovation

Strategic use of the UK's health data will be fundamental in realising the Government's NHS 10-year plan and the Life science sector plan. But in order to fully harness our health data for the benefit of patients and the public we need to keep pace with the evolving opportunity technology and life sciences can offer. This paper, as a supplement to the BIA's health data vision 2035: harnessing a national asset for growth, outlines how strategic Government investment over the coming years can build a world-leading, competitive health data ecosystem that fuels innovation, strengthens the UK economy, improves NHS productivity, and delivers better outcomes for patients.

The 5 pillars of a modern health data ecosystem

By 2035, the UK's health data ecosystem will function as a unified, intelligent platform powering both high-quality healthcare and cutting-edge life science innovation. This platform should be composed of five interdependent pillars:

1. **A secure and robust governance and funding framework**, including affordable access for UK SMEs, free from competing revenue targets, that ensures secure, legally compliant and ethical use, value return to the UK and industry co-development.
2. **Collection and standardisation of comprehensive data** including genomics: standardised across longitudinal multimodal sources.
3. **Seamless integration across systems and data types**, with up-to-date records and common standards across all four UK nations, supported by defined interoperability standards governance.
4. **Secure and efficient access models** that are streamlined and timely, offering different data and access models for different users and applications, underpinned by clear lawful bases for processing and transparent consent and opt-out frameworks.
5. **Resilient, flexible and scalable technical architecture** that can enable AI and other innovative clinical solutions.



Each pillar reinforces the others. It is fundamental that each of these pillars continues to be developed with patient and public involvement, user input, sustained investment, and clear measures of progress. By scaling these functions and supporting UK innovators to use UK health data responsibly, we can create real value for patients, the NHS, UK companies, and the wider economy, as outlined in the BIA health data vision 2035 report.

This paper outlines each pillar, the work that is needed, and example metrics to measure progress.

1. A secure and robust governing framework

Strong, trusted governance is essential for the future of the UK's health data ecosystem. By 2035, the UK has an opportunity to lead with a governance framework that is both ethical and enabling – designed to support secure data access while delivering benefits and value to the NHS, patients, and the UK economy.

Rather than being constrained by legacy models, governance will need to evolve to match the dynamism of modern research users. This means moving beyond static, one-size-fits-all controls toward proportionate, transparent, and risk-based oversight, with patient and public involvement at its centre. Crucially, it also means maintaining the integrity of patient records while enabling transformative uses of data to improve lives.

Principles for trusted and enabling health data governance

A future-ready governance framework will:

- Embed public trust, transparency, accountability, and reporting from the outset
- Operate proportionately to the risks and purposes of data use
- Support timely, responsible access for clinical and research needs
- Be co-designed and regularly reviewed with patients, clinicians, regulators, and industry users
- Provide safeguards and auditability without delaying innovation.

Public confidence through ethical stewardship

Public trust will continue to be paramount. The system will require transparent governance frameworks for data usage, strong ethical oversight, and mechanisms for public engagement and redress. These systems should be co-designed with the public, building on DHSC'S recent public deliberations[i].

Clear and consistent policies for data ownership, consent (with granular options), anonymisation/pseudonymisation, and data sharing agreements will be essential to ensure responsible innovation and maintain public confidence. This includes proactive strategies to identify and mitigate biases in data and algorithms that could perpetuate or worsen health inequalities.

Data output and return of results

To enable a broad range of research and innovation use cases, data environments should allow outputs to leave securely and efficiently, with upfront clarity on permitted result types. Key requirements include:

- Clear, pre-agreed rules on what types of data inputs and outputs are allowed
- Automated airlock systems for low-risk projects to avoid delays
- Specific support for use cases such as precision medicine sub-population analysis, AI model validation, and ancestry-based insights
- Governance mechanisms for data ingress/egress, including port connectivity and organisational accreditation
- Technical monitoring with human in the loop for high risk or new innovative projects.

Pro-innovation accreditation framework

A streamlined, centralised accreditation process for users will support safe, timely access to data across platforms. Models such as HDRUK's SOURCD and UCL's Data Safe Haven can underpin a national standard for:

- Verifying training and qualifications
- Streamlining multi-dataset access
- Auditing and revoking access when required.

