

Health and care: Template data protection impact assessment (DPIA)

Background

A [data protection impact assessment \(DPIA\)](#) will help you to identify and mitigate potential data protection risks to an acceptable level before using or sharing (processing) data that identifies individuals (personal data).

A DPIA will also help you meet a number of data protection legal requirements including:

- [Data protection by design](#) - privacy and data protection issues must be considered at the start, or in the design phase, of a new system, product or process, then continuously while it exists.
- [Accountability](#) - your organisation is responsible for showing how it complies with data protection laws.
- [Transparency](#) - personal data must be used and shared in a transparent way.
- [Security](#) - adequate measures need to be in place to protect data. This can range from policies and procedures to technical security measures such as encryption of data.

DPIAs are mandatory when there is a high risk to individuals, such as when using the health and care data of a large number of people. However, health and care organisations are strongly advised to complete a DPIA when using and sharing personal data in a new or substantially changed way.

A DPIA involves a risk assessment. If a high-level risk remains after applying mitigations, then you must consult with the Information Commissioner's Office (ICO) for further advice before starting to collect, use or share the data.

A DPIA is a live document - you must update it if there are any changes to:

- the purpose - why you are proposing to use or share personal data
- the manner - how you will use or share the data
- who is involved - the organisations using and sharing personal data

This is a template DPIA for health and care organisations. We encourage organisations to adopt it. The template is written so that it is easy to use without needing expertise in data protection. It is the responsibility of the organisation which is deciding on why and how the data is being used and shared (known as the controller), to ensure that the DPIA is completed appropriately.

In the case of research, the sponsor is the controller. See Health Research Authority (HRA) guidance on [controllers](#) and research. HRA guidance on [DPIAs](#) sets out that sponsors should complete a DPIA for the broad range of health and care research they sponsor and ensure that individual research projects are designed in accordance with the DPIA. Individual DPIAs should only need to be completed for individual research projects that involve activities beyond the generic research DPIA. Where the study deviates from the established processes (for example, where it is intended that a project uses a new technology for the processing of personal data, or requires that

safeguards set out in standing policies cannot be applied), the sponsor should consider whether a study specific DPIA is appropriate to address the level of risk, or whether updating existing DPIA(s) will be sufficient. Research sites should not complete DPIAs or request researchers to complete individual DPIAs for each research project, as they are not the controller.

“Personal data” retains its meaning under UK GDPR (relating to living individuals only) but the scope of this policy also includes information relating to deceased individuals where relevant.

Text with **green background** is guidance text.

Contents

<i>Data protection impact assessment (DPIA)</i>	4
SECTION 1 – Screening questions.....	4
SECTION 2 – Why do you need the data?	9
SECTION 3 – What data do you want to use or share?	12
SECTION 4 – Where will data flow?.....	20
SECTION 5 – Is the intended use of the data lawful?	22
SECTION 6 – How are you keeping the data secure?	24
SECTION 7 – How long are you keeping the data and what will happen to it after that time?.....	28
SECTION 8 – How are people’s rights and choices being met?	30
SECTION 9 – Which organisations are involved?.....	35
SECTION 10 – What data protections are there and what mitigations will you put in place?	37
SECTION 11 – Review and sign-off.....	42
SECTION 12 – Appendices	43

Data protection impact assessment (DPIA)

Data protection impact assessment (DPIA) title:	DPIA for onboarding General Practice data to Wessex SDE
---	---

SECTION 1 – Screening questions

1. Do you need to do a DPIA?

- ☒ Yes
☐ No

A DPIA is required for the Wessex SDE GP data onboarding because the processing meets multiple high-risk criteria under UK GDPR and ICO guidance:

- **Use of special category data at large scale**
The project involves extensive use of identifiable and pseudonymised health and care data, including full GP records and sensitive attributes (e.g. health status, ethnicity, genetic data).
- **Large-scale processing affecting many individuals**
Data is collected from all GP practices across Dorset and Hampshire & Isle of Wight, meaning a very large population is impacted.
- **Systematic and extensive processing for research purposes**
Data is linked, cleaned, and analysed within the SDE for population health, clinical research, and service improvement, including linkage across multiple datasets.
- **Use of new technology and infrastructure (DPE/TRE model)**
The Secure Data Environment introduces a new technical architecture (DPE/TRE, cloud-based platforms, controlled researcher access) for managing and analysing sensitive data.
- **Potential risk of re-identification and complex data flows**
Although data is pseudonymised before researcher access, identifiable data is processed for linkage and opt-outs, and data is moved between organisations and environments, creating inherent risks that require formal assessment.
- **Regulatory expectation for research programmes**
Health research sponsors are expected to complete a DPIA for the overall research framework, particularly where large-scale health data processing is involved.
- **Governance change**
The governance model has evolved under the Secure Data Environment (SDE) approach to provide a more structured, standardised, and centralised process

for the review and sign-off of research studies before socialisation with General practice which still maintains controllership and can opt-out.

The internal SDE review process is more robust providing an enhanced review of studies and researchers.

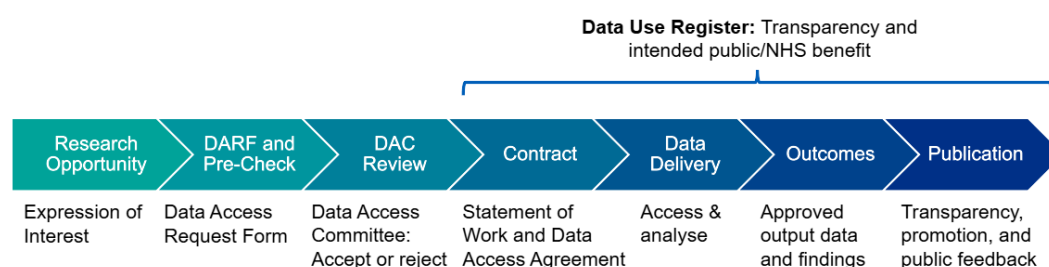
a. Summary of how data will be used and shared

Data is collected by General Practices through their electronic patient record systems. This data is then securely transferred to the Wessex Secure Data Environment (SDE), where it is consolidated, validated, cleaned, and prepared for research and analysis.

Within the SDE, identifiable data may initially be processed in a restricted Data Processing Environment (DPE) solely for purposes such as data linkage, quality assurance, deduplication, and application of opt-outs (type 1, local and national). This processing is undertaken only by authorised technical staff under strict controls.

Following these steps, datasets are pseudonymised and minimised and moved into a Trusted Research Environment (TRE), where approved researchers can access the data in secure, project-specific workspaces. Researchers cannot access identifiable data, remove data from the environment, or access data outside the scope of their approved project.

Data is used for approved purposes aligned with national SDE governance processes, including population health research, clinical research, service improvement, and evaluation. All research projects must undergo robust approval processes (including Data Access Committee review, technical and operational review, and organisational and researcher validation) before access is granted.



Data is not shared externally in raw form. Instead, it remains within the secure SDE infrastructure. Only non-identifiable, aggregated outputs (such as statistical analyses, tables, and reports) are permitted to leave the environment, and these are subject to disclosure controls and manual review to minimise the risk of re-identification.

In some cases, data may be linked with other health and care datasets (e.g. hospital or national datasets) within the SDE to provide a more complete view of patient pathways, but this linkage occurs under strict governance and only for approved studies.

Where artificial intelligence (AI) is used, this will be on a study-by-study basis within the secure environment. AI tools may be used to support analysis or generate insights, but they will operate within the SDE controls. Organisational data will not be used to train external or publicly accessible AI models; any use of AI will follow SDE policies, including appropriate governance, ethical review, and risk assessment.

Overall, this approach shifts from sharing copies of data by organisations externally to a secure data access model, where data remains protected within the SDE (where the data is held) while enabling only approved users to generate insights and outputs that support improved patient care, research, and service delivery.

b. Is any of the data identifiable, pseudonymised or anonymised?

Put an ☒ next to all that apply.

☒ Directly identifiable data

The SDE processes data in different forms at different stages, in line with ICO guidance on identifiability.

