Options appraisal to deliver a health data research service in England

January 2025

HDRUK
Health Data Research UK

abpi

Foreword

With our diverse population and universal healthcare system, NHS data should be the UK's jewel in the crown, not only to enable the highest quality of person-centred care, but also to drive research and development (R&D) and innovation to improve population health and attract inward investment. Despite this potential, the existing reality does not match the rhetoric and falls short of expectations.

Our nation boasts internationally acclaimed data assets, such as UK Biobank and the Clinical Practice Research Datalink (CPRD), which have long delivered high-quality data services to academia, government, and industry worldwide. Yet, as highlighted in the Sudlow Review 'Uniting the UK's Health Data', accessing the majority of NHS data for R&D is fraught with challenges, resulting in missed opportunities for patients, the NHS, researchers, and the UK economy.

In response to the Sudlow Review's recommendation, the government has signalled an openness to establishing a health data research service. This initiative presents a crucial opportunity to transcend our current complicated and fragmented landscape to create an internationally competitive research service. This service should be grounded on the principles of user-centred design and patient engagement, genuinely addressing the needs of all stakeholders.

The task is substantial, and it is tempting to settle for optimising our existing investments, despite their known flaws. To support the government in this ambitious endeavour, we have considered three principles for the design of a national data service. Firstly, prioritising the most critical datasets for R&D to provide a strong foundation for a data service that can then be scaled and expanded. Secondly, defining the essential criteria for a sustainable health data research service, setting a clear framework for future developments. Lastly, leveraging and building on the expertise and experience of existing data services that already deliver to an international standard.

These principles have guided our evaluation of seven potential health data research service options, enabling us to offer informed recommendations for government consideration. We are poised at the threshold of potentially creating something that is not just national but truly world-leading in health data research. Our insights aim to provide a robust foundation to assist the government's decision-making, ensuring that NHS data fully realises its potential to catalyse health improvement and stimulate economic growth.

Dr Richard Torbett MBE
Chief Executive –
Association of the British
Pharmaceutical Industry

Professor Andrew Morris CBE
Director – Health Data
Research UK

What this work adds

What this work does not seek to do

Currently there are two published documents proposing the need for a centralised health data service supporting research uses via a **National Data Trust** or a **National Health Data Service**.

This work builds on published and unpublished discussions by:

- **Distilling the minimum national datasets that should form the basis of a health data research service** and can be expanded from this foundation. This approach has the greatest potential to demonstrate immediate value and avoids the danger of trying to 'boil the ocean' at the outset, whilse enabling the ambition to scale and link datasets more widely.

- **Defining the essential criteria for a viable, internationally competitive and sustainable health data research service**, which can serve as an evaluation tool for any data service model under consideration.

- **Appraising different health data research service model options**, outlining the pros and cons of each option. The aim is to demonstrate to policy makers that there are multiple options that can be explored to take this work forward, some of which are already efficiently delivering to international standards.

1 This work outlines seven health data research service model options and outlines pros and cons of each, It does not make a recommendation for which model should be adopted.

This is because:

- **To the best of our knowledge, there is no model that meets all the essential criteria for a service**. There are different models that are delivering internationally recognised data services and models where the challenges to overcome to achieve the ambition are substantial.

- **The decision is for policy makers** and our intention is to inform the debate.

2 While we drew on significant expertise and experience in the development of this work, we do not claim to represent the views of patients and the public, healthcare professionals and the full range of research data users. This is why we recommend that a wider discussion informed by this work and the work of others should take place.

Intended audience for this work: Policy makers, government officials, think tanks and experts in health data research policy across charity, government, industry and academia

Use of health data for R&D is an enabler of two government missions

Labour's five missions to rebuild Britain

1 **Kickstart economic growth** to secure the highest sustained growth in the G7 – with good jobs and productivity growth in every part of the country making everyone, not just a few, better off.

