Secure Data Environments for Research and Development

Expression of Interest

September 2022

Table of Contents

1. Summary

NHS England is investing in the development of an interoperable network of sub-national Secure Data Environments for Research and Development (SDEs for R&D) across England. SDEs for R&D will enable approved researchers secure access to privacy protected NHS data.

We are now inviting Expressions of Interest (EOI) from NHS partnerships who are implementing, or are planning to implement, a Secure Data Environment for Research and Development.

This pack is for NHS partnerships who are already working, or want to work in partnership, with others in their locality to develop an SDEs for R&D.

It sets out:

- vision, context, and benefits of SDEs and specifically, SDEs for R&D
- expectations for SDEs for R&D and localities
- funding and timescales

The Expression of Interest Questions and Response template that localities should complete can be found in **Appendix A**. Evaluation criteria are included in **Appendix D**.

If localities have existing documentation relating to a question in the EOI Questions and Response template, they are strongly encouraged to submit this as evidence and reference it in the EOI Questions and Response template. There is no need to duplicate evidence between the EOI Questions and Response template and other submitted documents/weblinks. If localities are less advanced and do not already have documentation, they are expected to complete the EOI Questions and Response template.

The deadline for receipt of submissions is **17:00 on Friday 21st October 2022.**

2. Secure Data Environments

2.1 Policy Context

In the Data Saves Lives strategy, we committed to implementing Secure Data Environments (SDEs) as the default way to access NHS health and social care data for research and development. Moving to a system built on data access will allow for greater privacy, security and transparency, as well as meet core recommendations of the Goldacre Review and data commitments in the Life Sciences Vision.

2.2 What is a Secure Data Environment?

Secure Data Environments (SDEs) are data storage and access platforms, which uphold the highest standards of privacy and security of NHS health and social care data when used for research and analysis. They allow approved users to access and analyse data without the data leaving the environment.

Secure Data Environments allow organisations to control:

- Who can become a user to access the data
- The data that users can access
- What users can do with the data in the environment
- The information users can remove

A range of different users will benefit from improved access to NHS health and social care data. These users have different data requirements and skill sets and need to access data to produce different outputs. Broadly we categorise these differences into two groups:

1. SDEs for planning and population health management, such as the NHS COVID-19 data store, the planned 'federated data platform', and regional / ICS data platforms. Primary use is for internal planning and management, for instance by integrated care groups, NHS, and public health analysts.

2. SDEs for research and development, such as the platforms created by NHS Digital, OpenSafely and SAIL. Primary users include academic and industry researchers, as well as policy analysts, with a specific research question.

2.3 How Secure Data Environments will be delivered

The Secure Data Environment policy aims to simplify a complex, rapidly developing landscape. A number of key investments are being made to ensure this policy works in practice.

2.3.1 NHS Digital's (NHSD) National Secure Data Environment

NHS Digital is the current safe haven for health and care data. NHS Digital is currently piloting a National Secure Data Environment, which provides approved researchers from trusted organisations with timely and secure access to NHS health and social care data.

This national environment currently supports the work of over 100 users from across the NHS, academia, industry, and charity sectors. For example:

- The British Heart Foundation is researching the impact and effects of the COVID-19 pandemic on cardiovascular diseases.
- DATA-CAN - an innovation hub funded by Health Data Research UK - is working to understand the impact of COVID-19 on people affected by cancer.

With support from the NHS Transformation Directorate's Data for Research & Development programme, enhancements to this environment are planned throughout 2022 and beyond. This will include expanding the pilot to accommodate more users, with the aim that all data held nationally is managed through a Secure Data Environment when used for research and planning purposes.

2.3.2 Federated Data Platform

NHS England intends to procure a Federated Data Platform (FDP), which is an ecosystem of technologies and services to be implemented across the NHS in England. This will be an essential enabler for transformational improvements across the NHS.

The FDP will enable, and must apply, Secure Data Environment policy for any use of NHS health and social care beyond direct patient care. For example, when using data to support population health management and operational planning. This procurement will also support Integrated Care Systems to implement Secure Data Environment policy.

The FDP has five use cases:

- Population health and person insight
- Care co-ordination (ICS)
- Elective recovery (NHS Trust)
- Vaccines and immunisation
- Supply chain

The national FDP programme does not include a research use case and is not funded to do so.

2.3.3 Sub-national Secure Data Environments for Research and Development (SDEs for R&D)

During 2022, as a result of a competitive process, four localities received investment to scope and define how NHS-owned SDEs for R&D might work best at a regional level. This work considered how to provide researchers and analysts with access to NHS health and social care data at a significant 'regional' scale, maintain patient confidentiality, and ensure connectivity to local communities and NHS care teams. The exploratory work on 'regional' SDEs for R&D has helped to create a community of practice from which we can learn and has informed the development of SDE policy.