The SDE comprises two key environments which are physically separate from one:

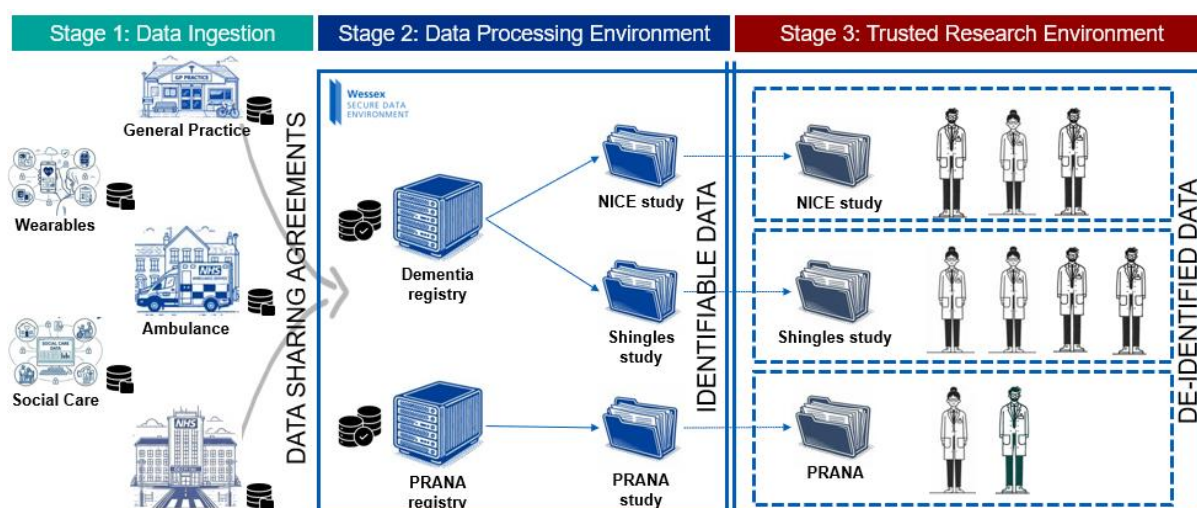
- 1) **Data Processing Environment (DPE)** is where any identifiable data is processed within a restricted environment accessible exclusively by Wessex SDE technical personnel:
 - Identifiable personal data (e.g. NHS number, name, date of birth) is used for:
 - Data ingestion and validation
 - Record linkage across datasets
 - Application of national and local opt-outs
 - This data is NOT accessible to or shared with researchers within the Trusted Research Environment (TRE) – discussed below.
 - Personal identifiable data will be processed exclusively by the SDE's Technical Team in the DPE with the following steps undertaken:
 - Acquisition of the raw primary care data
 - Versioning and cleansing of the raw data
 - Merging the data, such that it is in a consistent format regardless of the system of origin.
 - Tracking and application of opt-outs, including:
 - ✓ National Data opt out
 - ✓ Type 1 opt-outs
 - ✓ A local Wessex SDE Opt-Out – by patient or practice

All data opt outs above will be implemented in line with guidance within 21 days of receipt of notification or sooner within the DPE. This will include removal of data held where a patient has previously provided data but has now decided to opt out.

Where data has already been pseudonymised and provided to researchers in the TRE (see below), the opt out will not be applied retrospectively, but no additional data pertaining to the opted-out patient will be provided for that (or any other) study.

2) **Trusted Research Environment (TRE)** is where research datasets are staged for approved researcher access. The datasets are:

- Pseudonymised (identifiers replaced with codes or hashes)
- Minimised to only what is necessary for the study
- Researchers only have access to the de-identified data applicable to the research study that has been approved by the Data Access Committee (DAC)



☒ Pseudonymised data

As described above, the data is pseudonymised within the DPE before it is made available for research in the TRE, following a structured, standardised process designed to minimise the risk of re-identification. These steps are carried out under an agreed set of procedures set out in the SDE's set of Standard Operating Procedures (SOPs).

- **Pseudonymisation using hashing (SHA256)**
Identifiable data elements (such as NHS number and other direct identifiers) are transformed using SHA256 hashing, replacing them with unique coded values (pseudonyms). This method ensures that individuals cannot be directly identified without access to the separate re-identification key. Each research project applies a distinct hashing approach, meaning pseudonyms are not reusable across projects, further reducing the risk of re-identification.
- **Separation of identifiers and keys**

The additional information required to reverse the pseudonymisation (e.g. encryption keys or mapping tables) is held separately in a secure environment with strict access controls, ensuring that the pseudonymised dataset alone cannot identify individuals.

- **Data minimisation prior to access**

Before datasets are transferred from the Data Processing Environment (DPE) to the Trusted Research Environment (TRE), they are minimised to include only the data necessary for the approved research question. This reduces the amount of potentially identifying information available and lowers overall risk.

- **Governance and approval controls**

The data fields and structure of each pseudonymised dataset are defined within a Data Access Request Form (DARF) and reviewed through the SDE's Pre-Check and Data Access Committee (DAC) approval process, ensuring appropriate use and proportionality.

- **Controlled research environment**

Once pseudonymised, data is transferred to the Trusted Research Environment (TRE), where researchers can access it securely. Researchers cannot access identifiable data or the re-identification keys, and all analysis takes place within this controlled environment. Only fully anonymised aggregated outputs leave the SDE.

☒ Anonymised data

Data within the Wessex Secure Data Environment (SDE) is anonymised at the point of release, following a robust, multi-step process to ensure that individuals cannot be identified by any reasonably likely means.

How the data has been anonymised

- **Aggregation of data**

Data is presented in aggregated form, so that it relates to groups rather than individuals. This removes the ability to identify a specific person.

- **Removal and generalisation of identifiers**

All direct identifiers (e.g. name, NHS number, full postcode) and indirect identifiers (e.g. rare conditions, small populations) are removed or generalised before release.

- **Small number suppression**

Statistical disclosure controls are applied, including:

- Suppression of small numbers (e.g. low counts that could risk identification)
- Secondary suppression where needed to prevent values being inferred
This ensures that no individual can be identified from small or unique groups.

- **Data minimisation**

Outputs contain only the minimum necessary information required to answer the approved research question, reducing identifiability risk.

- **Disclosure control (“airlock”) process**

All outputs are subject to a formal review process before release:

- Automated and manual checks are applied
- Outputs are assessed for re-identification risk
- Only safe, non-disclosive outputs are approved

The data that is shared outside of the SDE consists only of fully anonymised, aggregated outputs. Recipients cannot identify individuals, either directly or indirectly, and do not receive any pseudonymised or identifiable data.

Where a DPIA is not required but you are documenting your decision and the risks, [go to section 10](#) and [11](#) – the other sections do not need completing.

SECTION 2 – Why do you need the data?

2. What are the purposes for using or sharing the data?

The purpose of using and sharing data within the Wessex Secure Data Environment (SDE) is to enable safe, secure use of health and care data to support research, improve services, and deliver better outcomes for patients and the wider population.

At a high level, the purposes are:

- **To improve patient care and outcomes**

Data is used to support research into new approaches to diagnosis, treatment, and prevention, helping clinicians and the NHS deliver safer, more effective care.

- **To understand population health and ensure appropriate service provision**

The data enables analysis of health trends, inequalities, and outcomes across populations, supporting planning so that the right services are in the right place at the right time.

- **To support health and care research and innovation**

Data is used for approved research purposes, including clinical research, population health studies, and service evaluation, contributing to new treatments, improved pathways, and innovation within the NHS.

- **To improve the accuracy and completeness of data**

Processing within the SDE includes data validation, linkage, and standardisation, ensuring higher-quality data that can be relied upon for decision-making and research.

- **To ensure data is available securely when needed**

The SDE provides a secure, controlled environment where approved users can access relevant data without widespread data sharing, improving efficiency while maintaining confidentiality.

- **To reduce burden and support efficient use of NHS resources**

By enabling controlled reuse of data for multiple approved projects, the SDE reduces repeated data transfers and administrative burden on GP practices and the wider NHS.

Data will be used for **six approved purposes** as defined in the Wessex SDE's (Health Research Authority) HRA approval and listed publicly on the Wessex SDE website, including:

- ***New approaches to diagnosis***, making better use of blood tests, medical images, and other information already in patient records.
- ***New approaches to treatment***, including changes to clinical pathways, better support for clinical decisions, and the introduction of new therapies.
- ***New approaches to patient records***, improving the accuracy and relevance of the data recorded and the management of patients across services.
- ***Clinical trial feasibility*** using previously collected clinical and operational data. Identification of potentially eligible trial subjects would take place using pseudonymised data within the SDE DPE but communication with subjects would only be via their care provider (GP, hospital consultant etc) - SDE staff would not approach subjects directly.
- ***Population health research*** to identify factors to aid understanding of health outcomes of a group of individuals including the distribution of such outcomes within groups including study of health determinants, health behaviours and the impact of health policies and interventions.
- ***Operational healthcare research*** using health and social care data to improve decision making, efficiency, and effectiveness in health services and systems.

3. What are the benefits of using or sharing the data?

The SDE Network is a key part of the UK Government's national commitment to move to an operating model of 'data access' rather than 'data sharing' as the default approach to the use of NHS data for research.