2 **Make Britain a clean energy superpower** to cut bills, create jobs and deliver security with cheaper, zero-carbon electricity by 2030, accelerating to net zero.

3 **Take back our streets** by halving serious violent crime and raising confidence in the police and criminal justice system to its highest levels.

4 **Break down barriers to opportunity** by reforming our childcare and education systems, to make sure there is no class ceiling on the ambitions of young people in Britain.

5 **Build an NHS fit for the future** that is there when people need it; with fewer lives lost to the biggest killers; in a fairer Britain, where everyone lives well for longer.

"Businesses in the life sciences industry generated £108.1 billion turn over in 2021/22. There were 304,200 people employed in the UK life sciences sector in 2021/22."

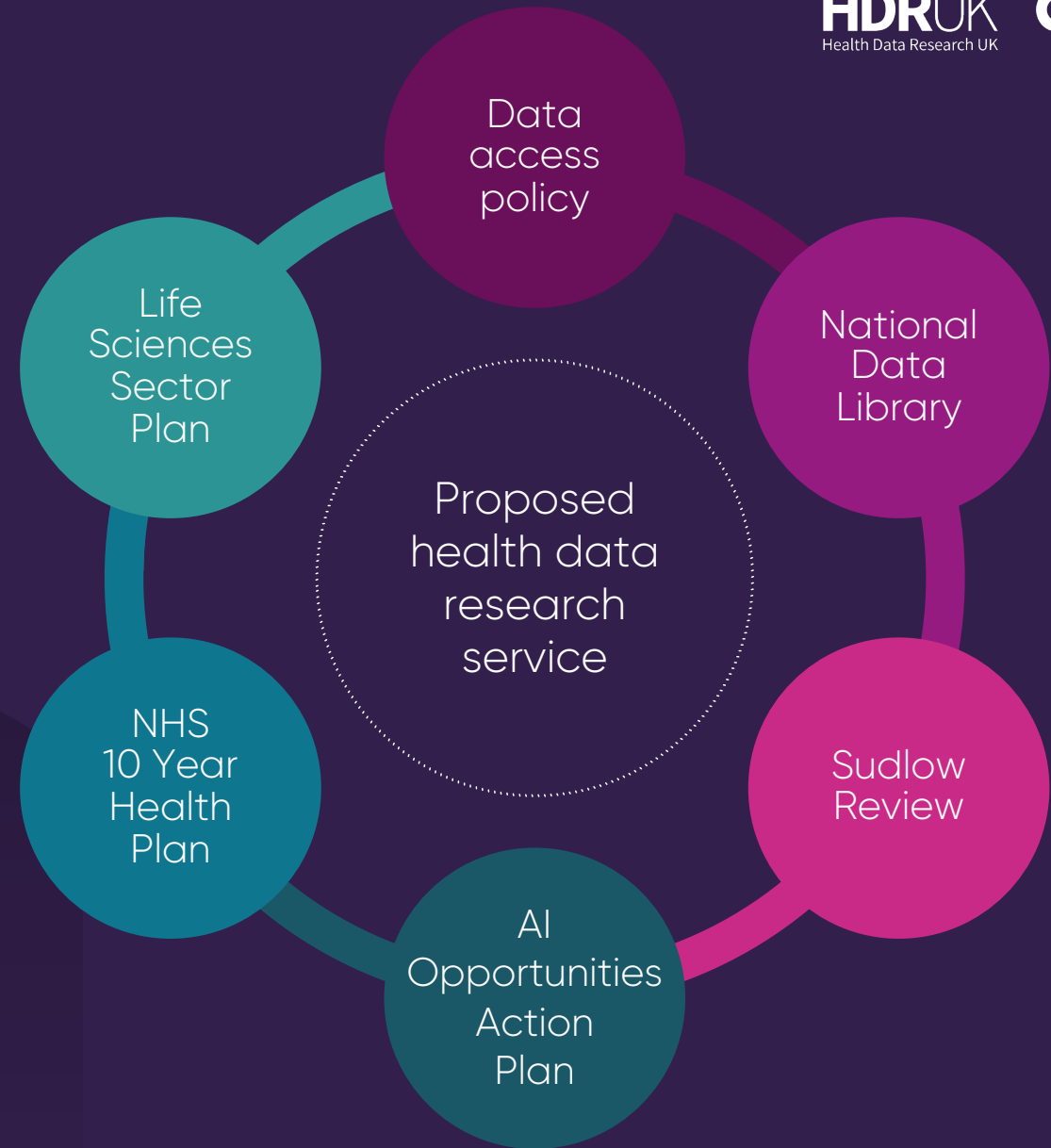Biosciences and health technology sector statistics 2021-2022, May 2024