2.4 Secure Data Environment Policy Guidelines

Policy guidelines set out expectations for how Secure Data Environments will be used to access NHS health and social care data. They have been developed in collaboration with leading experts in the field.

These guidelines have been developed to:

- Strengthen public confidence and trust in the transition to using Secure Data Environments to access NHS health and social care data
- Provide additional information about the use of Secure Data Environments, as outlined in the Data Saves Lives Strategy.
- Describe the foundations on which the NHS Transformation Directorate will further develop Secure Data Environment policy, in collaboration with the public and expert stakeholders.

- Communicate the direction of travel for Secure Data Environment policy signalling areas that require further development.
- Communicate the fundamental principles which Secure Data Environments must adhere to.

Some choices have already been made to ensure successful implementation, including all NHS health and social care data being accessed through a Secure Data Environment, with exceptions strictly limited. We also commit to establishing an accreditation process and an organisation that will ensure compliance, which in turn will standardise and limit the number of platforms that can provide access to NHS data. We recognise that these guidelines do not contain the full details to support a transition that will require significant changes in behaviour and process. The next phase of work will involve translating these high-level ambitions into workable practice, which will be supported by broad engagement.

2.5 The Five Safes Framework

Policy guidelines are arranged according to the Five Safes framework developed by the Office for National Statistics (ONS).

The Fives Safes are widely regarded as representing best practice in data protection, they include:

1. **Safe Settings** - the environment prevents inappropriate access, or misuse.
2. **Safe Data** - information is protected and is treated to protect confidentiality.
3. **Safe People** - individuals accessing the data are trained, and authorised, to use it appropriately.
4. **Safe Projects** - research projects are approved by data owners for the public good.
5. **Safe Outputs** - summarised data taken away is checked to ensure it protects privacy.

Full details of the policy guidelines can be found at:
https://www.gov.uk/government/publications/secure-data-environment-policy-guidelines/secure-data-environment-for-nhs-health-and-social-care-data-policy-guidelines

3. Sub-national Secure Data Environments for Research and Development

NHS England's Data for Research & Development (R&D) programme will, subject to final HMT approval, invest approximately £100 million over three years (2022-25) to support the development and operation of an interoperable network of a small number of sub-national SDEs for R&D.

This will be underpinned by critical work related to information governance policy, commercial principles for data access by academic and industry partners, and public, patient, and professional involvement and engagement, related directly to the Data for R&D programme and the wider Data Saves Lives strategy implementation.

SDEs for R&D are expected to:

- enable NHS organisations to maximise the use of their data for research and innovation and increase opportunities for a wider range of health and care teams and populations to become involved in research
- ensure connectivity to local communities and NHS care teams and facilitate a learning health system where data insights can be actioned to drive real-world change
- improve understanding of diseases, interventions, and pathways, which will translate to improved patient care locally and nationally
- increase public confidence in health and care data research activities and related benefits to the NHS
- make it easier for researchers to understand which data assets are available, whether they meet their requirements, and create a more streamlined and timely mechanism for obtaining approvals to access data
- improve efficiency and value for money by enabling creation of research-ready data assets and avoiding duplicative data curation processes.

3.1 Vision for SDEs for R&D

Our vision for this network follows some key principles:

- This investment will provide the infrastructure to ensure that SDEs for R&D become the default route for researchers to access NHS health and social care data (with limited exceptions that are being formally defined). It is closely linked to other investments in NHS data infrastructure, to ensure efficiency and value for money. This will improve understanding of diseases, interventions, and pathways, which will translate to improved patient care locally and nationally.

- There should be as few SDEs for R&D as possible, each meeting minimum target populations, fostering collaboration and aligning to patient flows, avoiding duplicative investments in similar infrastructure and data curation, and ensuring services offered are at a scale at least comparative to international centres of excellence.

- Over time, allow users to conduct approved research on de-identified, linked, longitudinal, multi-modal NHS health and social care data at significant scale, both within and collaboratively between the SDEs.

- SDEs for R&D will form part of an interoperable network, adhering to common core technology, governance, and operational process standards. This is to ensure users can conduct analysis within and across SDEs. To enable this, some functions will be coordinated by a central ‘hub’.

- The development and deployment of this network is not purely a technology endeavour. It is a programme of transformation with technology, operations, information governance, financial and workforce elements that will need to be delivered together across ICS, regional and national teams to achieve success.