Approved Users need insight from this vital data to analyse and link data for research, clinical service evaluation, clinical audit, health surveillance, novel diagnostics and therapeutics, clinical trials enabling, population health and system improvement. Taken together these approaches will reduce burden on the NHS, improve outcomes for patients and help save lives.

Supporting such research via the SDE will reduce the burden on General Practice, the wider NHS (including hospital, community and mental health services), as well as provide increasing support to social care services; all at a time when demand on resources is high.

Using and sharing data through the Wessex Secure Data Environment (SDE) provides significant benefits for individuals, healthcare organisations, and the wider public by enabling safe, efficient, and evidence-driven use of health and care information.

Benefits for individuals (patients and service users)

- *Improved health outcomes and quality of care*
The use of linked data supports research and analysis that leads to earlier diagnosis, more effective treatments, and improved clinical pathways.
- *Better understanding of individual health needs*
Linking data across services provides a more complete picture of a patient's journey, enabling more coordinated and personalised care.
- *Earlier identification of risks and preventative care*
Analysis of large datasets helps identify patterns and risk factors earlier, enabling proactive interventions and prevention strategies.
- *Improved patient experience*
Insights from data can inform service improvements, helping to reduce delays, improve access, and ensure care is delivered more effectively.

Benefits for organisations (GP practices, NHS bodies and researchers)

- *Facilitates joined up working across systems*
Bringing together data from multiple sources supports collaboration across primary care, secondary care, and other services.
- *Improves accuracy, completeness, and reliability of data*
Data processing, linkage, and validation within the SDE enhance data quality, leading to better decision-making.
- *Speeds up processes and access to data*
The SDE model provides secure, timely access to data for approved users, reducing delays associated with traditional data sharing methods.
- *Reduces administrative burden on staff*
Centralised, secure data access reduces the need for repeated data extractions and transfers, freeing up staff time for clinical and operational priorities.
- *Supports more effective use of resources*
Better insights enable organisations to plan services more efficiently and allocate resources where they are most needed.

Benefits for the wider public and health system

- *Improves population health and reduces inequalities*
Data analysis helps identify trends and disparities in health outcomes, supporting targeted interventions and more equitable service provision.
- *Enables innovation and research at scale*
Access to large, high-quality datasets supports development of new treatments, technologies, and evidence-based improvements in care.
- *Enhances system-wide planning and service design*
Data insights support commissioners and providers to ensure services are delivered in the right place, at the right time, for the population's needs.
- *Maintains public trust through secure data use*
The SDE approach ensures data is not widely shared but accessed securely, protecting confidentiality while delivering public benefit.

SECTION 3 – What data do you want to use or share?

4. Can you use anonymous data for your purposes? If not, explain why.

Put an ☒ next to the one that applies.

- ☐ Yes
☒ No
☐ Unsure

Anonymised data cannot be used for the core purposes of the Wessex SDE because it would not support the necessary processing activities required for research, data quality, and safe use of health and care data.

Key reasons include:

- **Data linkage requires identifiable information**
To create meaningful datasets, information from different sources (e.g. GP records, hospital data) needs to be linked at an individual level.
 - This requires identifiers such as NHS number, date of birth, or postcode.
 - If the data were fully anonymised, records could not be reliably matched across datasets, meaning patient pathways and outcomes could not be analysed.
- **Data quality, validation, and accuracy depend on identifiable data**
Before data can be used for research, it must be checked, cleaned, deduplicated, and standardised.

Identifiable data is required to:

- Detect and remove duplicate records
- Validate and correct errors

- Ensure consistent patient records across systems

Anonymous data would prevent these processes, leading to poorer quality datasets.

- **Application of opt-outs and legal obligations**

The SDE must apply national and local data opt-outs. This requires identifying individuals to:

- Remove those who have opted out
- Ensure legal compliance with confidentiality requirements

Anonymous data cannot support this, as individuals could not be identified and excluded appropriately.

- **Research requires longitudinal and individual-level analysis**

Many research purposes depend on understanding patient journeys over time. This includes:

- Tracking outcomes across services
- Studying disease progression
- Identifying risk factors and treatment effectiveness

Anonymous data (e.g. aggregated statistics only) does not allow this level of analysis.

5. Which types of personal data and deceased and deceased data do you need to use and why?

Put an ☒ next to all that apply.

- ☒ Forename
- ☒ Surname
- ☒ Address
- ☒ Postcode full
- ☒ Postcode partial
- ☒ Date of birth
- ☒ Date of death
- ☒ Age
- ☒ Gender
- ☒ Physical description, for example height
- ☐ Phone number
- ☐ Email address
- ☒ GP details
- ☐ Legal representative name (personal representative)
- ☒ NHS number
- ☐ National insurance number
- ☐ Other numerical identifier

- ☐ Photograph / picture of people
- ☐ IP address
- ☒ Other location data

Examples of 'other location data' would include data such as the location of a road traffic collision as recorded in the patient's notes, Ambulance, or RTC data.

- ☐ Audio recordings
- ☐ Video recordings
- ☐ Other
- ☐ None

The Wessex Secure Data Environment (SDE) requires the use of both **personal (identifiable)** and **special category (health) data** to support data linkage, research, and service improvement. The types of data used and the reasons for their inclusion are outlined below.

• **Identifiable personal data**

The following identifiers are required:

- Name (forename and surname)
- Address and postcode (full or partial)
- Date of birth / age
- Date of Death
- Gender
- NHS number
- GP/practice details

Why this is needed:

- To accurately link records across multiple datasets (e.g. GP, hospital, and national datasets) to create a complete patient pathway
- To ensure data quality and accuracy, including deduplication and validation of records
- To enable application of national, local, and GP opt-outs and meet legal and confidentiality requirements
- To support safe and reliable dataset preparation prior to pseudonymisation

These identifiers are only used within the secure Data Processing Environment (DPE) and are not accessible to researchers.

• **Health and care data (special category data)**

This includes:

- Clinical information (diagnoses, conditions, procedures)

- Medications and treatment details
- Laboratory and test results
- Vital signs and clinical measurements
- Patient records (including coded data and, where necessary, free text)

Why this is needed:

- To identify disease patterns, risk factors, and outcomes
- To support research into diagnosis, treatment, and prevention of disease
- To evaluate clinical pathways and service effectiveness
- To enable population health analysis and service planning

- **Additional sensitive data (where required and proportionate)**

In some cases, limited use of additional information may be required, such as:

- Ethnicity and demographic characteristics
- Genetic/genomic data (where relevant to a study)
- Other factors influencing health outcomes (e.g. religion where clinically relevant)

Why this is needed:

- To identify and address health inequalities and disparities
- To improve understanding of risk factors and treatment effectiveness across populations
- To support targeted interventions and personalised approaches

6. Data protection laws mean that some data is considered particularly sensitive. This is called special category data. Data that relates to criminal offences is also considered particularly sensitive. Which types of sensitive data do you need to use or share?

Put an ☒ next to all that apply.

- For all sensitive data, safeguards applied:
 - Sensitive data is used only where necessary and justified within each project
 - It is processed initially in a secure environment and then pseudonymised and minimised before researcher access
 - Access is controlled through robust governance (DARF, Pre-Check, DAC approval)
 - All outputs are fully anonymised before release in line with the Wessex SDE Safe People and Safe data SOP held internally and [SDE Network guidance](#).

- ☒ Information relating to an individual's physical or mental health or condition, for example information from health and care records

[Expand to see guidance notes](#)

Be specific where possible, for example diagnostic data, care plans, medication details, test results, vitals readings are needed in order to...

As above – see answer to Question 5.

Only agreed on a study-by-study basis, as required.

- ☐ Biometric information in order to uniquely identify an individual, for example facial recognition

N/A

- ☒ Genetic data, for example details about a DNA sample taken as part of a genetic clinical service

As above – see answer to Question 5.

Only agreed on a study-by-study basis, as required.

- ☒ Information relating to an individual's sexual life or sexual orientation

As above – see answer to Question 5.

Only agreed on a study-by-study basis, as required.

- Data is collected and processed except:
 - SNOMED Refset for 'General Practice summary data sharing exclusion for gender related issues' 999004371000000109
 - SNOMED Refset for 'General practice summary data sharing exclusion for assisted fertilisation' 999004351000000100
 - SNOMED Refset for 'General Practice summary data sharing exclusion for termination of pregnancy' 999004361000000107
 - All children of the SNOMED code 118199002 'Finding related to sexuality and sexual activity'

Use of the above data is strictly limited, proportionate, and subject to enhanced governance and approval.

- ☒ Racial or ethnic origin

As above – see answer to Question 5.

Only agreed on a study-by-study basis, as required.