"A national health data service will accelerate research and analysis that benefits society."

Sudlow Review: Uniting the UK's Health Data: A Huge Opportunity for Society, November 2024

Current health data policy context in England

It is vital we seize this opportunity to ensure:

1 **a shared vision** of what we want to achieve and a realistic implementation plan to achieve it

2 **sufficient investment** in underpinning health data research infrastructure

3 **user-centred design** to create an internationally competitive offer

4 **the maximising of our assets and expertise** by joining up our fragmented data ecosystem

Data access policy

Life Sciences Sector Plan

National Data Library

Proposed health data research service

NHS 10 Year Health Plan

Sudlow Review

AI Opportunities Action Plan

HDRUK
Health Data Research UK

abpi

6

Reality of the national health data research ecosystem in England

Despite the promise of NHS data being a UK unique selling point to power research and innovation, it remains largely untapped because in England there is:

inefficient government funding of national data infrastructure

- no centralised comprehensive GP data collection
- services fragmented across NHS trusts, NHS England & MHRA (CPRD)
- variable approaches between NHSE SDE and regional SDEs
- disparity in funding mechanisms and commercial models

lack of strategic join up across government

- each department operates in a silo, no system-wide planning leading to duplication, gaps and confusion
- limited stakeholder engagement and understanding of researcher data and service requirements
- pharmaceutical industry and international regulators use CPRD (~50% users) and few use current NHSE services (<10% users), which have received most recent investment

NHSE mergers and internal constraints

- disruption due to multiple mergers
- deprioritisation of research services when other NHS issues arise
- NHSE capacity restrictions due to limited headcount and skills causes long delays accessing data
- NHSD/E unable to link data for CPRD between 2021-2024, leading to major issues for regulatory uses

Joint ABPI HDR UK project to scope options for a sustainable health data research service

Current data services in England fragmented and not adequately delivering academic or industry research needs

This project takes forward **Sudlow Review** recommendation for a national health data service

Focus on creating a **minimum viable product health data research service** that can be scaled over time

Create a service that delivers a **trustworthy health data research service in England** that meets researcher needs

Purpose

Consider options to create a dedicated, internationally competitive, fit-for-purpose **national health data research service in England** that will:

- realise the value of NHS data for patients, the NHS and economic growth

- be trusted by patients, the public and healthcare professionals

- provide data and services that can be scaled over time and linked to other data assets and government data services

- inform future government funding decisions

Project approach to developing essential criteria and service model options

Determine priorities and essential criteria for a health data research service

- consult with experts from NHS, government, charities, academia and industry
- validate which datasets should be the initial priority
- refine essential criteria for a national research service that meets needs of stakeholders and users based on expert feedback‡

Develop service model options
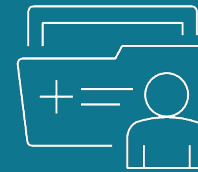
- gather evidence on existing data services' performance and delivery, financial sustainability and research user base
- in consultation with experts, identify viable service model options for a national research service

Evaluate service model options

- consider pros and cons of each option based on essential criteria for a service
- assess service model options, drawing on evidence of existing sustainable data services and options evaluation
- propose recommendations for next steps

‡This report is the work of the ABPI and HDR UK. An expert advisory group with representatives from NHSE, DHSC, HRA, AMRC, and general practitioners were asked to comment on ABPI and HDR UK assumptions on potential service model options, essential criteria for the service, and the accuracy of pros and cons proposed for each model. The role of these organisations was advisory on these aspects of the exercise only and they were not asked to endorse this report, or any recommendations regarding the potential strategic direction for a health data research service.