Risk-based access approvals will allow rapid progression for lower-risk projects, freeing up capacity for more complex or novel studies. Protocols for recontacting research consented patients will be embedded from the outset.

Driving interoperability and reusability

A robust regulatory foundation will support an ecosystem where:

- All clinical data is interoperable and machine-readable
- Primary care and hospital systems integrate via modern APIs and formats
- Research data is not needlessly duplicated (e.g. WGS only once per patient)
- Proven tools and consent models can be reused across sites and studies.

This will require both technical standards and policy levers, such as incentives for adoption (e.g. procurement priorities, digital maturity funding etc.)

Access models that foster innovation

As outlined in the BIA health data vision 2035, the commercial access model should be carefully considered to ensure that UK innovators can flourish.

The economic potential of a subscription model appears relatively easy to calculate: a limited number of large pharma companies may be able to afford a 'gold' level of access, some biotech companies may be able to afford a 'silver' level, and SMEs may only be able to afford a more limited 'bronze' tier. However, the real value created by health data is not determined simply by the number of subscriptions sold. It depends on how many organisations can use the data productively, what access rights and technical capabilities are included, and whether commercial terms allow companies to protect background IP and generate downstream value.

If pricing is too high, or if access terms are too restrictive, the model risks concentrating meaningful use among only the largest companies while limiting SME participation to

lower-value access tiers. This would reduce innovation, weaken UK-based value creation, and undermine the long-term health and economic benefits that HDRS is intended to unlock. The value created by downstream innovation is therefore likely to significantly exceed the revenue captured through access fees alone.

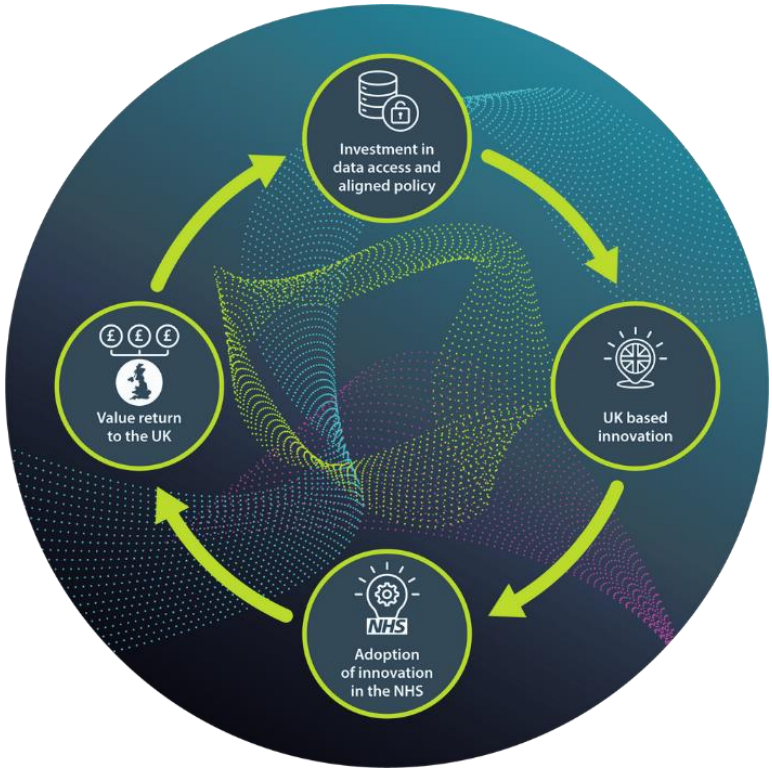
Pricing is also affected by data quality and global competition. Globally, similar resources are becoming significantly more numerous, accessible, and better tailored to specific research use cases. A first 'minimally viable' Health Data Research Service (HDRS) product based only on integrated births, deaths, care and prescription records, even at massive national scale, may have limited value for precision medicine, preventive medicine, and diagnostics unless it is linked to genomics, imaging, disease-specific or outcomes data. This means the first HDRS iterations are likely to be much less valuable for research users than later datasets, although those will have more competition by the time they come to market.