- ☐ Political opinions
- N/A

- ☒ Religious or philosophical beliefs

Religious or philosophical belief data may be used only where it is relevant and necessary for a specific approved purpose, for example:

- To support individual care needs and preferences (e.g. dietary requirements, end-of-life care)

- To enable research into how beliefs may influence health behaviours, access to services, or outcomes
- To help identify and address health inequalities across different groups
- To understand potential impact of decisions on treatment and outcomes e.g. resistance to vaccines, blood transfusions, etc.

Use of this data is strictly limited, proportionate, and subject to enhanced governance and approval.

☐ Trade union membership
N/A

☒ Information relating to criminal or suspected criminal offences

Criminal offence data is vital in supporting individuals who face socioeconomic hardship and poorer physical and mental health outcomes driven by a cyclical relationship with poverty, trauma and institutional barriers.

- Information relating to criminal offences or legal status (e.g. safeguarding-related indicators or justice-linked data where lawfully available)

Only where relevant to an approved project, for example:

- Safeguarding and protecting vulnerable individuals
- Understanding wider determinants of health
- Supporting public health or service planning

Use of this data is strictly limited, proportionate, and subject to enhanced governance and approval.

☐ None of the above

a. Embed a description of the dataset if available, unless special category data is covered in your embedded description in response to question 5.

As above – see answer to Question 5.

Only agreed on a study-by-study basis, as required.

7. Who are the individuals that can be identified from the data?

Put an ☒ next to all that apply.

- ☒ Patients or service users
- ☐ Carers
- ☐ Staff
- ☐ Wider workforce
- ☐ Visitors
- ☐ Members of the public
- ☐ Other – please state

The data processed within the Wessex Secure Data Environment (SDE) relates to and may identify the following groups of individuals:

- **Patients registered with GP practices**
This includes individuals whose health and care data is recorded within primary care systems and forms the core dataset.
- **Individuals receiving NHS care or treatment**
Data may include patients who have accessed services across primary care, secondary care, community/mental health or other NHS-funded services, enabling analysis of care pathways.
- **Next of kin, carer and NHS staff data**
Any key stakeholders involved in the patient's health and care journey may be visible in the Data Processing Environment but none of this information will enter the Trusted Research Environment (TRE) and be available to researchers.

8. Where will your data come from?

The data used within the Wessex Secure Data Environment (SDE) is sourced from health and care organisations, primarily within NHS services, and is brought together to support approved research and analysis.

Primary data sources

- General Practice
The main source of data is electronic health record (EHR) held by practices, including coded clinical data and associated patient information using the standard extract procedure from the relevant GP clinical system supplier.

Currently, the suppliers are Optum (EMIS) and SystmOne TPP, but other suppliers introduced to market will be covered in future iterations of the DPIA.

Data Extraction

- The data will be extracted by a data extraction partner:
 - Dorset Intelligence and Insight Service (DiiS)
 - HIOW Cerner HealthIntent
 - Wessex SDE
 - Another SDE within the SDE network
 - Commercial third-party provider

This will **always be done with practice oversight** and be captured in the data sharing agreement. If the data extraction processor is another SDE or a third-party provider, they will have no access to the data, they will just be extracting the data encrypted from the GP clinical system supplier or population health management platform and be providing it (still encrypted) to the SDE.

9. Will you be linking any data together?

Put an ☐ next to the one that applies.

- ☒ Yes - provide an explanation below
- ☐ No - [go to question 10](#)
- ☐ Unsure - try to provide an explanation of what you think

Where required and approved, GP data may be linked with other datasets, including:

- Secondary care (hospital) data
Information relating to hospital admissions, outpatient activity, procedures, and treatments.
- Community and mental health services data
Data relating to care delivered in community settings or mental health services.
- National datasets (where relevant and approved)
This may include datasets such as NHS Digital collections or disease registries to provide a more complete view of patient pathways.
 - Hospital Episode Statistics (HES)
 - Emergency Care Data Set (ECDS)
 - Maternity Services Data Set (MSDS)
 - Secondary Uses Service (SUS)
 - National Cancer Registration and Analysis Service (NCRAS)
 - [Data sets - NHS England Digital](#)

a. Will it become possible, as a result of linking data, to be able to identify individuals who were not already identifiable from the original dataset?

Put an ☒ next to the one that applies.

- ☐ Yes
- ☒ No
- ☐ Unsure

Unlikely – while it is not possible to guarantee there will be zero risk of re-identification, the Wessex SDE follows all national and NHS standards to minimise this risk.

Risk of re-identification is actively managed.

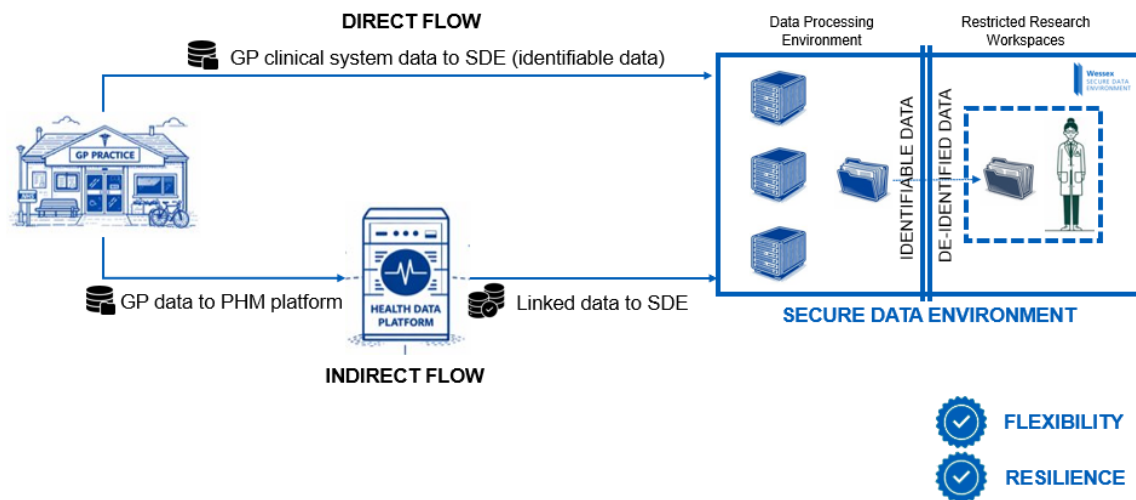
- **After linkage:**
 - Data is pseudonymised and minimised
 - Direct identifiers are removed before researcher access
 - Only pseudonymised datasets are made available within the TRE
 - No increased identifiability for researchers or outputs
Researchers cannot identify individuals, and no identifiable data is released externally; outputs are fully anonymised.
- **Governance**
 - Key risk mitigations are managed as part of the SDE's Live Risk Management process.

- Access to the Wessex SDE is governed by a formal End User Licence Agreement (EULA), which explicitly prohibits any attempt to re-identify individuals using the data made available within the environment. Users are required to adhere to strict data usage and confidentiality obligations, ensuring that all data is used solely for approved and lawful purposes. Any attempt to circumvent these controls or to re-identify patient data constitutes a breach of the agreed terms and is subject to formal disciplinary action and applicable sanctions in accordance with the EULA.
- A live register is maintained and managed by our IG and Governance Lead, and the SDE Board is updated on a monthly basis.
- Any additional risks identified on a study-by-study basis are triaged appropriately, risk assessed for likelihood, severity and impact and managed or escalated as appropriate.

The key risks in relation to GP data sharing and re-identification are captured in the risk table at the end of this document.

SECTION 4 – Where will data flow?

10. Describe the flows of data.



A more detailed version is available in [Appendix A](#).

Data flows within the Wessex Secure Data Environment (SDE) follow a structured, secure pathway from source systems through processing and analysis to controlled outputs.