What should we prioritise to maximise the value of NHS data for patients, the NHS and the economy?

Value of NHS data lies in **longitudinal coverage** across diverse population

These data are valuable for discovery and identifying unmet need, population health, patient pathways, real world analysis and regulatory uses, which **can be generalised to the UK and other countries**

Timely access to national GP and secondary care data is critical

National GP and secondary care data collections encompass the population and contain a wealth of longitudinal information about symptoms, diagnosis, tests, interventions, prescriptions and vaccinations

Linkage to these national datasets is essential to maximise the value of 'omic' and imaging data, clinical and population cohorts and rich hospital datasets within regional SDEs

The priority therefore needs to be addressing this gap and creating a research service, starting initially with national health datasets, that can be scaled to link to other data sources

What should be the initial scope for a minimum viable national health data research service?

National datasets in England

- comprehensive GP data

- national secondary care datasets including hospital prescribing data, mortality data, audits and disease registries

Research purposes only

- access by academics, regulators, and industry researchers

- Secure Data Environment/Safe Haven by default

Research service offering

- linked data for observational research, AI and machine learning

- data linkage for participants in con`sented cohorts and clinical trials

Scalable

- initial focus on national health datasets

- aim to link with other UK national, regional,'omic' and non-health datasets including ONS and national data library in the future

- future services to support clinical trials feasibility and recruitment

Future scope for data linkage or inclusion: regional SDE hospital level data, 'omics' data, community, mental health and social care data, data in devolved nations, socioeconomic data.

Essential criteria for a health data research service

Operate in partnership with the NHS

- comprehensive longitudinal GP and national secondary care data collected by NHSE flows into service
- leverages operational efficiencies and reduces dependencies on GP software suppliers

Considered trustworthy by patients, the public and healthcare professionals

- Secure Data Environment/Safe Haven by default
- responsible for data controllership, with the ability to hold identifiable patient data

Deliver internationally competitive research services

- autonomy to set priorities and focus on research service delivery for all sectors
- workforce capacity and skills capability to meet user demands and support service innovation