Current competitive datasets provide useful benchmarks for SME affordability:

Data set	Scale	Cost and terms	Accessibility
UK Biobank	500,000 participants	Approximately £9k per year plus TRE usage fees, with no downstream IP or commercial rights	Widely accepted
All of Us	1 million participants	Approximately $15k per year plus TRE usage fees, with no downstream IP or commercial rights	Widely accepted
Other major US health systems	250,000 – 1 million+ participants	Approximately $25k-$50k per year, often additional administrative and investigator costs, with no downstream IP/rights (inclusive of contribution to IRB fees and principal investigator time)	Less popular subscriptions than UKBiobank/All of Us due to the time, effort and cost of subscribing

For some resources (e.g. Our Future Health, UK Biobank Pharma Proteomics Project) larger subscription fees have been paid by pre-competitive pharma consortia in return for a 9-36 temporary exclusivity. Such exclusivity-based model would be difficult to justify for wider NHS datasets and could risk limiting UK-based innovation and value creation.

More affordable and predictable access to data for approved research will enable more innovation, better health outcomes, and greater value creation for the NHS, patients and the UK economy.



UK industry leadership and co-development

The future data ecosystem will require extensive user engagement. Deep involvement of users will be essential, including through a formal Industry Advisory Group – a key recommendation of the Sudlow Review[ii].

A robust Industry Advisory Group should reflect the full breadth of the life sciences sector, with a specific focus on SMEs. SMEs are the primary engines of innovation within the UK life sciences ecosystem; supporting these companies in bringing commercially successful products to market is a critical for both economic growth and improved patient outcomes. This forum should:

- Represent the full life sciences value chain
- Provide early input on regulatory, technical, and operational design
- Identify opportunities to align public and commercial goals
- Ensure that environments work for UK startups and global-scale companies alike.

Early, meaningful, and accessible engagement with UK small and scaling businesses will help ensure the data ecosystem fosters innovation rather than creating new barriers. These users bring a wealth of practical experience gained from global collaborations regarding the technical challenges, regulatory processes, and significant investment required to transform raw health data into viable clinical products.

UK leadership in 2035: HDRS as a sovereign resource

To realise the benefits outlined in the BIA health data vision 2035, HDRS should be positioned as the UK's strategic gateway for health data. To fulfil this role, it will require:

- **Statutory authority** over data access and contractual terms
- **Clear mandate** to prioritise UK value creation alongside research excellence
- **Operational capacity** to assess UK value creation
- **Governance integration** across MHRA, NICE, DHSC, DBT, Treasury
- **Enforcement mechanisms** for value covenant compliance.

Access to UK health data and NHS collaboration should be **conditional on demonstrable UK value creation**, including, for example one of the following:

- Meaningful UK-based R&D investment
- Development activity and high-value jobs located domestically
- Clinical trials and co-development partnerships with UK institutions
- Development of solutions that can create value for the NHS.

This is not about protectionism. It is about stewardship: ensuring sustainable value for the UK public from the responsible use of personal health data. By ensuring that data access creates UK value, the country can sustain public support, fund future services, and remain a magnet for responsible global collaboration.

Example progress metrics:

- *Number of opt-outs and consents for research*
- *Time from analysis to return of results*
- *Time to user accreditation*
- *Number of UK SMEs on advisory groups*
- *Number of completed projects*
- *Ratio of SME applications to SME completed projects*
- *Mechanism for tracking UK value return*
- *Number of HDRS-enabled innovations adopted or evaluated by the NHS*
- *Investment and jobs created through HDRS-enabled projects.*

2. Comprehensive standardised data

The 2035 data ecosystem will be primarily defined by the quality, depth, breadth, and usability of the data it can surface. Comprehensive, longitudinal multimodal data will be fundamental to health innovation. The depth, complexity, and diversity of healthcare data available at scale will expand significantly, building on major initiatives such as Our Future Health. To deliver translational impact, the life science and healthcare sectors will require longitudinal multimodal data that goes far beyond traditional electronic health records. Large-scale integrated clinical and genomic data is already the established global standard for research-relevant patient data collections, including Our Future Health[iii], All of Us[iv], FinnGen[v], Kaiser Permanente Research Database[vi], the Veteran Association's Million Veteran Program[vii], PRECISE[viii], and the Emirati Genome Program[ix].