Summary of data flow

- **Data collection from GP clinical systems (source systems)**
Data is entered by GP staff and other primary care users into GP clinical systems.
- **Data extraction**

For participating practices (with signed DSA/DPIA), data is extracted:

- Directly from GP systems (e.g. EMIS, TPP SystmOne), or
 - Via intermediary platforms (e.g. Dorset Intelligence & Insight Service – DiiS)
 - The SDE will always use the latest data available. If patients move between practices either within county or between counties, the data will be updated accordingly.
- **Transfer to SDE (DPE)**
 - Data is securely transferred into the Data Processing Environment (DPE) hosted by Wessex SDE (UHS).
 - **Data preparation within DPE**

Data undergoes:

 - Cleansing, validation, versioning and standardisation
 - Application of patient opt-outs (National, Type 1, local where applicable)
 - Pseudonymisation (e.g. hashing)
 - Data minimisation
 - Linkage with other datasets (where approved) – some cohort identification data may have to leave the SDE to other NHS trusts to support linkage, however this data will be identifier only and not contain any clinical information
 - **Transfer to TRE (analysis stage)**
 - Prepared, pseudonymised datasets are transferred into the Trusted Research Environment (TRE)
 - Approved researchers access and analyse de-identified data for their approved study only securely within the TRE
 - **Outputs and dissemination**
 - Only aggregated, anonymised outputs (with small number suppression) are allowed to leave the SDE following disclosure checks
 - **Data sharing with other SDEs (where approved)**
 - Pseudonymised datasets may be shared with, or received from, other SDEs for multi-region research (subject to approvals)
 - **Retention and disposal**
 - Data is retained in line with NHS Records Management Code of Practice for Health and Social Care
 - Project datasets are archived or destroyed after use

Example table of data flows

Data flow name	Source (from)	Destination (to)	Description
GP EHR	GP clinical systems	Wessex SDE (DPE at UHS)	Full GP care record transferred securely via API or approved extraction route (Airlock)

UHS – Wessex SDE	Wessex SDE Data Processing Environment (DPE)	Wessex SDE Trusted Research Environment (TRE)	Pseudonymised, minimised, research-ready datasets accessible in SDE TRE.
Research outputs	Wessex SDE TRE	Research organisations / external users	Aggregated, anonymised outputs (tables, graphs) with data minimisation and low number suppression applied in line with GDPR and Wessex SDE Safe People and Safe data SOP held internally.
Cross-SDE sharing	Wessex SDE	Other SDE environments	Pseudonymised datasets shared for approved nationally agreed research.

11. Confirm that your organisation’s information asset register (IAR), record of processing activities (ROPA) or your combined information assets and flows register (IAFR) has been updated with the flows described above.

Put an ☒ next to the one that applies.

- ☒ Yes
☐ No
☐ Unsure - add as a risk in section 10 with an action to find out

12. Will any data be shared outside of the UK?

Put an ☒ next to the one that applies.

- ☐ Yes
☒ No - [go to question 13](#)
☐ Unsure - add as a risk in [section 10](#) with an action to find out, then [go to question 13](#)

a. If yes, give details, including any safeguards or measures put in place to protect the data whilst outside of the UK

No GP data will be shared outside the UK.

The SDE operates a “data stays within the environment” model, so raw data is not exported or shared externally.

Fully anonymised aggregated outputs from the research studies that have been approved may be shared outside of the UK.

This will be managed in line with NHS standards and as per Wessex SDE Safe People and Safe data SOP held internally.

SECTION 5 – Is the intended use of the data lawful?

13. Under Article 6 of the UK General Data Protection Regulation (UK GDPR) what is your lawful basis for processing personal data?

Put an ☒ next to the one that applies.

- ☐ (a) **We have [consent](#)**
- ☐ (b) **We have a contractual obligation**
- ☐ (c) **We have a legal obligation**
- ☐ (d) **We need to protect a person's vital interests**
- ☒ (e) **We need it to perform a public task**
- ☐ (ea) **We have a recognised legitimate interest**

Put an ☒ next to the one that applies.

The recognised legitimate interest falls under:

- ☐ Sharing with another organisation for its public task or official functions
- ☐ National security, public security and defence
- ☐ Emergencies
- ☐ Detecting, investigating or preventing crime
- ☐ Safeguarding vulnerable individuals

- ☐ (f) **We have another legitimate interest**

14. If you have indicated in question 6 that you are using special category data, what is your lawful basis under Article 9 of the UK GDPR?

Put an ☒ next to the one that applies:

- ☐ (a) **We have [consent](#)**
- ☐ (b) **We need it for employment purposes**
- ☐ (c) **We need it to protect a person's vital interests**
- ☐ (d) **We need it for our not-for-profit organisation's activities, and it relates to our current or former members, or people who have regular contact with us**
- ☐ (e) **The person who the data is about has made that data public**
- ☐ (f) **We need it for legal claims, to seek legal advice or judicial acts**
- ☐ (g) **There is a [substantial public interest](#), as set out in [this list](#)**
- ☒ (h) **We need it to provide or manage health or social care services**
- ☒ (i) **We need it for public health purposes**
- ☒ (j) **We need it for archiving, research and statistics where this is in the public interest**
- ☐ **Not applicable**

15. What is your legal basis for using and sharing this health and care data under the common law duty of confidentiality?

Put an ☒ next to the one that applies.

- ☐ [Implied consent](#) - [go to question 16](#)
- ☐ [Explicit consent](#)
- ☒ **Section 251 support**
- ☐ **The law either requires or allows the data to be used and shared**
- ☐ **Overriding public interest**
- ☐ **Not applicable** - [go to question 16](#)

The Wessex SDE uses Section 251 permission specifically to link disparate health records from various locations.

Section 251 (NHS Act 2006)

Under common law, patient data is confidential. Section 251 gives the Secretary of State for Health and Social Care the legal power to bypass this requirement so confidential patient information can be used for medical research and SDE operations without explicit patient consent.

Confidentiality Advisory Group (CAG)

The Confidentiality Advisory Group (CAG) is an independent advisory body to the Health Research Authority (HRA) which reviews research applications and SDE data-linkage processes. The CAG ensures that utilising confidential patient information without consent is in the public interest and that robust data privacy and opt-out mechanisms are in place before granting SDE approval.

HRA (Health Research Authority)

The HRA relies on the CAG's expert recommendations to issue formal Section 251 approvals. Once the CAG reviews an SDE's process (such as its procedures for pseudonymization and security), the HRA provides the final green light required to operate and share these datasets securely.

Wessex Secure Data Environment (SDE) successfully obtained approvals from the NHS Health Research Authority (HRA), marking a significant step in ethical and legal compliance. The Research Ethics Committee granted ethical approval for the project. Whilst the HRA's Confidentiality Advisory Group has given the Wessex SDE permission to access confidential patient information without individual consent (this permission is known as a Section 251 notice).

Section 251 – Health Research Authority
CAG (IRAS ID: 334643)
REC (IRAS ID: 339034)

SECTION 6 – How are you keeping the data secure?

16. Are you storing information?

Put an ☒ next to the one that applies.

- ☒ Yes
☐ No - [go to question 17](#)

a. How will information be stored?

Put an ☒ next to all that apply.

- ☐ Physical storage, for example filing cabinets, archive rooms etc
☐ Local organisation servers
☒ Third party storage
☐ Other

Third party cloud storage partners, with approval from Health Research Authority – Research Ethics Committee and Clinical Advisory Group

Data is stored and transmitted in a highly secure, NHS-owned cloud-based digital lab, leveraging Microsoft Azure services for storage and analytics. Data is de-identified before entering the environment and does not leave this secure, ISO 27001-certified system during research.

17. Are you transferring information?

Put an ☒ next to the one that applies.

- ☒ Yes
☐ No - [go to question 18](#)

a. How will information be transferred?

Data is transmitted and stored in a highly secure, NHS-owned cloud-based digital lab, leveraging Microsoft Azure services for storage and analytics. Data is de-identified before entering the environment and does not leave this secure, ISO 27001-certified system during research.

Data may also be shared with other SDEs for research purposes where agreed and subject to their having the same security and governance controls in place.

18. How will you ensure that information is safe and secure?

Put an ☒ next to all that apply.

- ☒ Encryption

All data is encrypted at rest using AES-256 and in transit using TLS 1.2 or higher.

All data is protected through the application of **robust, industry-aligned encryption controls** both **at rest and in transit**, ensuring confidentiality and integrity throughout its lifecycle.

Data in transit is secured using **TLS 1.2 or higher**, with strong cryptographic standards (including SHA-256) to safeguard against interception and unauthorized access. Data at rest is encrypted using approved mechanisms consistent with organizational security and information governance requirements.

A **Privacy-Preserving Key Encryption System (PPKES)** is implemented to enforce strict access controls and data segregation. This ensures that each SDE retains exclusive control over the decryption of keys which are limited to approved use cases such as data linkage and cohort identification. Encryption keys are logically segregated and are not accessible across SDE boundaries, preventing unauthorized decryption by other SDEs.

These measures are supported by established governance controls, ensuring that access, key management, and data handling are tightly controlled, monitored, and restricted to authorized purposes only."

☒ User identity verification and authentication

MFA is in place for both researchers and engineers working within the Secure Data Environment. This is supported by a one-time password (OTP) sent via email, or Microsoft Authenticator application.

☒ Role based access controls (RBAC)

Access roles for users are allocated in line with Wessex SDE Safe People and Safe data SOP, ensuring consistency, fairness, and proper governance.