Remove fragmentation in the research data ecosystem

- maximise government funding by combining government assets into a single fit-for-purpose service
- build on existing expertise and services that already deliver to international standards
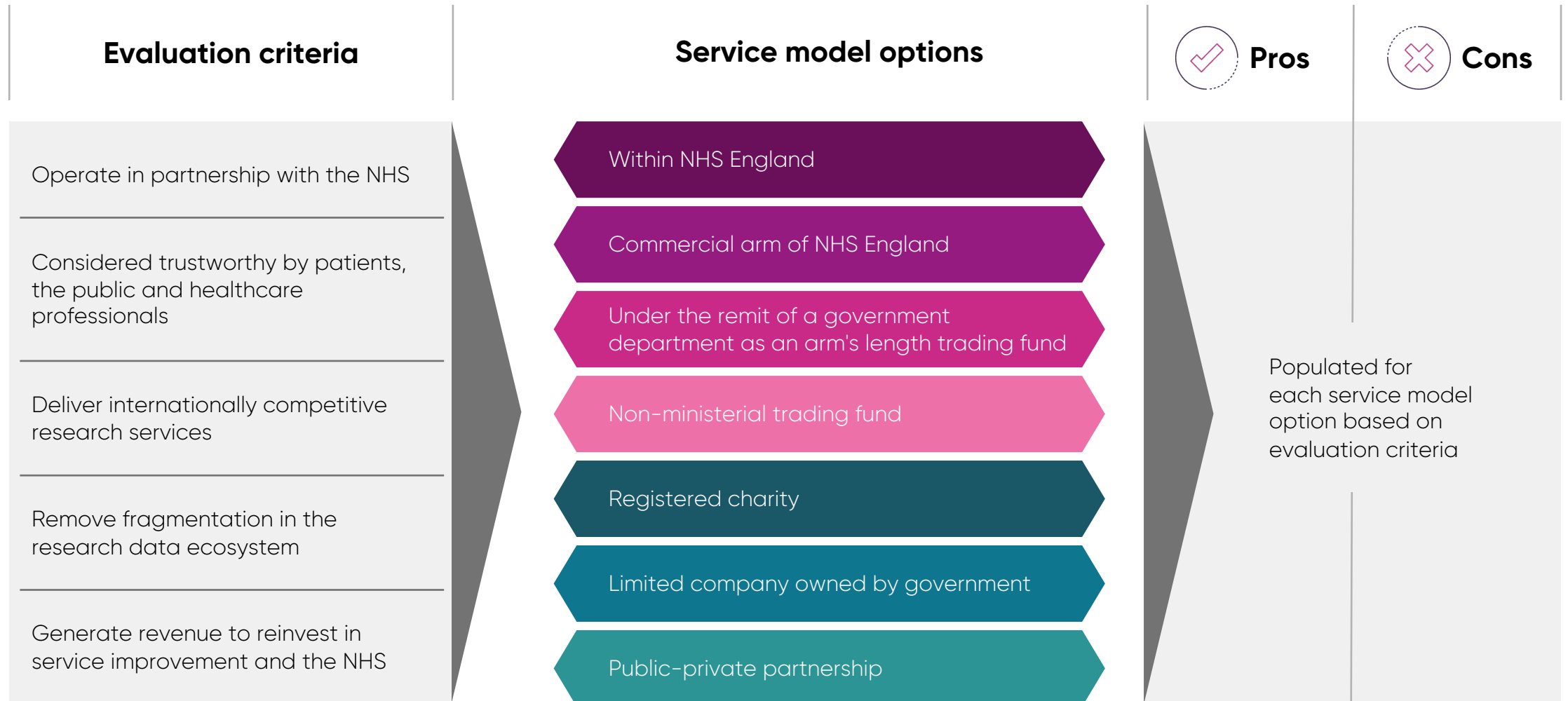
Generate revenue to reinvest in service improvement and the NHS

- dedicated budget, ability to generate and retain income to reinvest in service expansion and the NHS
- government funding required to establish the service, with goal of achieving financial sustainability

Options for health data research service model

Within NHS England
- Expansion of NHSE SDE
- Includes the research component of DARS

Commercial arm of NHS England
- Akin to NHS commissioning support units
- Body within NHSE with delegated powers and freedoms

Under the remit of a government department as an arm's length trading fund
- Akin to MHRA as arm's length body of department of health
- Previous status as trading fund

Non-ministerial trading fund
- Akin to Office for National Statistics
- Independent of direct ministerial oversight

Registered charity
- Akin to UK Biobank
- Majority government funding

Limited company owned by government
- Akin to Genomics England
- Wholly government owned company

Public-private partnership
- Company jointly owned by government and private sector

Evaluation criteria informed the assessment of each service model

HDRUK
Health Data Research UK

abpi

Evaluation criteria

Operate in partnership with the NHS

Considered trustworthy by patients, the public and healthcare professionals

Deliver internationally competitive research services

Remove fragmentation in the research data ecosystem

Generate revenue to reinvest in service improvement and the NHS

Service model options

Within NHS England

Commercial arm of NHS England

Under the remit of a government department as an arm's length trading fund

Non-ministerial trading fund

Registered charity

Limited company owned by government

Public-private partnership

Pros

Cons

Populated for each service model option based on evaluation criteria

Service model option 1
Within NHS England

Pros

- Existing remit to support secondary uses of national data collections, including issuing directions on behalf of the Secretary of State for Health and Social Care.