Genomic data: Genomic data is a key enabler of precision and preventative medicine. Over time, patients will become familiar with understanding their genetic risk through user-friendly personalised education and wellness programmes via apps and digital health avatars. This shift will embed genomics into routine care and self-management, empowering individuals to understand and act on their inherited health risks to maintain their health for longer.

Whole Genome Sequencing (WGS) will be routine for newborns, individuals at risk of inherited conditions, and cancer patients. WGS data is the gold standard for clinical care and research, but it remains expensive to collect and maintain, creates ethical and logistical challenges, and requires specialist interpretation. It is also not required for many diagnostic, prognostic and public health applications. In these cases, genotype data is easier to collect and manage at large population scale. To realise its full value, robust infrastructure will be needed to enable secure, long-term storage of a range of genetic data with a flexibility to support future applications.

Clinical data: Deep, longitudinal electronic health records (EHRs) including primary, secondary and tertiary care, hospital records, prescription and dispensing information, specialist consultations, outcomes data, and detailed phenotyping. This data must be harmonised, up to date, highly structured, and interoperable.

Functional genomics: By 2035, a comprehensive, causal atlas of genetic variants—including coding and non-coding regions—mapped across cell types, physiological conditions, and disease-relevant models will provide a fundamental shift in our ability to translate genomic information into precision health outcomes. This will enable diagnosis of more diseases with confidence, accelerating and de-risking of drug discovery, and new personalised treatments. This requires integration of genomics data with cell type specific

epigenetics, 3D genomics, and longitudinal transcriptomics, proteomics, and immunophenotyping data[x].

Imaging data: High-resolution medical images (MRI, CT, X-ray, pathology slides) combined with advanced image analysis and feature detection using computer vision AI, offering automated detection, quantification, and predictive insights.

Wearable and sensor data: Continuous streams of physiological data from wearables (heart rate, sleep patterns, activity levels), smart implants, and home monitoring devices, enabling early detection of deviations from baseline health and personalised interventions.

Environmental and lifestyle data: Data on diet, exercise, pollution exposure, socioeconomics/deprivation, social determinants of health, and other geographical factors, integrated to provide a holistic view of an individual's health influences.

Real-world evidence (RWE): Large-scale, real-world data from diverse patient populations, capturing outcomes and experiences outside of traditional clinical trials, to inform drug development, post-market surveillance, and health policy.

Unstructured data: Advanced Natural Language Processing (NLP) and Large Language Models (LLMs) will extract valuable insights from free-text clinical notes, research papers, and patient narratives, converting them into structured, analysable formats.

Synthetic and control data: Synthetic data can provide a 'staging area' that derisks earlier access for feasibility and study design, speeding up research including clinical trials. High-fidelity synthetic data, that mimics real-world data distributions and relationships without compromising patient privacy, will be fundamental for research and AI model training. Similarly, construction of control datasets from within patient cohorts that have been treated and characterised in the same way is essential.

Health equity must be embedded as a core principle. This means addressing historical underrepresentation and bias in datasets, avoiding algorithmic biases that could lead to differential care or misdiagnosis, and ensuring that digital-first approaches do not inadvertently widen existing health disparities.

Depth of data should be considered alongside breadth during the rollout of HDRS. Innovators will need to access deep, linked multimodal data sets as – connected to local communities and clinical teams - well as national-scale clinical data sets.

Example progress metrics:

- *Number of WGS and genotype records available*
- *Number and scale of linked multimodal datasets*

- *Linked imaging, pathology and outcomes data available*
- *Synthetic data available*
- *Data quality measures: completeness, timeliness, standardisation and linkage quality.*

3. Seamless integration

Seamless, secure access to diverse, high-quality datasets from multiple sources, and shared across institutions and borders will be essential in attracting and supporting innovation. Our vision is for HDRS to provide a single, national-scale resource of high-integrity, real-time health data. From a user perspective, this will function as a unified source—accessible via trusted hubs within a secure environment—capable of supporting advanced life sciences research.