☐ Physical controls

N/A

☒ Security policies

The Wessex SDE is hosted on ISO 27001–certified platforms (ARO Business Communications Limited) and CIRU (Clinical Information Research Unit at UoS)), ensuring secure, governed, and externally verified backups. The certification covers internal security policies and physical security controls, with additional assurance provided through two ISO-compliant organisations and also relevant UHS Security Policies.

☒ Business continuity plans

The Wessex SDE is hosted across ISO 27001–certified platforms (ARO Business Communications Limited) and CIRU (Clinical Information Research Unit at UoS)), ensuring backup processes are securely managed, compliant with information governance requirements, and subject to external assurance.

☐ Other

N/A

19. How will you ensure the information will not be used for any other purposes beyond those set out in [question 2](#)?

Put an ☒ next to all that apply and provide details.

Governance and approvals

- Data Access Request Form (DARF):
Every project must clearly define its purpose, scope, and required data before access is granted.
- Pre-Check and Data Access Committee (DAC) approval:
All requests are independently reviewed to ensure the data use is appropriate, proportionate, and aligned with approved purposes.
- Project-specific access:
Data is only provided for the specific approved use case and cannot be reused for other purposes without further approval.

Access controls and environment restrictions

- Data is accessed only within the secure Trusted Research Environment (TRE)
- Users are granted role-based access and can only view data required for their approved study
- Researchers cannot download or transfer raw data outside the environment
- All activity is logged, monitored, and auditable

Data minimisation and separation

- Only the minimum necessary data is provided for each project
- Data is pseudonymised before access, ensuring researchers cannot identify individuals
- Different projects use separately pseudonymised datasets, preventing cross-project linkage or reuse – researchers will only have access to the datasets for their approved research study.

Legal and contractual controls

- Users must agree to data access agreements and confidentiality obligations
- There are clear sanctions and penalties for misuse of data
- Processing is carried out under defined legal bases (e.g. public task, research purposes)

Output controls

- All results must go through a formal disclosure control (“airlock”) process
- Only anonymised, aggregated outputs can leave the TRE
- This ensures no data can be repurposed or misused outside the approved project

Auditing and oversight

- Regular audit and compliance checks are carried out
- Any inappropriate use can be detected through activity monitoring and audit logs

- Wessex SDE maintains oversight of all approved projects

☒ Contract

Data Hosting:

- Microsoft Azure
- ARO Technology

☒ Data processing agreement

The SDE may hold data processing agreements as required to extract the data from the GP systems and transfer it to the DPE, as set out above.

These may be with the following:

- 1) In Dorset the Dorset intelligence and insight Service (DiiS)
- 2) In HIOW the HIOW ICS population health management (PHM) service
- 3) The Wessex SDE itself
- 4) Another SDE
- 5) A commercial third party

☒ Data sharing agreement

Individual GP practices will be asked to sign a DSA on behalf of their organisation

☐ [Data sharing and processing agreement \(DSPA\)](#)

N/A

☒ Audit

The SDE is subject to regular audit by

- ISO
- Health Research England

☒ Staff training

The training and authorisation of the data analysts that carry out these procedures, as well as the SOPs themselves are covered under the SDE's ISO27001 certification.

Both the researcher's organisation and the individual researchers must complete the necessary authorisation in order to gain access and are granted access only to data which is required for their individual study.

☐ Other

N/A

SECTION 7 – How long are you keeping the data and what will happen to it after that time?

20. How long are you planning to use the data for?

Once the research has finished, the dataset (and any copies) will be archived and destroyed according to [Wessex SDE policy](#). The data will be retained for as long as the purpose(s) described above remains valid or a new legal purpose agreed. This will be in line with the: [NHS Records Management Code of Practice 2021](#) and updates 2023 A guide to the management of health and care records.

21. How long do you intend to keep the data?

Once the research has finished, the dataset (and any copies) will be archived and destroyed according to [Wessex SDE policy](#). The data will be retained for as long as the purpose(s) described above remains valid or a new legal purpose agreed. This will be in line with the: [NHS Records Management Code of Practice 2021](#) and updates 2023 A guide to the management of health and care records.

22. What will happen to the data at the end of this period?

Put an ☐ next to all that apply.

☒ Secure destruction

Wiping hard drives with certificate of sanitisation - All deleted in line with [Wessex SDE policy](#).

☐ Long term preservation by transferring the data to a National Archives' approved Place of Deposit

N/A

☐ Transfer to another organisation

N/A

☐ Extension to retention period

N/A

☐ It will be anonymised and kept

N/A

☐ The controller(s) will manage as it is held by them

N/A

☐ Other

N/A

SECTION 8 – How are people’s rights and choices being met?

23. How will you comply with the following individual rights (where they apply)?

The right to be informed

The right to be informed about the collection and use of personal data.

Put an ☒ next to all that apply.

☒ Privacy notice(s) for all relevant organisations

The Wessex SDE’s privacy notice can be found on our website [here](#).

☒ Information leaflets

☒ Posters

☐ Letters

☐ Emails

☐ Texts

☒ Social media campaign

☐ DPIA published (best practice rather than requirement)

☐ Other

☐ Not applicable

a. Describe how the communication methods above meet your obligation to comply with the right to be informed.

The Wessex SDE has a continuous programme of Communications and Public and Patient Engagement and Involvement (PPIE) to ensure that members of the public are aware of how their data is being used for research. This includes clearly set transparency and publication commitments within our [Transparency policy](#).

All GP surgeries contributing data are provided with a Public Notification Pack as part of their onboarding process which includes posters to be put up in publicly accessible areas. GPs agree to making this information available to their patients when they sign the DSA and DPIA as part of our commitments to the HRA.

The right of access

The right to access details of data use and receive a copy of their personal information - this is commonly referred to as a subject access request.

Right of Access (Data Subject Access Requests) are publicised via our website, managed through our 'Incoming Communications and Enquiries SOP' and implemented under the guidance of UHS’s Information Governance Team as the host of the Wessex SDE.

The right to rectification

The right to have inaccurate personal data rectified or completed if it is incomplete.

N/A

The right to erasure

The right to have personal data erased, if applicable.

N/A

The right to restrict processing

The right to limit how their data is used, if applicable.

N/A

The right to data portability

The right to obtain and re-use their personal data, if applicable.

N/A

The right to object

The right to object to the use and sharing of personal data, if applicable.

Patients/Public

- Should patients object to the use and sharing of personal data they can either:
 - Opt out of sharing their data via the National, local or GP opt out routes as publicised on our website – [Your choices](#)
 - Contact the SDE directly regarding a specific concern or objection via our [contact form](#) or emailing WessexSDE@uhs.nhs.uk.

Note: any objections raised would be shared with the HRA as part of our Annual Update

Practices

- Formal notification should be received in 2-week window after approval of study by DAC:
 - Practices must formally notify Wessex SDE of their decision to opt out.
- Agreement and governance route:
 - Local opt-out is managed through agreed governance arrangements.
- Handling of existing data:
 - Any data already within the SDE will be handled in line with:
Existing legal basis and governance approvals
Agreed retention and deletion policies

Deceased patients

The Wessex SDE CAG approval includes the authorisation to hold data for patients that have died, including information about date of death. This data is critically important to our researchers. As a result, Wessex SDE will not be routinely destroying data relating to these patients.

Although UK GDPR applies only to living individuals, confidentiality obligations and common law duty of confidentiality still apply after death. The SDE will continue to

handle a deceased patient's data carefully, lawfully, and ethically. Any previous data opt-outs will still be strictly adhered to.

The SDE will respond to requests from next of kin and implement a local data opt out if requested.

Out of area

The Wessex SDE CAG approval includes data for every patient whose care is, or has been, delivered by NHS organisations within the locality unless they have chosen that their data should not be used through a type-1, local or national data opt-out which will be strictly adhered to.

Any previously collected data prior to the patient moving out of area will be retained and used for research analysis.

24. Will the national data opt-out need to be applied?

Put an ☒ next to the one that applies.

☒ Yes

The following opt-outs will be applied within Wessex SDE:

- ✓ National Data opt out
- ✓ Type 1 opt-outs
- ✓ A local Wessex SDE Opt-Out – by patient or practice

Opt-outs are tracked daily and recorded in a Master Patient Index held by Wessex SDE and only accessible by approved members of our Technical Team. Opt-outs are applied to the data as close as possible to pseudonymisation within the TRE before sharing de-identified data with approved researchers. This is to ensure that all opt-outs have been practically applied to the data prior to being used for a study.

☐ No

N/A

☐ Unsure - add as a risk in [section 10](#) with an action to find out

25. Will any decisions be made in a purely automated way without any meaningful human involvement (automated decision-making)?

Put an ☒ next to the one that applies.