- Local teams will be responsible for most SDEs for R&D delivery and management. The national team ‘central hub’ will focus on unblocking common challenges faced across the network, for example, related to Information Governance (IG) and commercial challenges.

3.2 Our expectations

- SDEs for R&D should provide researchers with access to near real time, multi-modal NHS health and social care data at a significant ‘regional’ scale. Multimodal data includes, but is not limited to, structured and unstructured data from electronic health records and case management systems, diagnostic test results and images, wider social determinants of health and data from wearables and devices.

- SDEs for R&D must remain under NHS control but a collaborative partnership of organisations within the region is expected to contribute to successful delivery. This is likely to include NHS Trusts, Integrated Care Boards, Local Authorities, Higher Education Institutions, Academic Health Science Networks and Research Collaboratives as a minimum.

- It is expected each SDE for R&D will cover a population of no less than 3 million local people initially with plan to scale to 5+ million by 2025.

- The development and operation of the SDE for R&D must be led by an NHS organisation and overseen by a board that includes representation from participating NHS organisations and patients and the public within the region.

- NHS health and social care data must remain under the control of NHS health and social care organisations in the SDE for R&D. NHS data should not be made available to researchers outside an SDE for R&D except for limited exceptions, which will be defined in policy guidelines.

- The flow of data from participating NHS organisations, and its subsequent processing for the purposes of research, should meet the requirements of the Data Protection Act 2018 and the Common Law Duty of Confidentiality.

- SDEs for R&D should have a single Data Access Committee, with patient and public representation, where decisions are made as to which researchers/projects should be granted access, subject to appropriate data protection requirements. This should encourage access decisions to be made more rapidly, consistently, and at a larger

scale than is often currently the case. Localities will also need to have a clear and robust process for ethics approval for research.

- SDEs for R&D must involve patient and public representatives in governance and decision making. Ongoing public engagement is expected to build trust and increase public confidence in health and care data research activities and related benefits to the NHS.

- SDEs for R&D should provide the infrastructure and analytical tooling to support the following types of research use cases:

    - AI/algorithm development (testing, training, and validation)
    - Clinical trial activities (feasibility, recruitment, efficacy through short- and long-term trial follow up)
    - Real world studies (safety, effectiveness, cost effectiveness)
    - Translational research (academic discovery and implementation of discovery into practice)
    - Epidemiological studies (large cohorts for population health research)
    - Health systems research (evaluation of systems or processes, including operational and applied research)

- SDEs for R&D must have an environment where data is processed and curated to create 'research-ready' datasets and secure analytical workspaces with appropriate tooling where researchers conduct analysis on de-identified data.

- SDEs for R&D must incorporate NHS SDEs Technical capability specifications into their design and have a roadmap to meet the mandatory requirements for SDE accreditation. See **Appendix C** for full detail of requirements.

- Data should be discoverable so that potential users can understand the high- level contents of the data available, how it has been used previously and relevant details of its lineage.

- SDEs for R&D must have robust identity management, attribute-based access control and ability to suspend or revoke access.

- SDEs for R&D must preserve individuals' privacy through data minimisation, use of privacy enhancing technologies, avoidance of disclosive information or ability to export information about individuals in a re-identifiable format.

- Patient and public trust must be maintained through transparency of governance and decision making, time limited access permissions, audit logs of activity and data use registry.

- SDEs for R&D must adhere to cybersecurity requirements and mandatory standards, including:
    - ISO 27001 for infrastructure and hosting layer, and the NHSD cloud security good practice guide where applicable.
    - NHS Digital Data Security and Protection Toolkit compliant.
    - Cyber Essentials Plus: Supplier has been independently assessed and verified by a Government approved external body that it meets the Cyber Essentials implementation profile [BIS/14/696].
    - The Security of Network and Information Systems Regulations (NIS Regulations) 2018.
    - Data Centre Alliance Class 3 Facility European Code of Conduct (EUCOC): Compliant.

- SDEs for R&D must have the following supporting services in place
    - Technical and data services
    - Communications and business development
    - User permissions
    - Commercials and finance
    - Technical and data services
    - Communications and business development
    - User permissions
    - Commercials and finance
    - Patient and public engagement

- Operational management
- Customer/account management and researcher support

- Data partnerships, including commercial partnerships, entered must be transparent, non-exclusive and deliver benefit to the NHS. Local frameworks developed should support fair value revenue sharing with other locality partners as outlined in the Department of Health & Social Care's Five Principles.

- SDEs for R&D must commit to following the commercial principles being developed by the Centre for Improving Data Collaboration (CIDC), as outlined in **Appendix B.**

- SDEs for R&D teams are expected to participate in a collaborative community of practice; sharing best practice and learning, developing operational blueprints that can be used by other teams, and highlighting challenges requiring support from the NHS England 'central hub'.