Delivering this ambition will require both the standardisation and integration of the source data, alongside clear governance structures for managing the clinical data sources. Well before 2035, the UK should have established a fully interoperable data infrastructure enabling secure, near-real-time access to clinical and genomic data across the boundaries of healthcare providers, research institutions, and all four UK nations.

This will depend on the consistent adoption of common data standards (e.g., FHIR), interoperable APIs, and the widespread use of identifiers, such as the NHS number. Linking data from multiple sources at both individual and population levels will enable real-time analytics, discovery science, and more personalised, predictive and preventive care.

Key components of seamless data integration in 2035 will include:

- **Harmonised data access process:** with rapid, predictable and transparent approval for data access
- **The single patient record:** every patient will have a single, secure, and comprehensive lifelong digital record that they can access and control, supporting coordinated and personalised care that includes data from primary and secondary care, genomics, social care, environmental exposures, and lifestyle factors.
- **Deidentification of patient records for research applications:** best practice processes to redact patient-identifiable information, while preserving the ability to recontact patients where a legal basis to do so is present.
- **Data and cyber literacy training for patients**: patients with access to their own, healthcare data will need clear support to understand what it is appropriate to share, with whom and for what purposes

- **Data and cyber literacy training for clinical staff**: A healthcare workforce will need enhanced digital, data science and AI literacy. This includes understanding both the uses and limitations of such technologies rather than assumptions.

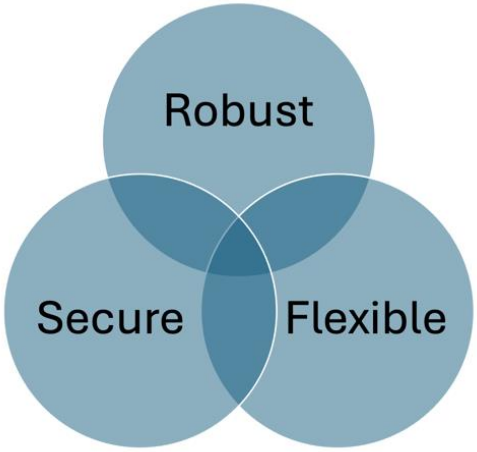
Example progress metrics:

- *Percentage of datasets meeting agreed standards.*
- *Percentage of records using unique identifiers.*
- *Time to data access approval.*
- *Percentage of access requests completed within agreed timelines.*
- *Time to de-identification and recontact.*
- *Number of staff and patients receiving data/cyber literacy training.*

4. Secure and efficient access models

By 2035, the UK will have a modern, flexible and high-performance technical architecture capable of handling large-scale, multimodal health data. This infrastructure will support both care delivery and world-leading research, while reflecting the sophistication of real-world data environments - from genomics to imaging, wearables, and unstructured health records.

Access models must reflect the fact that clinical care and research have different requirements. Clinical care requires identifiable data to support direct patient management. Research and innovation typically require de-identified data, secure analysis environments, and clear governance.



The access models supporting these applications and users must be:

- **Robust** – supporting high-volume, high-stakes clinical and research use.
- **Secure** – protecting privacy, intellectual property, and trustworthiness.
- **Flexible** – enabling safe, efficient access across a range of use cases.

The two main user groups are clinical care staff and research teams. These require different types of data, application support, and access mechanisms:

- **Identifiable data for clinical care:** By 2035, care staff will need to access identifiable, integrated health and care data to deliver high-quality clinical services. Data infrastructure will need to:
    - be secure and well-resourced, with built-in real-time links to care teams
    - support real-time analytics and proactive patient management
    - align with UK regulatory frameworks for access, sharing and storage
    - overcome local administrative and regulatory barriers through consistent national rules
- **De-identified data for research:** The core infrastructure supporting research users must provide secure, de-identified long-term storage for up-to-date healthcare data. Trusted research environments (TREs) are the backbone of a secure research ecosystem, providing secure access to de-identified data for research. To maximise their value, these environments will need to:
    - support long-term, secure storage and linkage across primary, secondary, and prescribing datasets
    - include high-quality data as out lined above
    - enable routine inclusion of clinical outcomes and drug response data