☐ Yes

☒ No - [go to question 26](#)

☐ Unsure - add as a risk in [section 10](#) with an action to find out

a. Please provide an explanation of how the decisions are made and who it impacts.

N/A

b. How will the quality of the outputs and decisions be monitored and recorded?

N/A

c. Will the effect of the decisions on the individual be [significant](#)?

Put an ☒ next to the one that applies.

☐ Yes

☒ No – go to [question 26](#)

d. How will you provide people with information about the decision?

N/A

e. How can people express their views and contest the decision?

N/A

f. How can people ask for a human intervention in the decision?

N/A

g. Confirm whether the decision is:

Put an ☒ next to the one that applies.

☐ Based on the individual's consent

☐ Needed for a contract with the individual

☐ Required or authorised by law (public task, legal obligation)

☒ None of the above – add as a risk in [section 10](#) with an action to address

h. Are you using any special category data as part of automated decision making?

Put an ☒ next to the one that applies.

☐ Yes

☒ No – go to [question 26](#)

i. Confirm which of the following conditions apply to the decision:

Put an ☒ next to the one that applies.

☐ The individual has explicitly consented (can only apply if you have selected consent for question 25e)

☐ Substantial public interest is our Article 9 lawful basis

☐ None of the above – add as a risk in [section 10](#) with an action to address

N/A

26. Detail any stakeholder consultation that has taken place (if applicable).

A range of stakeholder engagement and consultation activities have been undertaken to ensure transparency, accountability, and that the use of data within the Wessex Secure Data Environment (SDE) reflects the needs and expectations of patients, the public, and partner organisations.

General Practice (GP) stakeholders

- Engagement with GP practices, including clinicians and practice managers, to:
 - Explain the purpose and operation of the SDE
 - Seek input on data flows and operational processes
 - Address concerns around workload, governance, and patient confidentiality

NHS organisations and partner bodies

- Collaboration with NHS trusts, Integrated Care Systems (ICS), and other health and care organisations to:
 - Align on governance arrangements and data sharing approaches
 - Ensure consistency with wider NHS data strategies
 - Support joined up working across systems

Information governance and legal teams

- Consultation with IG professionals and legal advisors to:
 - Ensure compliance with UK GDPR, Data Protection Act 2018, and common law duty of confidentiality
 - Develop appropriate policies, processes, and safeguards

Public and patient involvement and engagement (PPIE)

- Engagement with patients and the public through:
 - Public engagement events and workshops
 - Patient representative groups and advisory panels
 - Feedback mechanisms to understand public attitudes to data use

This helps ensure that data use is transparent, acceptable, and in line with public expectations.

Ongoing engagement and review

- Stakeholder consultation is ongoing, with feedback used to refine processes and communications
- Mechanisms are in place to:
 - Keep stakeholders informed of developments
 - Respond to concerns or feedback
 - Update materials (e.g. privacy notices, communications) as needed

SECTION 9 – Which organisations are involved?

27. List the organisation(s) that will decide why and how the data is being used and shared (controllers).

GP surgeries will be the Data Controllers up to the point at which data enters the Wessex SDE.

In contributing data, they transfer Data Controller responsibility to UHS for ensuring that any data shared for research is accurate, up-to-date, minimised, and lawfully disclosed under the UK GDPR, Data Protection Act 2018, and the Common Law Duty of Confidentiality.

Access to data is controlled through the Wessex Data Access Process and Data Access Committee (DAC), made up of public and professional members.

Following approval from the DAC, GPs will be notified that a new study is being undertaken. GPs will have up to 10 working days to notify the Wessex SDE if they wish for their data to be excluded from the study.

28. List the organisation(s) that are being instructed to use or share the data (processors).

The organisations whose data is being requested for sharing in relation to this DPIA are:

- All GP surgeries across Hampshire, Isle of Wight and Dorset ICBs, including their system providers (System One, EMIS)
- Dorset ICB, in relation to Dorset GP data, approved by the GP surgery for sharing that is currently captured within the Dorset DiiS system

Data Processors for the Wessex SDE

- University of Southampton (UoS), Clinical Informatics and Research Unit (CIRU);
- BC Platforms AG;
- ARO
- Microsoft Azure

29. List any organisations that have been subcontracted by your processor to handle data

N/A

30. List any organisation involved but not covered above. For example, a commissioner.

N/A

31. Explain the relationship between the organisations set out in questions [27](#), [28](#), [29](#), and [30](#) and any other relevant organisation. Explain what activities they do

University Hospital Southampton (UHS) – Data Controller; As the host organisation for the Wessex SDE, they have determined the scope of research the SDE will support (the ‘purposes’ for processing) and decided on the detail of the system(s), technical controls and overall data management (the ‘means’ of processing).

GP surgeries – Data Contributors; approved data-holding organisation, that supplies data to the Wessex SDE for the purposes of approved research, analysis, or service improvement.

In contributing data, they share Data Controller responsibility with UHS for ensuring that any data shared for research is accurate, up-to-date, minimised, and lawfully disclosed under the UK GDPR, Data Protection Act 2018, and the Common Law Duty of Confidentiality.

GP EPR providers (System One, EMIS) – Data Processors; sending data on behalf of the GP surgeries

BCP and ARO – Data Processors; receiving data on behalf of the SDE

CIRU – Data Processors; managing the data inside the SDE as the SDE’s core partners and Technical Services.

An Accountability Matrix to outline the above information is provided in Appendix B.

32. What due diligence measures and checks have been carried out on any processors used?

Put an ☒ next to all that apply.

☒ **Data Security and Protection Toolkit (DSPT) compliance**

UHS: DSPT: 24/25 Approaching Standards (ODS RHM)

UoS - DSPT: 24/25 Standards Met (ODS: EE133879-CIRU)

ARO - DSPT: 24/25 Standards Exceeded (ODS 8J121)

☒ **Registered with the Information Commissioner’s Office (ICO)**

UHS - ICO reg.: Z4989884, expiry 27th November 2026

UoS - ICO reg.: Z6801020, expiry 20th June 2026

BCP - ICO reg.: ZB928057, expiry 1st July 2026

ARO - ICO reg.: Z9552844, expiry 27th July 2026

☐ **Digital Technology Assessment Criteria (DTAC) assessment**

N/A

☒ **Stated accreditations**

UoS(CIRU) is ISO 270001 certified

ARO (Arrow Business Communications Limited) is ISO 270001 certified

☒ **Cyber Essentials or any other cyber security certification**

University of Southampton – Clinical Informatics Research Unit (CIRU) - Cyber Essentials: CE, expiry 6th November 2026 (ID: f24f8bb1-ad11-47e9-af42-1652c8d0bde8)

ARO - Cyber Essentials: CE+, expiry 22nd October 2026 (ID: 1b2adc2c-4104-4659-93b5-35908194c9b4)

☐ **Other checks**

SECTION 10 – What data protections are there and what mitigations will you put in place?

33. Complete the [risk assessment table](#). Use the [risk scoring table](#) to decide on the risk score.

Risk assessment table

Risk ref no.	Description	Likelihood (L) score	Impact score (I)	Risk score* (L x I)	Mitigations	Risk score* with mitigations applied
1	Data is being transferred out of the UK.	3	3	9	No transfer of data outside the UK. Assurance obtained that the servers are in the UK.	$1 * 3 = 3$
2	Privacy notice is not in place or is incorrect.	3	2	6	Ensure Privacy Notice reflects the programme. Transparency obligations on partners documented in DSA. Communications pack available and used to assure compliance with transparency obligations under the UK GDPR.	$1 * 2 = 2$
3	Post go live risks are not monitored actively.	3	3	9	Post go-live risks relating to NFRs and all other relevant risks are actively monitored, and any exceptions are appropriately escalated/reported. Completion and sign off, of all relevant Data Protection documentation including this DPIA and the relevant Data Sharing Agreements.	$2 * 3 = 6$
4	Processor may choose to use an unaccredited system and therefore risk introducing security weaknesses.	3	3	9	Processor contracts mitigate this risk by requiring all processors to implement and evidence appropriate technical and organisational security measures, obtain prior approval before introducing new or unaccredited systems, and remain subject to audit, assurance, and remedial action where security standards are not met.	$1 * 3 = 3$
5	Patient data is not linked properly across different systems.	5	3	15	SOP in place with effective checks to ensure that any data that needs to be linked between or across systems relating to a single patient is robust. This is	$2 * 3 = 6$

					done prior to researchers being given access to a particular workspace.	
6	System security is compromised.	3	5	15	CIRU have SOPs and Policies in place to manage the infrastructure, including the system security. These are followed and reviewed internally. In addition, they have recently passed an ISO 27001 external audit.	$1 * 5 = 5$
7	Staff do not understand the importance of data security and confidentiality with regards to patient identifiable data.	3	3	9	All Trust R&D staff are mandated to do IG training on an annual basis via an online system (including audit trail and notifications). All staff with a UHS honorary contract (i.e. contractors) should receive a recommended reading list - the trigger for this is when their honorary contract is issued. 3rd party contractors/consultants currently working with SDE do not have access to clinical/patient data, or access to systems providing this data. CIRU staff undergo regular mandatory training in Data Protection (every 1-2 years depending on role) and Cyber Security (annual) which is managed centrally by the University. The SDE has implemented an End User Licence Agreement which contains clauses regarding data protection.	$1 * 3 = 3$
8	Electronic transfer of data is not controlled effectively.	3	3	9	SOPs in place, service level agreement in place, regular testing in place.	$1 * 3 = 3$
9	Data is destroyed with no evidence.	3	3	9	SOP in place which details how data is to be archived and destroyed in line with UHS policy.	$1 * 3 = 3$