4. Number of SDEs for R&D

We plan to invest in a small number of SDEs for R&D, which will operate at significant geographic scale.

We recognise there are multiple factors to consider, and trade-offs to make, when determining what the exact number should be.

These include:

- An ambition for data assets to cover population sizes of 5 million+ (to meet the needs of academic and life science researchers), while recognising that many of the existing health and care networks and digital collaboratives operate across smaller geographies.

- Improving the timeliness of access to data and reducing the number of data access approval committees that researchers need to navigate, while ensuring local data

controllers remain involved in decision making and can retain the trust of their local populations and health and care stakeholders.

- Aligning where possible with existing investment in SDE infrastructure by NHS trusts and ICSs, while incentivising collaboration to maximise return on investment and ensure services established are comparable to international exemplars.

- Encouraging localities who are more digitally mature to progress at pace, while ensuring that those who are less ready to progress can still contribute to key design and governance decisions.

Through close working with NHS teams and research stakeholders over the last nine months, and considering international examples, our analysis suggests there should be approximately 10 SDEs for R&D across England that provide high levels of population coverage.

5. Funding Opportunity

The Data for R&D programme will invest approximately £100m in SDEs for R&D, subject to final approval of the Programme Business Case by HM Treasury.

Each NHS SDE for R&D will receive up to £10 million over three years. Funding will be allocated in three tranches aligned to a set of agreed deliverables and timescales. These are referred to as Work Packages 1-3.

The exact amount of funding allocated to each SDE for R&D may vary depending on the size of the population within the geography, anticipated delivery timescales and the total number of SDEs for R&D that receive investment.

Investment will be offered to a locality within each NHS England region to ensure all regions are able to commence work on implementation of a SDE for R&D. It is not expected that investment will be allocated to more than two localities in the same NHS England region. **It is strongly advised that localities within the same region who are planning to submit separate Expressions of Interest Questions and Response template should consider**

opportunities for collaboration before they submit their bids. If needed this can be discussed with the NHSE Data for R&D team.

Funding will be a combination of Public Dividend Capital and revenue, with a capital: revenue split of approximately 30:70 over the three-year funding period. Funding allocated to an SDE for R&D each year must be spent within the same financial year and cannot be carried over to the following year.

Public Dividend Capital can only be allocated to an NHS Trust. The nominated Finance Lead for the locality is therefore expected to be an NHS Trust Director of Finance. The NHS Trust is expected to account for any assets created using Public Dividend Capital in line with Standard Financial Instructions. The NHS Trust which receives investment on behalf of the locality should have a mechanism in place to distribute funding to other NHS, or non-NHS, delivery partners and to receive contributions from other partners towards capital charges and depreciation costs, such as through a Memorandum of Understanding.

There is no requirement for financial matched funding. However, it is expected that localities will be able to demonstrate significant 'in-kind' investment. This may be through existing or planned local investment in technical or data infrastructure or provision of resources and subject matter experts who will contribute to implementation and operational delivery of the SDE for R&D. SDE for R&D teams will be asked to include an estimate of the total value of local 'in-kind' investment with their strategic plans and business case.

5.1 SDEs for R&D Work Packages and Delivery Waves

SDE for R&D funding will be aligned to three work packages.

Work Package 1: Strategic Design Phase (2-6 months). This phase of work is focused on completion of an SDE for R&D business case and associated documents outlining the SDE for R&D design and operating model.

Work Package 2: Delivery of Minimum Viable Product (MVP) (12 months). This phase of work is focused on technical delivery of the SDE for R&D; establishing the services required to run the SDE for R&D; setting up the Data Access Committee and associated Information

Governance requirements; patient and public engagement and enabling a small number of research projects to commence using the MVP SDE for R&D.

Work Package 3: Continued Delivery and Interoperability (12 months). This phase of work focuses on continued delivery of the SDE for R&D; onboarding of new and varied data assets; expansion of support for new research projects and demonstration of interoperability between SDEs for R&D.

Localities chosen to receive investment in each NHS region will be allocated to two delivery Waves.

Wave 1: These localities will already be at advanced stages of planning for, or delivery of, their SDE for R&D. They are expected to make demonstrable progress at pace within the next 12 months and to work together as a community of practice, creating service blueprints and contributing materials to a resource hub. They are expected to be early adopters of policy guidance and commercial principles and will identify common delivery challenges that the Data for R&D team will help to resolve.

Wave 1 SDEs for R&D should be able to share detailed strategic plans and a business case with the Data for R&D team by 16th December 2022 at the latest.