- Has statutory right to collect identifiable health data.

- Builds on existing infrastructure and can leverage economies of scale.

- NHSE recognised by patients and the public as having a role in collecting and using health data.

Cons

- Budget is not ring-fenced and can be/has been utilised to support frontline services.

- Current resources subject to diversion to other NHS priorities, leading to persistent issues providing timely data access for research.

- Full range of requisite specialist expertise not currently available in house. Head count restrictions prevent effective delivery of current service offerings and would not support expanded services.

- Operations constrained by annual budgets, with inability to carry forward any surplus, which creates uncertainty and hampers delivery and planning.

- Perceived risk by GPs that collection of data for secondary uses may also be used for performance management.

- Service would require creation of new senior leadership post with internal accountability akin to Chief Scientific Advisor Genomic Medicine Service role.

Service model option 2
Commercial arm of NHSE

Pros

- Remit dedicated to research services without competing NHS delivery priorities.

- No legislative change to collect identifiable health data.

- Ability to generate income and flow surplus into continuous improvement and potentially to the NHS.

- Potential for workforce capacity flexibilities to meet service demand.

- Builds on existing infrastructure and can leverage economies of scale.

- NHSE recognised by patients and the public as having a role in collecting and using health data.

- Ability to create senior leadership post directly accountable to NHSE Executive.

Cons

- May be challenging for budget to be ring-fenced and not subject to diversion to support core NHSE or frontline services.

- Establishment of separate NHSE data body may be perceived as recreating NHS Digital, and be at heightened risk of merger back into NHSE as part of future NHS budget efficiencies.

- Salaries would follow NHS pay scales and would need to apply for exemptions for specialist roles.

- While there can be some delegated freedoms, operations may be restricted by government operational frameworks.

- Perceived risk by GPs that collection of data for secondary uses may also be used for performance management.

Service model option 3
Government department arm's length body trading fund

Pros

- Remit focused on research service delivery without competing priorities.

- Budget dedicated to delivering research services.

- Ability to generate income and invest surplus into expansion of services and continuous improvement.

- Flexibility to hire staff with requisite skills with potential to expand capacity to meet service demand.

- Government services generally trusted by healthcare professionals and the public.

- Ministerial support within the context of departmental responsibilities.

- Service led by senior leader.

Cons

- Overhead costs associated with supporting an independent government body, including establishing own research, IT and governance infrastructure.

- Salaries restricted to civil service pay scales and headcount may be subject to government restrictions.

- Not part of the NHS, which could affect trust by the public and healthcare professionals.

- Ability to collect identifiable data might necessitate legislative change (but might be achievable with a direction from the Secretary of State or alternative means).

Service model option 4
Non-ministerial trading fund

Pros

- Remit focused on research service delivery without competing priorities.

- Budget dedicated to delivering research services.

- Ability to generate income and invest surplus into expansion of services and continuous improvement.

- Flexibility to hire staff with requisite skills with potential to expand capacity to meet service demand.

- Government services generally trusted by healthcare professionals and public.

- Service led by senior leader.

Cons

- Overhead costs associated with supporting an independent government body, including establishing own research, IT and governance infrastructure.

- Salaries restricted to civil service pay scales and headcount may be subject to government restrictions.

- Not part of the NHS, which could affect trust by the public and healthcare professionals.

- Ability to collect identifiable data might necessitate legislative change (but might be achievable with a direction from the Secretary of State or alternative means).

Service model option 5
Registered charity

Pros

- Remit focused on research service delivery without competing priorities.

- Budget dedicated to delivering research services.

- Able to receive funding from third parties.

- Depending on the degree and extent of public sector control, may have flexibility to operate outside public body policies, including bringing in requisite skills at competitive pay scales, with capacity to expand to meet service demand.

- Service led by senior leader.