Critically, TREs will need to evolve beyond basic analytical functionality and support the secure deployment of proprietary analytic and AI tools and the import and co-analysis of third-party datasets. The next generation of discovery will require:

- support for AI-driven analytics and large-scale multimodal data
- the ability to import and deploy validated tools and models using containerised systems (e.g., Docker)
- the ability to import and co-analyse a wide variety of research datasets
- a research environment that is fast, flexible, developer-ready, and not restricted to legacy methods

Example progress metrics:

- *Time from approved project application to data provision.*
- *Percentage of projects completed within agreed timelines.*
- *Percentage of users/projects meeting access requirements on first submission.*
- *TRE capability to store and analyse large-scale multimodal data.*
- *Time to approve and deploy proprietary tools.*

- *Turnaround time for cTRE accreditation.*
- *User satisfaction among SMEs, academia, and industry.*

5. Resilient, flexible and scalable technical architecture

As the volume and variety of health data and applications grows, so too does the infrastructure challenge. The platform will need to support demanding analytics and AI, during both research and deployment phases, flexibly and securely.

A resilient, future-proof data system will need to support diverse data access and integration models, analytical methods, high-throughput AI workflows, and collaborative projects across multiple sectors and geographies. Research almost always involves bringing together multiple datasets and applying new methods that are not part of standard workflows or libraries.

Innovators will need to have access to high-performance elastic compute power and to be able to securely bring their proprietary analytical and AI tools to the data – whether within their own environments or securely via externally hosted Federated Analysis or Certified Trusted Research Environments (TREs) (see box below).

Secure, scalable infrastructure enables analysis wherever data resides, with clear governance to protect value and trust.

Flexibility in data environment is key for R&D.

Companies rely on a wealth of data from various sources, across the UK and internationally. The ability to use these data together is vital to developing tools and treatments that are representative of different populations and care systems.

However, data-sharing agreements and legal constraints often prevent this data from being analysed out of territory or moved into centralised platforms. Future infrastructure will therefore need to allow secure analysis to occur wherever the data resides—whether that is in federated public environments or within cTRE platforms.

Trusted research environments

What is a TRE?

TREs are secure cloud computing environments where approved researchers can access and work with (typically de-identified) health data within a protected environment, without allowing those data to be copied or downloaded outside of the TRE. TREs have strict controls that limit access to data, minimise the risk that an individual person can be identified, and enable monitoring of the actions of researchers to ensure that data access and analysis rules are followed. They are designed to make sure that health data is always being accessed securely, used appropriately, and that people's privacy is being protected.

A TRE is a closed platform with an airlock that acts as a barrier between the TRE and the external world to control data import and export processes to prevent unauthorised transfers, protect participant privacy, and maintain system integrity. The airlock typically inspects all imports for malware, embedded identifiers, or security risks before allowing entry. Results to be outputted from the TRE (usually just group level or population summary results) will be checked by a "disclosure control" team to confirm that ethical standards have been met, that the data to be exported is appropriate and required for the purpose of the study, and that no re-identifiable data is present.

What is an accredited or certified TRE?

While many well-known TREs are hosted by the data custodians (for example the NHS, Our Future Health, or a cohort study), in some cases the data custodian can accredit or certify externally owned TREs. These certified TREs (cTREs) are similar cloud computing environments to the primary TRE but designed and operated by a third-party (a TRE provider or the end-user themselves). cTREs can be designed to be more flexible and secure than standard TREs, but they must meet or exceed the same strict data security and operational standards as the primary TRE. The custodian will typically use a specialist third-party company to audit adherence to these standards before allowing datasets to be shared and research studies to be run in a cTREs.