10	National and local Data Opt-Out is not applied to data packages ahead of sharing with researchers.	5	3	15	SOP in place as to how to prepare data packages, reporting to DPO as required.	$2 * 3 = 6$
11	Identifiable information is being made available to researchers in error.	2	4	8	Most studies will include structured data. If data is unstructured the use of Pterodactyl to redact the identifiable information in unstructured test fields/documents will be used or NLP tools can be used by the provider organisation on identifiable data to then extract structured facts to pseudonymise and send to the SDE.	$1 * 4 = 4$

***Risk scoring table**

		Impact (I)				
		Negligible (1)	Low (2)	Moderate (3)	Significant (4)	Catastrophic (5)
Likelihood (L)	Rare (1)	1	2	3	4	5
	Unlikely (2)	2	4	6	8	10
	Possible (3)	3	6	9	12	15
	Likely (4)	4	8	12	16	20
	Almost certain (5)	5	10	15	20	25

34. Detail any actions needed to mitigate any risks, who has approved the action, who owns the action, when it is due and whether it is complete.

Risk ref no.	Action needed	Action approver	Action owner	Due date	Status for example, outstanding / complete

SECTION 11 – Review and sign-off

Reviewer sign off

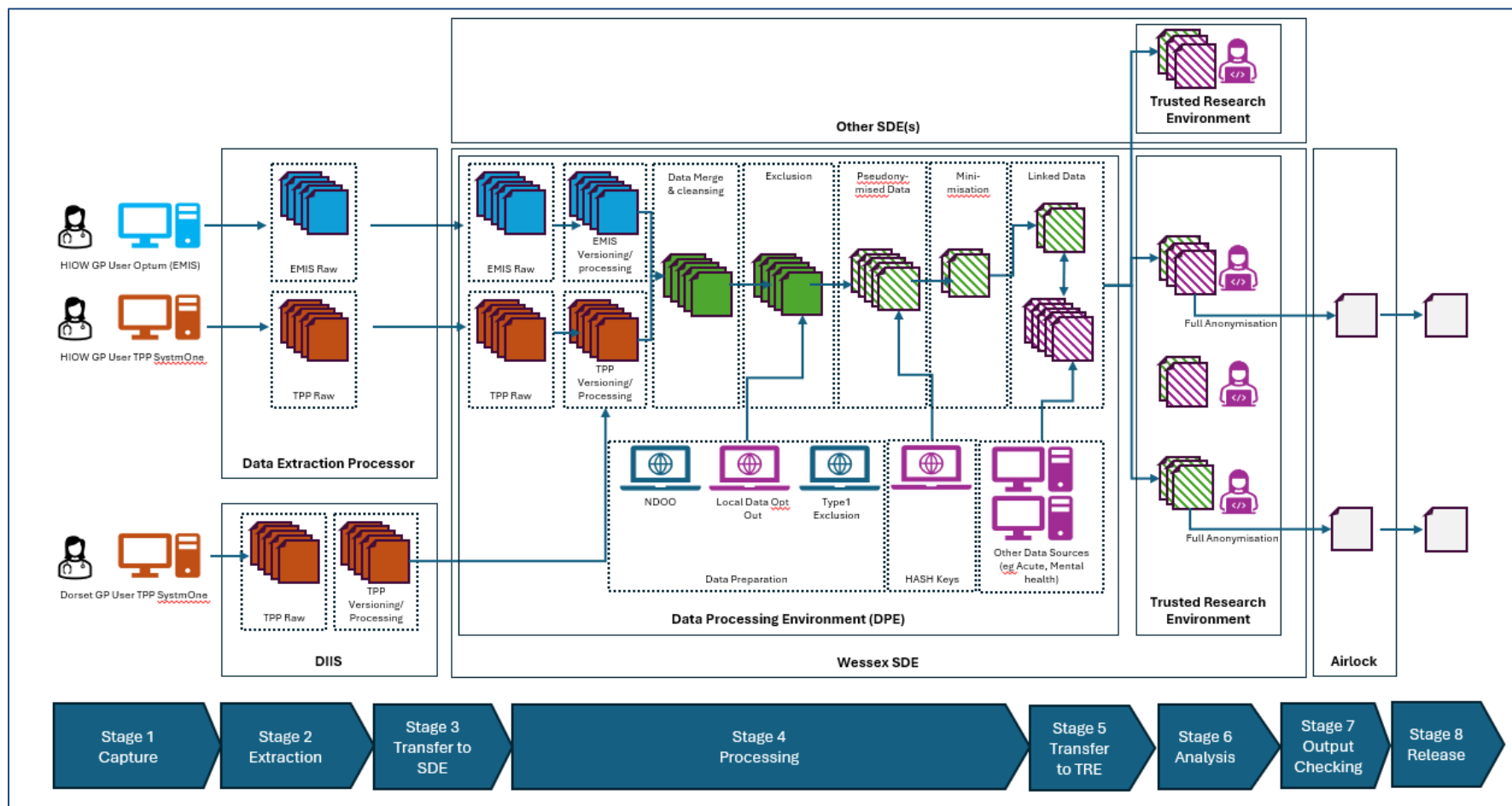
Reviewer name:	Emile A Armour
Reviewer job title:	Head of Information Governance & Data Protection Officer
Reviewer contact details:	Dataprotection@uhs.nhs.uk
Date of review:	13.07.2026
Comments:	None
Date for next review:	Any significant changes to processing or 2029.

Approver sign-off

Approver name:	Emile A Armour
Approver job title:	Head of IG & DPO
Approver contact details:	Dataprotection@uhs.nhs.uk
Date of approval:	14.07.2026
Signature:	<i>EA Armour</i>

SECTION 12 – Appendices

Appendix A - GP Data Flow



Appendix B – Accountability Matrix

	Stage 1 & 2 Capture & Extraction	Stage 3 Transfer to SDE	Stage 4 Data Processing	Stage 5 Transfer to TRE	Stage 6 Analysis	Stage 7 Output Checking	Stage 8 Release
Accountable (Controller)	GP Practices (Primary controller)	Transition point → UHS (Wessex SDE host becomes controller for processing)	UHS (Wessex SDE)	UHS (Wessex SDE)	- UHS (as SDE controller) - Research Sponsor (for study-level governance)	UHS (Wessex SDE)	- UHS (initial release decision) - Receiving organisation becomes controller for outputs
Processor	- GP system providers (EMIS, TPP) - Extraction Partners	ARO / BCP / Azure / CIRU (hosting and transfer infrastructure)	Platform providers	TRE Platform provider (CIRU)	TRE infrastructure provider	CIRU	CIRU
Risk Owner	- GP Practice Caldicott Guardian (primary) - Supported by GP Practice Data Controller / Practice Lead - Extraction Partners	- UHS Wessex SDE Technical Lead - Supported by Extraction Partner	- Information Asset Owner (IAO) – Wessex SDE GP Data Asset - Supported by SDE IG Lead / UHS Wessex SDE Technical Lead	UHS Wessex SDE Data Lead	- Data Access Committee (DAC) Chair (governance accountability) - Research Sponsor / Chief Investigator (use accountability)	CIRU	- SDE Governance Lead (pre-release) - Receiving Organisation Data Controller (post-release)
Key Risks	- Data accuracy - Lawful basis and transparency - Opt-out capture	- Secure transmission (encryption, integrity) - Unauthorised access during transfer - Data integrity in transit	- Incorrect linkage - Failure to apply opt-outs	- Dataset minimisation failure - Inappropriate data exposure	- Misuse of data - Attempted re-identification risk - Scope creep	- Residual disclosure risk - Small number identification	- Misinterpretation or misuse of outputs - Secondary disclosure risk