It is expected that three localities will be identified as Wave 1 SDEs for R&D, who will each receive a total of £2.5 million Public Dividend Capital and £1.3 million revenue this financial year. Wave 1 SDEs for R&D **must** be able to manage this investment in year, as it cannot be carried over to 23/24.

We would expect localities submitting EOI Questions and Response template (**Appendix A**) for Wave 1 funding to be able to provide evidence of the following:

- Commitment from an NHS Trust Finance Director that the regional partnership can spend £1.3 revenue and £2.5m PDC this financial year.

- Commitment from NHS leadership teams that this is a system priority and sufficient resources will be allocated to facilitate delivery at pace.

- Established governance structure with representation from all NHS stakeholders to ensure effective decision making.

- Evidence of patient and public involvement in governance and decision making.

- A single Data Access Committee in place or plans to implement one.

- Plans to ensure data for research can be processed legally.

- A technical strategy for development of the SDE for R&D, including plans to build, buy or develop different components.

- A roadmap to ensure the SDE for R&D meets the mandatory SDE accreditation criteria described in **Appendix C**.

- Existing strategic delivery plans, or a business case, which have been approved through local governance structures.

Wave 2: These localities will have commenced planning or will be at an early stage of planning for their SDE for R&D and should have begun to establish a governance structure which enables effective decision making across the locality, with input from all key stakeholders. Localities should have a leadership team in place who are progressing the development of strategic plans. Localities may have undertaken some exploratory work to determine what technical and research infrastructure already exists across the locality that could contribute to the SDE for R&D design.

Wave 2 SDEs for R&D will receive £300,000 revenue this financial year to develop their strategic plans and business case. This work should be completed by 31st March 2023.

It is expected that up to seven localities will be identified as Wave 2 SDEs for R&D, ensuring there is at least one SDE for R&D in each NHS England region.

A Memorandum of Understanding with be developed with each SDE for R&D team which will specify the exact funding allocation and expected deliverables. This will be developed in collaboration with each SDE for R&D team. Localities will not be expected to submit further bids for funding bids.

SDE for R&D teams are expected to liaise regularly with the Data for R&D team and to include representatives of the team on their main governance committee so there is continual oversight of progress. All SDE for R&D teams are expected to participate in the community of practice. The Data for R&D team will approve release of each tranche of funding subject to satisfactory progress with each phase of work.

6. Submission and evaluation process

6.1 Submission

Localities who wish to submit an Expression of Interest, should complete, and submit the EOI Questions and Response Template (**Appendix A**) along with any supplementary supporting information to Data.research@nhs.net.

Localities who have existing documentation are strongly encouraged to submit this as evidence and reference it in the EOI Questions and Response template. There is no need to duplicate evidence between the EOI Questions and Response template and other submitted documents/weblinks. Localities who do not have existing documentation should respond by completing the EOI Questions and Response Template form.

The deadline for receipt of submissions is **17:00 on Friday 21st October 2022.**

6.2 Evaluation

Submissions will be assessed and scored based on the EOI Evaluation Criteria outlined in **Appendix D.**

If you have any questions regarding this EOI process, please send them to Data.research@nhs.net

7. Timetable

Date:	Event:
21st September 2022	Launch Expression of Interest.
21st September – 20th October 2022	Regional partnerships complete EOI submissions and submit any questions requiring clarification.
21st September – 20th October 2022	NHSE Data for R&D team available to present at Data & Analysis Regional Roadshows.
Between 11th and 13th October 2022	Live Q&A session open to all localities.
21st October 2022	Deadline for Expression of Interest submissions to NHS England.
31st October 2022	Complete assessment of Expression of Interest submissions.
1st November	Announce Wave 1 SDEs for R&D and develop MOUs to release funding for Work Package 1.
1st - 30th November	Follow up and clarification with regions to confirm Wave 2 SDEs for R&D.
1st December 2022	Announce Wave 2 SDEs for R&D and develop MOUs to release funding for Work Package 1.
16th December 2022 (latest)	Wave 1 SDEs for R&D submit strategic plans and business case for approval to release Work Package 2 funding.
31st March 2023 (latest)	Wave 2 SDEs for R&D submit strategic plans and business case for approval to release Work Package 2 funding.

8. Appendices

Appendix A- SDEs for R&D EOI Questions and Response Template

Appendix B- SDEs for R&D EOI Commercial Principles

Appendix C- (Draft) SDEs for R&D Technical Capability Specifications

Appendix D- SDEs for R&D EOI Evaluation Criteria

Appendix E- SDEs for R&D EOI FAQs

Appendix F- SDEs for R&D EOI Data Assets Table