Cons

- Overhead costs associated with supporting an independent registered charity, including establishing own research, IT and governance infrastructure.

- Subject to Charity Commission requirements and obligations. Rules may restrict spending of any surplus.

- Depending on the degree and extent of public sector control, charity could be classified as a public body, which would limit some operational flexibilities.

- Not part of the NHS, which could affect trust by the public and healthcare professionals.

- Ability to collect identifiable data might necessitate legislative change (but might be achievable with a direction from the Secretary of State or alternative means).

Service model option 6
Limited company owned by government

Pros

- Remit focused on research service delivery without competing priorities.

- Budget dedicated to delivering research services.

- Ability to generate income.

- Depending on the degree and extent of public sector control, may have flexibility to operate outside public body policies, including retaining and investing surplus into expansion of services and continuous improvement, and bringing in requisite skills at competitive pay scales, with capacity to expand to meet service demand.

- Supported by government advocacy.

- Service led by senior leader.

Cons

- Overhead costs associated with supporting an independent limited company, including establishing own research, IT and governance infrastructure.

- May need to comply with government operational frameworks and salary scales.

- Not part of the NHS, which could affect trust by the public and healthcare professionals.

- Ability to collect identifiable data might necessitate legislative change (but might be achievable with a direction from the Secretary of State or alternative means).

Service model option 7
Public-private partnership

Pros

- Remit focused on research service delivery without competing priorities.

- Depending on the degree and extent of public sector control, may have flexibility to operate outside public body policies, including retaining and investing surplus into expansion of services and continuous improvement, and bringing in requisite skills at competitive pay scales, with capacity to expand to meet service demand.

- Government does not have to bear the full cost and has potential to become fully financially sustainable.

- Partnership offers opportunity to have a more competitive operating model and service offering.

- Service led by senior leader.

Cons

- Overhead costs associated with supporting an independent company, including establishing own research, IT and governance infrastructure.

- Heightened risk of public and healthcare professional trust concerns due to private sector involvement.

- Pressure to achieve financially sustainability and return on investment for private sector partners could result in higher data access costs.

- Private sector investment could result in preferential data access demands and ongoing tensions between government and private investor priorities.

- Ability to collect identifiable data might necessitate legislative change (but might be achievable with a direction from the Secretary of State or alternative means).

Assessment of service models based on evaluation criteria and deliverability

	Pros / Cons	Likelihood of sustainable delivery	Service model	Rationale
	Populated for each service model option based on evaluation criteria	**Existing models delivering to international standards**	government trading fund: - **arm's-length body** - **non-ministerial body** - **government limited company** - **charity**	Have examples of existing models (CPRD, Genomics England and UK Biobank) that demonstrate these models can deliver timely data for researchers to internationally competitive standards.
		Unproven	- **commercial arm of NHSE**	No current comparable model. Independence from NHSE would need to be defined to deliver essential criteria for service.
		Major challenges to overcome	- **within NHSE** - **public private partnership (PPP)**	Constraints within NHSE causing current issues, and there are trust and conflicting priorities issues associated with a PPP. These issues are unlikely to be resolved in near-to-medium term.

Recommendations for next steps

HDRUK
Health Data Research UK

abpi

Recommendation 1

Develop a shared vision across stakeholders

- appoint the Data for R&D programme Strategic Advisory Group, to develop a shared vision and aligned priorities for the UK health data research ecosystem
- agree realistic phased implementation plan to address greatest needs first and maximise return from government spending on health data research infrastructure

Recommendation 2

Engage with the public, patients and GPs

- consult with GPs, patients and members of the public to hear their views on service model options, governance, and public participation and involvement

Recommendation 3

Build on services that are already delivering

- build on experience and expertise of existing research data services with a track record of delivering researcher needs to academia, regulators and industry worldwide e.g. CPRD and UK Biobank
- avoid recreating the wheel or perpetuating fragmentation in all existing services by adopting user-centred design to drive investment decisions