Core infrastructure requirements include:

- **Elastic compute capacity,** including CPUs, GPUs, and TPUs, to scale with varying project needs—from intense AI training to lighter-touch validation tasks
- **A flexible multi-provider, cloud-first strategy,** ensuring resilience, scalability, and cost-effectiveness

- **Open, developer-friendly environments** (e.g. Linux) with well-maintained hardware and software support that allow users to install, test, and update tools and libraries directly
- **Support for integration of third-party data sources**
- **Support for federated analysis** where data cannot be physically moved due to agreements or sensitivity
- **High-performance, independently audited platforms** enabling secure, logged access to sensitive datasets.

Security and IP protection will be built in by design:

- Regular OS updates and security patch maintenance
- Independent cybersecurity audits and adherence to ISO27001
- Regular independent penetration testing and reporting
- Strong IP governance, including contractual protections for proprietary assets (where innovators need to transfer their IP into third-party environments)
- Auditable access logs to verify that only authorised contributors can/have access/ed data or environments
- Service-level guarantees for service, compute and data availability

Enabling secure, performant access across this data ecosystem is ***vital for keeping the UK competitive*** as a destination for data-driven discovery.

Example progress metrics:

- *Time to compute provision*
- *Percentage of projects supported by adequate compute provision*
- *Ability to link or federate external data sets*
- *Time to agree IP protection terms*
- *Availability of auditable access*
- *Time to agree service-level guarantees.*

Summary conclusion

HDRS is a critical opportunity to build the national infrastructure needed for the next generation of healthcare and life sciences innovation. Done well, it can help the UK move from fragmented data assets to a coordinated ecosystem that improves patient outcomes, strengthens NHS sustainability, supports UK SMEs, and drives economic growth.

The five pillars set out in this paper — comprehensive data, seamless integration, trusted access, scalable architecture, and enabling governance — are mutually reinforcing. None will succeed in isolation. Together, they provide the foundation for a health data ecosystem that supports research, innovation, clinical improvement, and value return to the UK.

The UK already has many of the assets needed to lead globally: the NHS, world-class research infrastructure, genomics capability, trusted regulation, and a vibrant life sciences ecosystem. But leadership is not guaranteed. The next five years will determine whether HDRS becomes a strategic national platform for innovation and growth, or simply another access mechanism in a fragmented landscape.

The opportunity is clear. HDRS should be built as a national asset that enables responsible innovation, supports UK companies, delivers value back to patients and the NHS, and secures the UK's position as a global leader in health data-driven life sciences.

References

[i] NHS England National public engagement on the use of health and care data (Oct 2025) https://transform.england.nhs.uk/key-tools-and-info/data-saves-lives/national-public-engagement-on-the-use-of-health-data/

[ii] HDRUK Uniting the UK's Health Data: A Huge Opportunity for Society https://www.hdruk.ac.uk/helping-with-health-data/the-sudlow-review/ (2024)

[iii] Cook MB, Sanderson SC, et al. Our Future Health: a unique global resource for discovery and translational research. *Nat Med.* 2025 Mar;31(3):728-730. doi: 10.1038/s41591-024-03438-0.

[iv] The All of Us Research Program Genomics Investigators. Genomic data in the All of Us Research Program. *Nature* 627, 340–346 (2024). https://doi.org/10.1038/s41586-023-06957-x

[v] Kurki M.I., et al. FinnGen provides genetic insights from a well-phenotyped isolated population. *Nature* 2023 Jan;613(7944):508-518. doi: 10.1038/s41586-022-05473-8.

[vi] https://researchbank.kaiserpermanente.org/for-researchers/

[vii] Justice AC, McMahon B, Madduri R, et al. A landmark federal interagency collaboration to promote data science in health care: Million Veteran Program-Computational Health Analytics for Medical Precision to Improve Outcomes Now. *JAMIA Open*. 2024;7(4):ooae126. doi: 10.1093/jamiaopen/ooae126.

[viii] https://www.biorxiv.org/content/10.1101/2025.03.13.642552v1

[ix] https://www.medrxiv.org/content/10.1101/2025.09.12.25334546v1

[x] BIA, Functional genomics explainer: A strategic opportunity for the UK (2025) https://www.bioindustry.org/resource-report/functional-genomics-explainer.html