Recommendation 4

Choose a sustainable service model that will deliver all the essential criteria

- validate pros and cons assumptions in the options appraisal
- ensure service model chosen is scalable, sustainable and can meet all the essential criteria defined for a health data research service, including delivering services to an internationally competitive standard

Glossary

Clinical Practice Research Datalink (CPRD): A health data service providing de-identified patient data for research purposes. https://www.cprd.com/

Data access policy: Changes to how NHS data is accessed including the move from a system of data sharing to a system of data access. https://www.gov.uk/government/publications/data-access-policy-update/data-access-policy-update

Data Access Request Service (DARS): NHS England's gateway to accessing NHS health and social care data. https://digital.nhs.uk/services/data-access-request-service-dars

Data asset: A collection of data that has value for research, such as the UK Biobank. https://www.ukbiobank.ac.uk/

Data linkage: The process of connecting different datasets for a more comprehensive analysis. https://bmjmedicine.bmj.com/content/1/1/e000087

Data (Use and Access) Bill: A proposed Government Bill with implications for health data sharing. https://bills.parliament.uk/bills/3825

Genomics England: Company wholly owned by government that analyses sequenced genomes for the NHS and provides genomic data for research. https://www.genomicsengland.co.uk/

Health data research service: A service providing secure access to health data for research and development purposes.

Longitudinal data: Data collected over a long period to track changes and trends. https://www.dundee.ac.uk/stories/data-safe-havens-unique-longitudinal-data-drives-healthcare-innovation

National Data Library: A proposal from the Labour Party Manifesto to "bring together existing research programmes and help deliver data-driven public services. https://labour.org.uk/wp-content/uploads/2024/06/Labour-Party-manifesto-2024.pdf

NHS commissioning support units: Independent units that provide commissioned services back to the NHS. https://www.kingsfund.org.uk/insight-and-analysis/long-reads/what-commissioning-and-how-it-changing

NHS England Secure Data Environment: NHS England's Secure Data Environment platform providing access to nationally collected data sets, formerly the responsibility of NHS Digital https://digital.nhs.uk/services/secure-data-environment-service

NHS Digital (NHSD): Former national provider of data and IT systems for the NHS, now merged into NHS England. https://digital.nhs.uk/

NHS 10 Year Health Plan: Upcoming plan for how government will deliver an NHS fit for the future. https://www.england.nhs.uk/long-term-plan/

Omics data: Biological data related to genomics, proteomics, or other molecular-level information. https://rd-connect.eu/what-we-do/omics/

Glossary

OpenSAFELY: NHS England commissioned software platform for analysis of electronic health records data. https://www.opensafely.org/

Public-private partnership (PPP): A collaboration between government and private companies for service delivery. https://www.independent.co.uk/voices/nhs-private-public-partnership-b2158452.html

Regional SDE: Secure data environment providing access to health data at a regional level. https://digital.nhs.uk/data-and-information/research-powered-by-data/sde-network

Siloed data: Data stored separately without easy access across systems. https://www.computerweekly.com/news/252514852/Siloed-data-holding-back-coordinated-health-responses

Sudlow Review: An independent review highlighting the need for unified health data access in the UK. https://www.hdruk.ac.uk/helping-with-health-data/the-sudlow-review/

Trusted research environment (TRE), also referred to as Data Safe Haven or Secure Data Environment (SDE): A secure space where sensitive data can be accessed and analysed without leaving the environment. https://understandingpatientdata.org.uk/secure-data-environments

UK Biobank: A biomedical database and research resource, established as a charity in 2006, containing data and samples from half a million UK participants. https://www.ukbiobank.ac.uk/

User-centred design: A process that focuses on what users need throughout development. https://www.experienceux.co.uk/faqs/what-is-user-centred-design/

DHSC = Department of Health and Social Care
HRA = Health Research Authority
AMRC = Association of Medical Research Charities